Quasi-periodic standing wave solutions of gravity-capillary water waves

Massimiliano Berti, Riccardo Montalto

Abstract. We prove the existence and the linear stability of Cantor families of small amplitude time *quasi-periodic* standing wave solutions (i.e. periodic and even in the space variable x) of a 2-dimensional ocean with infinite depth under the action of gravity and surface tension.

Keywords: KAM for PDEs, Water waves, quasi-periodic solutions, standing waves.

MSC 2010: 76B15, 37K55, 76D45 (37K50, 35S05).

Contents

1 Introduction and main result

In this paper we prove the existence of non trivial, small amplitude, *quasi-periodic* in time, *linearly stable* gravity-capillary standing water waves of a 2-d perfect, incompressible, irrotational fluid with infinite depth, under periodic boundary conditions, and which occupies the free boundary region

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

More precisely we find quasi-periodic in time solutions of the system

$$\begin{cases} \partial_t \Phi + \frac{1}{2}|\nabla \Phi|^2 + g\eta = \kappa \frac{\eta_{xx}}{(1+\eta_x^2)^{3/2}} & \text{at } y = \eta(x) \\ \Delta \Phi = 0 & \text{in } \mathcal{D}_\eta \\ \nabla \Phi \to 0 & \text{as } y \to -\infty \\ \partial_t \eta = \partial_y \Phi - \partial_x \eta \cdot \partial_x \Phi & \text{at } y = \eta(x) \end{cases} \tag{1.1}$$

where g is the acceleration of gravity, $\kappa \in [\kappa_1, \kappa_2]$, $\kappa_1 > 0$, is the surface tension coefficient and

$$\frac{\eta_{xx}}{(1+\eta_x^2)^{3/2}} = \partial_x \left(\frac{\eta_x}{\sqrt{1+\eta_x^2}} \right)$$

is the mean curvature of the free surface. The unknowns of the problem are the free surface $y = \eta(x)$ and the velocity potential $\Phi : \mathcal{D}_\eta \to \mathbb{R}$, i.e. the irrotational velocity field $v = \nabla_{x,y}\Phi$ of the fluid. The first equation in (1.1) is the Bernoulli condition according to which the jump of pressure across the free surface is proportional to the mean curvature. The last equation in (1.1) expresses that the velocity of the free surface coincides with the one of the fluid particles.

In the sequel we shall assume (with no loss of generality) that the gravity constant $g = 1$.

Following Zakharov [44] and Craig-Sulem [20], the evolution problem (1.1) may be written as an infinite dimensional Hamiltonian system. At each time $t \in \mathbb{R}$ the profile $\eta(t,x)$ of the fluid and the value

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

of the velocity potential Φ restricted to the free boundary uniquely determine the velocity potential Φ in the whole $\mathcal{D}_\eta$, solving (at each t) the elliptic problem (see e.g. [2], [33])

$$\Delta\Phi = 0 \quad \text{in } \mathcal{D}_\eta, \quad \Phi(x+2\pi, y) = \Phi(x,y) \, , \quad \Phi|_{y=\eta} = \psi, \quad \nabla\Phi(x,y) \to 0 \text{ as } y \to -\infty \, . \tag{1.2}$$

As proved in [44], [20], system (1.1) is then equivalent to the system

$$\begin{cases} \partial_t \eta = G(\eta)\psi, \\ \partial_t \psi + \eta + \dfrac{1}{2}\psi_x^2 - \dfrac{1}{2}\dfrac{\big(G(\eta)\psi + \eta_x \psi_x\big)^2}{1+\eta_x^2} = \kappa \dfrac{\eta_{xx}}{(1+\eta_x^2)^{3/2}} \end{cases} \tag{1.3}$$

where $G(\eta)$ is the so-called Dirichlet–Neumann operator defined by

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

(we denote by η_x the space derivative $\partial_x \eta$.) The operator $G(\eta)$ is linear in ψ, self-adjoint with respect to the L^2 scalar product and semi positive definite, actually its Kernel are only the constants. It is well known since Calderon that the Dirichlet-Neumann operator is a *pseudo-differential* operator with principal symbol $|D|$, actually $G(\eta) - |D| \in OPS^{-\infty}$, see section 2.4.

Furthermore the equations (1.3) are the Hamiltonian system (see [44], [20])

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

$$\partial_t u = J \nabla_u H(u)\,, \quad u := \begin{pmatrix} \eta \\ \psi \end{pmatrix}\,, \quad J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}\,, \tag{1.5}$$

where ∇ denotes the L^2-gradient, and the Hamiltonian

$$H(\eta, \psi) := \frac{1}{2}(\psi, G(\eta)\psi)_{L^2(\mathbb{T}_x)} + \int_{\mathbb{T}} \frac{\eta^2}{2}\, dx + \kappa \int_{\mathbb{T}} \sqrt{1 + \eta_x^2}\, dx \tag{1.6}$$

is the sum of the kinetic energy

$$K := \frac{1}{2}(\psi, G(\eta)\psi)_{L^2(\mathbb{T}_x)} = \frac{1}{2}\int_{\mathcal{D}_\eta} |\nabla \Phi|^2(x, y) dx dy\,,$$

the potential energy and the energy of the capillary forces (area surface integral) expressed in terms of the variables (η, ψ).

The symplectic structure induced by (1.5) is the standard Darboux 2-form

$$\mathcal{W}(u_1, u_2) := (u_1, J u_2)_{L^2(\mathbb{T}_x)} = (\eta_1, \psi_2)_{L^2(\mathbb{T}_x)} - (\psi_1, \eta_2)_{L^2(\mathbb{T}_x)} \tag{1.7}$$

for all $u_1 = (\eta_1, \psi_1)$, $u_2 = (\eta_2, \psi_2)$.

The water-waves system (1.3)-(1.5) exhibits several symmetries. First of all, the mass $\int_{\mathbb{T}} \eta\, dx$ is a prime integral of (1.3). Moreover

$$\partial_t \int_{\mathbb{T}} \psi\, dx = -\int_{\mathbb{T}} \eta\, dx - \int_{\mathbb{T}} \nabla_\eta K\, dx = -\int_{\mathbb{T}} \eta\, dx$$

because $\int_{\mathbb{T}} \nabla_\eta K\, dx = 0$. This follows because $\mathbb{R} \ni c \mapsto K(c + \eta, \psi)$ is constant (the bottom of the ocean is at $-\infty$) and so $0 = d_\eta K(\eta, \psi)[1] = (\nabla_\eta K, 1)_{L^2(\mathbb{T})}$. As a consequence the subspace

$$\int_{\mathbb{T}} \eta\, dx = \int_{\mathbb{T}} \psi\, dx = 0 \tag{1.8}$$

is invariant under the evolution of (1.3) and we shall restrict to solutions satisfying (1.8).

In addition, the subspace of functions which are even in x,

$$\eta(x) = \eta(-x)\,, \quad \psi(x) = \psi(-x)\,, \tag{1.9}$$

is invariant under (1.3). Thanks to this property and (1.8), we shall restrict (η, ψ) to the phase space of 2π-periodic functions which admit the Fourier expansion

$$\eta(x) = \sum\nolimits_{j \geq 1} \eta_j \cos(jx)\,, \quad \psi(x) = \sum\nolimits_{j \geq 1} \psi_j \cos(jx)\,. \tag{1.10}$$

In this case also the velocity potential $\Phi(x, y)$ is even and 2π-periodic in x and so the x-component of the velocity field $v = (\Phi_x, \Phi_y)$ vanishes at $x = k\pi$, $\forall k \in \mathbb{Z}$. Hence there is no flux of fluid through the lines $x = k\pi$, $k \in \mathbb{Z}$, and a solution of (1.3) satisfying (1.10) describes the motion of a liquid confined between two walls.

Another important symmetry of the capillary water waves system is reversibility, namely the equations (1.3)-(1.5) are reversible with respect to the involution $\rho : (\eta, \psi) \mapsto (\eta, -\psi)$, or, equivalently, the Hamiltonian is even in ψ:

$$H \circ \rho = H\,, \quad H(\eta, \psi) = H(\eta, -\psi)\,, \quad \rho : (\eta, \psi) \mapsto (\eta, -\psi)\,. \tag{1.11}$$

As a consequence it is natural to look for solutions of (1.3) satisfying

$$u(-t) = \rho u(t)\,, \quad i.e. \quad \eta(-t,x) = \eta(t,x)\,,\ \psi(-t,x) = -\psi(t,x)\,,\ \forall t,x \in \mathbb{R}\,, \tag{1.12}$$

namely η is even in time and ψ is odd in time. Solutions of the water waves equations (1.3) satisfying (1.10) and (1.12) are called gravity-capillary *standing water waves.*

This is a small divisors problem. Existence of small amplitude time periodic pure gravity (without surface tension) standing wave solutions has been proved by Iooss, Plotnikov, Toland in [32], see also [28], [29], and in [39] in finite depth. Existence of time periodic gravity-capillary standing wave solutions has been recently proved by Alazard-Baldi [1]. The above results are proved via a Lyapunov Schmidt decomposition combined with a Nash-Moser iterative scheme.

In this paper we extend the latter result proving the existence of time *quasi-periodic* gravity-capillary standing wave solutions of (1.3), see Theorem 1.1, as well as their linear stability. The reducibility of the linearized equations at the quasi-periodic solutions is not only an interesting dynamical information but it is also the key for the existence proof in Theorem 1.1.

We also mention that existence of small amplitude 2-d travelling gravity water wave solutions dates back to Levi-Civita [34] (standing waves are not traveling because they are even in space, see (1.9)). Existence of small amplitude 3-d traveling gravity-capillary water wave solutions with space periodic boundary conditions has been proved by Craig-Nicholls [19] (it is not a small divisor problem) and by Iooss-Plotinikov [30]-[31] in the case of zero surface tension (in such a case it is a small divisor problem).

Existence of quasi-periodic solutions of PDEs (that we shall call in a broad sense KAM theory) with unbounded perturbations (i.e. the nonlinearity contains derivatives) has been developed by Kuksin [37] for KdV, see also Kappeler-Pöschel [36], by Liu-Yuan [35], Zhang-Gao-Yuan [45] for derivative NLS, by Berti-Biasco-Procesi [12]-[13] for derivative NLW. All these previous results still refer to semilinear perturbations, i.e. the order of the derivatives in the nonlinearity is strictly lower than the order of the constant coefficient (integrable) linear differential operator.

For quasi-linear (either fully nonlinear) nonlinearities the first KAM results have been recently proved by Baldi-Berti-Montalto in [7], [8], [9] for perturbations of Airy, KdV and mKdV equations. These techniques have been extended by Feola-Procesi [26] for quasi-linear perturbations of Schrödinger equations.

The gravity-capillary water waves system (1.3) is indeed a quasi-linear PDE. In suitable complex coordinates it can be written in the symmetric form $\mathtt{u}_t = \mathrm{i}T(D)\mathtt{u} + N(\mathtt{u},\bar{\mathtt{u}})$, $\mathtt{u} \in \mathbb{C}$, where $T(D) := |D|^{1/2}(1-\kappa\partial_{xx})^{1/2}$ is the Fourier multiplier which describes the linear dispersion relation of the water waves equations linearized at $(\eta,\psi) = 0$ (see (1.13)-(1.17)), and the nonlinearity $N(\mathtt{u},\bar{\mathtt{u}})$ depends on the highest order term $|D|^{3/2}\mathtt{u}$ as well, see sections (6.1)-(6.2) for the complex form of the linearized system.

We have not the space to report the huge literature concerning KAM theory for semilinear PDEs in one and also higher space dimension, for which we refer to [37], [18], [24], [16], [17].

Let us present rigorously our main result. As already said we look for small amplitude quasi-periodic solutions of (1.3). It is therefore of main importance the dynamics of the system obtained linearizing (1.3) at the equilibrium $(\eta,\psi) = (0,0)$ (flat ocean and fluid at rest), namely

$$\begin{cases} \partial_t \eta = G(0)\psi, \\ \partial_t \psi + \eta = \kappa \eta_{xx} \end{cases} \tag{1.13}$$

where $G(0) = |D_x|$ is the Dirichlet-Neumann operator for the flat surface $\eta = 0$, namely

$$|D_x|\cos(jx) = |j|\cos(jx), \quad |D_x|\sin(jx) = |j|\sin(jx)\,,\ \forall j \in \mathbb{Z}\,.$$

In compact Hamiltonian form, the system (1.13) reads

$$\partial_t u = J\Omega u\,, \quad \Omega := \begin{pmatrix} 1-\kappa\partial_{xx} & 0 \\ 0 & G(0) \end{pmatrix}, \tag{1.14}$$

which is the Hamiltonian system generated by the quadratic Hamiltonian (see (1.6))

$$H_L := \frac{1}{2}(u,\Omega u)_{L^2(\mathbb{T}_x)} = \frac{1}{2}(\psi, G(0)\psi)_{L^2(\mathbb{T}_x)} + \frac{1}{2}\int_{\mathbb{T}} \left(\eta^2 + \kappa\eta_x^2\right) dx\,. \tag{1.15}$$

The standing wave solutions of the linear system (1.13), i.e. (1.14), are

$$\eta(t,x) = \sum\nolimits_{j\geq 1} a_j \cos(\omega_j t)\cos(jx), \quad \psi(t,x) = -\sum\nolimits_{j\geq 1} a_j j^{-1}\omega_j \sin(\omega_j t)\cos(jx)\,, \tag{1.16}$$

$a_j \in \mathbb{R}$, with linear frequencies of oscillations

$$\omega_j := \omega_j(\kappa) := \sqrt{j(1+\kappa j^2)}\,, \quad j \geq 1\,. \tag{1.17}$$

The main result of the paper proves that most of the standing wave solutions (1.16) of the linear system (1.13) can be continued to standing wave solutions of the nonlinear water-waves Hamiltonian system (1.3) for most values of the surface tension parameter $\kappa \in [\kappa_1, \kappa_2]$. More precisely, fix an arbitrary finite subset $\mathbb{S}^+ \subset \mathbb{N}^+ := \{1, 2, \ldots\}$ (tangential sites) and consider the linear standing wave solutions (of (1.13))

$$\eta(t,x) = \sum_{j\in\mathbb{S}^+} \sqrt{\xi_j}\cos(\omega_j t)\cos(jx), \quad \psi(t,x) = -\sum_{j\in\mathbb{S}^+} \sqrt{\xi_j} j^{-1}\omega_j \sin(\omega_j t)\cos(jx)\,,\ \xi_j > 0\,, \tag{1.18}$$

which are Fourier supported in $\mathbb{S}^+$. In Theorem 1.1 below we prove the existence of quasi-periodic solutions $u(\tilde\omega t, x) = (\eta,\psi)(\tilde\omega t, x)$ of (1.3), with frequency $\tilde\omega := (\tilde\omega_j)_{j\in\mathbb{S}^+}$ (to be determined), close to the solutions (1.18) of (1.13), for most values of the surface tension parameter $\kappa \in [\kappa_1, \kappa_2]$.

Let $\nu := |\mathbb{S}^+|$ denote the cardinality of $\mathbb{S}^+$. The function $u(\varphi, x) = (\eta,\psi)(\varphi,x)$, $\varphi \in \mathbb{T}^\nu$, belongs to the Sobolev spaces of $(2\pi)^{\nu+1}$-periodic real functions

$$H^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2) := \big\{u = (\eta,\psi) : \eta,\psi \in H^s\big\}$$

$$H^s := H^s(\mathbb{T}^{\nu+1}, \mathbb{R}) = \Big\{ f = \sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} \widehat{f}_{\ell j}\, e^{\mathrm{i}(\ell\cdot\varphi + jx)} \,:\, \|f\|_s^2 := \sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} |\widehat{f}_{\ell j}|^2 \langle \ell, j\rangle^{2s} < +\infty \Big\} \tag{1.19}$$

where $\langle \ell, j\rangle := \max\{1, |\ell|, |j|\}$ with $|\ell| := \max_{i=1,\ldots,\nu} |\ell_i|$. For

$$s \geq s_0 := \Big[\frac{\nu+1}{2}\Big] + 1 \in \mathbb{N} \tag{1.20}$$

the Sobolev spaces $H^s \subset L^\infty(\mathbb{T}^{\nu+1})$ are an algebra with respect to the product of functions.

Theorem 1.1. (KAM for capillary-gravity water waves) *For every choice of finitely many tangential sites $\mathbb{S}^+ \subset \mathbb{N}^+$, there exists $\bar s > s_0$, $\varepsilon_0 \in (0,1)$ such that for every $|\xi| \leq \varepsilon_0^2$, $\xi := (\xi_j)_{j\in\mathbb{S}^+}$, there exists a Cantor like set $\mathcal{G} \subset [\kappa_1, \kappa_2]$ with asymptotically full measure as $\xi \to 0$, i.e.*

$$\lim_{\xi\to 0} |\mathcal{G}| = \kappa_2 - \kappa_1\,,$$

such that, for any surface tension coefficient $\kappa \in \mathcal{G}$, the capillary-gravity system (1.3) *has a time quasi-periodic standing wave solution $u(\tilde\omega t, x) = (\eta(\tilde\omega t, x), \psi(\tilde\omega t, x))$, with Sobolev regularity $(\eta,\psi)(\varphi,x) \in H^{\bar s}(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R}^2)$, of the form*

$$\begin{aligned}\eta(\tilde\omega t, x) &= \sum\nolimits_{j\in\mathbb{S}^+} \sqrt{\xi_j}\cos(\tilde\omega_j t)\cos(jx) + o(\sqrt{|\xi|}),\\ \psi(\tilde\omega, x) &= -\sum\nolimits_{j\in\mathbb{S}^+} \sqrt{\xi_j} j^{-1}\omega_j \sin(\tilde\omega_j t)\cos(jx) + o(\sqrt{|\xi|})\end{aligned} \tag{1.21}$$

with a diophantine frequency vector $\tilde\omega := \tilde\omega(\kappa,\xi) \in \mathbb{R}^\nu$ satisfying $\tilde\omega_j - \omega_j(\kappa) \to 0$, $j \in \mathbb{S}^+$, as $\xi \to 0$. The terms $o(\sqrt{|\xi|})$ are small in $H^{\bar s}(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R}^2)$. In addition these quasi-periodic solutions are linearly stable.

Theorem 1.1 follows by Theorems 4.1 and 4.2. Let us make some comments.

1. No global in time existence results concerning the initial value problem of the water waves equations (1.3) under *periodic* boundary conditions are known so far. The present Nash-Moser-KAM iterative procedure selects many values of the surface tension parameter $\kappa \in [\kappa_1, \kappa_2]$ which give rise to the

quasi-periodic solutions (1.21), which are defined for all times. Clearly, by a Fubini-type argument it also results that, for most values of $\kappa \in [\kappa_1, \kappa_2]$, there exist quasi-periodic solutions of (1.3) for most values of the amplitudes $|\xi| \leq \varepsilon_0^2$. The fact that we find quasi-periodic solutions restricting to a proper subset of parameters is not a technical issue. The gravity-capillary water-waves equations (1.3) are not expected to be integrable (albeit a rigorous proof is still lacking): yet the third order Birkhoff normal form possesses multiple resonant triads (Wilton ripples), see Craig-Sulem [21].

2. In the proof of Theorem 1.1 all the estimates depend on the surface tension coefficient $\kappa > 0$ and the result does not hold at the limit of zero surface tension $\kappa \to 0$. Because of capillarity the linear frequencies (1.17) grow asymptotically $\sim \sqrt{\kappa} j^{3/2}$ as $j \to +\infty$. Without surface tension the linear frequencies grow asymptotically as $\sim j^{1/2}$ and a different proof is required.

3. The quasi-periodic solutions (1.21) are mainly supported in Fourier space on the tangential sites $\mathbb{S}^+$. The dynamics of the water waves equations (1.3) restricted to the symplectic subspaces

$$H_{\mathbb{S}^+} := \Big\{ v = \sum_{j \in \mathbb{S}^+} \begin{pmatrix} \eta_j \\ \psi_j \end{pmatrix} \cos(jx) \Big\}, \quad H_{\mathbb{S}^+}^{\perp} := \Big\{ z = \sum_{j \in \mathbb{N} \setminus \mathbb{S}^+} \begin{pmatrix} \eta_j \\ \psi_j \end{pmatrix} \cos(jx) \in H_0^1(\mathbb{T}_x) \Big\}, \tag{1.22}$$

is quite different. We call $v \in H_{\mathbb{S}^+}$ the *tangential* variable and $z \in H_{\mathbb{S}^+}^{\perp}$ the *normal* one. On the finite dimensional subspace $H_{\mathbb{S}^+}$ we describe the dynamics by introducing the action-angle variables $(\theta, I) \in \mathbb{T}^\nu \times \mathbb{R}^\nu$, see (4.7).

This is a difference with respect to the previous papers [39], [28], [29], [30], [31], [32], [1], that follow the Lyapunov-Schmidt decomposition. The present formulation enables, among other advantages, to prove the stability of the quasi-periodic solutions.

4. **Linear stability.** The quasi-periodic solutions $u(\tilde\omega t) = (\eta(\tilde\omega t), \psi(\tilde\omega t))$ found in Theorem 1.1 are linearly stable. This is not only a dynamically relevant information but also an essential ingredient of the existence proof (it is not necessary for time periodic solutions as in [1], [28], [29], [32]). Let us state precisely the result. Around each invariant torus there exist symplectic coordinates

$$(\phi, y, w) = (\phi, y, \eta, \psi) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_{\mathbb{S}^+}^{\perp}$$

(see (5.27) and [15]) in which the water waves Hamiltonian reads

$$\omega \cdot y + \frac{1}{2} K_{20}(\phi) y \cdot y + \big(K_{11}(\phi) y, w \big)_{L^2(\mathbb{T}_x)} + \frac{1}{2} \big(K_{02}(\phi) w, w \big)_{L^2(\mathbb{T}_x)} + K_{\geq 3}(\phi, y, w) \tag{1.23}$$

where $K_{\geq 3}$ collects the terms at least cubic in the variables (y, w) (see (5.29) and note that at a solution $\partial_\phi K_{00} = 0$, $K_{10} = \omega$, $K_{01} = 0$ by Lemma 5.6). In these coordinates the quasi-periodic solution reads $t \mapsto (\omega t, 0, 0)$ (for simplicity we denote the frequency $\tilde\omega$ of the quasi-periodic solution by ω) and the corresponding linearized water waves equations are

$$\begin{cases} \dot\phi = K_{20}(\omega t)[y] + K_{11}^T(\omega t)[w] \\ \dot y = 0 \\ \dot w = J K_{02}(\omega t)[w] + J K_{11}(\omega t)[y] \, . \end{cases} \tag{1.24}$$

Thus the actions $y(t) = y(0)$ do not evolve in time and the third equation reduces to the PDE

$$\dot w = J K_{02}(\omega t)[w] + J K_{11}(\omega t)[y(0)] \, . \tag{1.25}$$

The self-adjoint operator $K_{02}(\omega t)$ (defined in (5.29)) turns out to be the restriction to $H_{\mathbb{S}^+}^{\perp}$ of the linearized water-waves operator $\partial_u \nabla H(u(\omega t))$, explicitly computed in (6.8), up to a finite dimensional remainder, see Lemma 6.1.

Denote $H^s_\perp := H^s_\perp(\mathbb{T}_x) := H^s(\mathbb{T}_x) \cap H^\perp_{\mathbb{S}}$ (real or complex valued). sections 6 and 7 prove the existence of bounded and invertible "symmetrizer" maps, see (7.96), such that $\forall \varphi \in \mathbb{T}^\nu$, $m = 1, 2$

$$\mathbf{W}_{m,\infty}(\varphi) : \Big(H^s(\mathbb{T}_x, \mathbb{C}) \times H^s(\mathbb{T}_x, \mathbb{C})\Big) \cap H^\perp_{\mathbb{S}_+} \to \Big(H^s(\mathbb{T}_x, \mathbb{R}) \times H^{s-\frac{1}{2}}(\mathbb{T}_x, \mathbb{R})\Big) \cap H^\perp_{\mathbb{S}_+}\,, \tag{1.26}$$

$$\mathbf{W}^{-1}_{m,\infty}(\varphi) : \Big(H^s(\mathbb{T}_x, \mathbb{R}) \times H^{s-\frac{1}{2}}(\mathbb{T}_x, \mathbb{R})\Big) \cap H^\perp_{\mathbb{S}_+} \to \Big(H^s(\mathbb{T}_x, \mathbb{C}) \times H^s(\mathbb{T}_x, \mathbb{C})\Big) \cap H^\perp_{\mathbb{S}_+}\,, \tag{1.27}$$

and, under the change of variables

$$w = (\eta, \psi) = \mathbf{W}_{1,\infty}(\omega t) w_\infty\,, \quad w_\infty = (\mathtt{w}_\infty, \overline{\mathtt{w}}_\infty)\,,$$

the equation (1.25) transforms into the diagonal system

$$\partial_t w_\infty = -\mathrm{i}\mathbf{D}_\infty w_\infty + f_\infty(\omega t)\,, \quad f_\infty(\omega t) := \mathbf{W}_{2,\infty}(\varphi)(\omega t)^{-1} J K_{11}(\omega t)[y(0)] = \begin{pmatrix} \mathtt{f}_\infty(\omega t) \\ \overline{\mathtt{f}}_\infty(\omega t) \end{pmatrix} \tag{1.28}$$

where, denoting $\mathbb{S}_0 := \mathbb{S}_+ \cup (-\mathbb{S}_+) \cup \{0\} \subseteq \mathbb{Z}$,

$$\mathbf{D}_\infty := \begin{pmatrix} D_\infty & 0 \\ 0 & -D_\infty \end{pmatrix}\,, \quad D_\infty := \mathrm{diag}_{j \in \mathbb{S}_0^c}\{\mu_j^\infty\}\,, \quad \mu_j^\infty \in \mathbb{R}\,, \tag{1.29}$$

is a Fourier multiplier operator of the form (see (8.40))

$$\mu_j^\infty := \mathtt{m}_3^\infty \sqrt{|j|(1 + \kappa j^2)} + \mathtt{m}_1^\infty |j|^{\frac{1}{2}} + r_j^\infty\,,\ j \in \mathbb{S}_0^c\,, \quad r_j^\infty = r_{-j}^\infty\,, \tag{1.30}$$

where, for some $\mathtt{a} > 0$,

$$\mathtt{m}_3^\infty = 1 + O(\varepsilon^{\mathtt{a}})\,,\ \mathtt{m}_1^\infty = O(\varepsilon^{\mathtt{a}})\,,\ \sup_{j \in \mathbb{S}_0^c} |r_j^\infty| = O(\varepsilon^{\mathtt{a}})\,,\ \forall |k| \leq k_0\,,$$

see (4.23)-(4.24), (4.27) and $k_0 \in \mathbb{N}$ is a constant fixed in section 3 once for all along the whole paper and which depends only on the linear frequencies $\omega_j(\kappa)$ defined in (1.17). The μ_j^∞ are the *Floquet exponents* of the quasi-periodic solution. The second equation of system (1.28) is actually the complex conjugated of the first one, and (1.28) reduces to the infinitely many decoupled scalar equations

$$\partial_t \mathtt{w}_{\infty,j} = -\mathrm{i}\mu_j^\infty \mathtt{w}_{\infty,j} + \mathtt{f}_{\infty,j}(\omega t)\,, \quad \forall j \in \mathbb{S}_0^c\,.$$

By variation of constants the solutions are

$$\mathtt{w}_{\infty,j}(t) = c_j e^{-\mathrm{i}\mu_j^\infty t} + \mathtt{v}_{\infty,j}(t) \qquad \text{where} \qquad \mathtt{v}_{\infty,j}(t) := \sum_{\ell \in \mathbb{Z}^\nu} \frac{\mathtt{f}_{\infty,j,\ell}\, e^{\mathrm{i}\omega\cdot\ell t}}{\mathrm{i}(\omega \cdot \ell + \mu_j^\infty)}\,, \quad \forall j \in \mathbb{S}_0^c\,. \tag{1.31}$$

Note that the first Melnikov conditions (4.25) hold at a solution so that $\mathtt{v}_{\infty,j}(t)$ in (1.31) is well defined. Moreover (1.26) implies $\|f_\infty(\omega t)\|_{H^s_x \times H^s_x} \leq C|y(0)|$. As a consequence the Sobolev norm of the solution of (1.28) with initial condition $w_\infty(0) \in H^{\mathfrak{s}_0}(\mathbb{T}_x)$, $\mathfrak{s}_0 < s$ (in a suitable range of values), satisfies

$$\|w_\infty(t)\|_{H^{\mathfrak{s}_0}_x \times H^{\mathfrak{s}_0}_x} \leq C(s)(|y(0)| + \|w_\infty(0)\|_{H^{\mathfrak{s}_0}_x \times H^{\mathfrak{s}_0}_x})\,,$$

and, for all $t \in \mathbb{R}$, using (1.26), (1.27), we get

$$\|(\eta, \psi)(t)\|_{H^{\mathfrak{s}_0}_x \times H^{\mathfrak{s}_0 - \frac{1}{2}}_x} \lesssim \|(\eta(0), \psi(0))\|_{H^{\mathfrak{s}_0}_x \times H^{\mathfrak{s}_0 - \frac{1}{2}}_x}$$

which proves the linear stability of the torus. Note that the profile $\eta \in H^{\mathfrak{s}_0}(\mathbb{T}_x)$ is more regular than the velocity potential $\psi \in H^{\mathfrak{s}_0 - \frac{1}{2}}(\mathbb{T}_x)$, as it is expected in presence of surface tension, see [2].

Clearly a crucial point is the diagonalization of (1.25) into (1.29). With respect to [1] this requires to analyze more in detail the pseudo-differential nature of the operators obtained after each conjugation and to implement a KAM scheme with second order Melnikov non-resonance conditions, as we shall explain in detail below.

5. *Hamiltonian and reversible structure.* It is well known that the existence of quasi-periodic motions is possible just for systems with some algebraic structure which excludes "secular motions" and friction phenomena. The most common ones are the Hamiltonian and the reversible structure. The water-waves system (1.3) exhibits both of them and we shall use both. The Hamiltonian structure is used in particular in section 5 to introduce the symplectic coordinates (ϕ, y, w) in (5.27) adapted to an approximately-invariant torus. On the other hand, for solving the second equation of the linear system (5.50) we use reversibility (we could exploit just the Hamiltonian structure as done in [8]-[9], [15]-[16]). Moreover the transformations $\mathbf{W}_{1,\infty}$, $\mathbf{W}_{2,\infty}$ which reduce the linearized operator to constant coefficients preserve the reversible structure (it is slightly simpler than to preserve the Hamiltonian one). Reversibility implies that several averaged vector fields are zero, for example a constant coefficient operator of the form $h \mapsto a\partial_x h$, $a \in \mathbb{R}$, is not compatible with the reversible structure of the water waves, and therefore it is zero. This leads to the asymptotic expansion of the Floquet exponents (1.30), in particular to the fact that they are purely imaginary. The linear stability of the quasi-periodic standing wave solutions of Theorem 1.1 is a consequence of the reversible structure of the water waves equations.

We prove the existence of quasi-periodic solutions by a Nash-Moser iterative scheme in Sobolev spaces formulated as a 'Théoréme de conjugaison hypothétique" á la Herman (section 4.1). In order to perform effective measure estimates in the surface tension parameter $\kappa \in [\kappa_1, \kappa_2]$ (section 4.2) we use degenerate KAM theory for PDEs (section 3). For the convergence of the Nash-Moser scheme (section 8) it is sufficient to have an "almost approximate" inverse of the linearized operators at each step of the iteration. We follow (section 5) the scheme proposed in [15]-[16], and implemented in [8]-[9], which reduces the problem to "almost approximately" invert the linearized operator restricted to the normal directions. The crucial PDE analysis is the reduction in sections 6-7 of the linearized operator to constant coefficients.

Let us present more in details some key ideas of the paper.

1. *Bifurcation analysis and Degenerate KAM theory.* A first key observation is that, for most values of the surface tension parameter $\kappa \in [\kappa_1, \kappa_2]$, the unperturbed linear frequencies (1.17) are diophantine and satisfy also first and second order Melnikov non-resonance conditions. More precisely the unperturbed tangential frequency vector $\vec{\omega}(\kappa) := (\omega_j(\kappa))_{j\in\mathbb{S}^+}$ satisfies

$$|\vec{\omega}(\kappa) \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau}, \ \forall \ell \in \mathbb{Z}^\nu \setminus \{0\},$$

and it is non-resonant with the normal frequencies $\vec{\Omega}(\kappa) := (\Omega_j(\kappa))_{j\in\mathbb{N}^+\setminus\mathbb{S}^+} = (\omega_j(\kappa))_{j\in\mathbb{N}^+\setminus\mathbb{S}^+}$, i.e.

$$|\vec{\omega}(\kappa) \cdot \ell + \Omega_j(\kappa)| \geq \gamma j^{\frac{3}{2}} \langle \ell \rangle^{-\tau}, \ \forall \ell \in \mathbb{Z}^\nu, \ j \in \mathbb{N}^+ \setminus \mathbb{S}^+ \,,$$
$$|\vec{\omega}(\kappa) \cdot \ell + \Omega_j(\kappa) \pm \Omega_{j'}(\kappa)| \geq \gamma |j^{\frac{3}{2}} \pm j'^{\frac{3}{2}}| \langle \ell \rangle^{-\tau}, \forall \ell \in \mathbb{Z}^\nu, j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \,.$$

This is a problem of diophantine approximation on submanifolds. It can be solved by degenerate KAM theory (explained below) exploiting that the linear frequencies $\kappa \mapsto \omega_j(\kappa)$ are *analytic*, simple, grow asymptotically as $j^{3/2}$ and are *non-degenerate* in the sense of Bambusi-Berti-Magistrelli [10] (another proof can be given by the tools of subanalytic geometry in Delort-Szeftel [23]). For such values of $\kappa \in [\kappa_1, \kappa_2]$, the solutions (1.18) of the linear equation (1.13) are already sufficiently good approximate quasi-periodic solutions of the nonlinear water waves system (1.3). Since the parameter space $[\kappa_1, \kappa_2]$ is fixed, the small divisor constant γ can be taken $\gamma = o(\varepsilon^a)$ with $a > 0$ small as needed, see (4.27). As a consequence for proving the continuation of (1.18) to solutions of the nonlinear water waves system (1.3), all the terms which are at least quadratic in (1.3) are yet perturbative (in (4.1) it is sufficient to regard the vector field $\varepsilon X_{P_\varepsilon}$ as a perturbation of the linear vector field $J\Omega$).

Actually along the Nash-Moser-KAM iteration we need to verify that the perturbed frequencies are diophantine and satisfy first and second order Melnikov non-resonance conditions. It is actually for that we find convenient to develop degenerate KAM theory as in [10] and we formulate the problem as a Théoréme de conjugaison hypothétique à la Nash-Moser as we explain below.

2. *A Nash-Moser Théoréme de conjugaison hypothétique.* The expected quasi-periodic solutions of the autonomous Hamiltonian system (1.3) will have shifted frequencies $\tilde{\omega}_j$ -to be found- close to the linear frequencies $\omega_j(\kappa)$ in (1.17), which depend on the nonlinearity and the amplitudes ξ_j. Since the Melnikov non-resonance conditions are naturally imposed on ω, it is convenient to use the functional setting formulation of Theorem 4.1 where the *parameters* are the *frequencies* $\omega \in \mathbb{R}^\nu$ and the *surface tension* $\kappa \in [\kappa_1, \kappa_2]$ and we introduce a counterterm $\alpha \in \mathbb{R}^\nu$ in the family of Hamiltonians H_α defined in (4.16).

   Then the goal is to prove that, for ε small enough, for "most" parameters $(\omega, \kappa) \in \mathcal{C}^\gamma_\infty$, there exists a value of the constants $\alpha := \alpha_\infty(\omega, \kappa, \varepsilon) = \omega + O(\varepsilon\gamma^{-k})$ and a ν-dimensional embedded torus $\mathcal{T} = i(\mathbb{T}^\nu)$ close to $\mathbb{T}^\nu \times \{0\} \times \{0\}$, invariant for the Hamiltonian vector field $X_{H(\alpha_\infty(\omega,\kappa,\varepsilon),\cdot)}$ and supporting quasi-periodic solutions with frequency ω. This is equivalent to look for a zero of the nonlinear operator $\mathcal{F}(i, \alpha, \omega, \kappa, \epsilon) = 0$ defined in (4.17). This equation is solved in Theorem 4.1 by a Nash-Moser iterative scheme. The value of $\alpha := \alpha_\infty(\omega, \kappa, \varepsilon)$ is adjusted along the iteration in order to control the average of the first component of the Hamilton equation (4.17), in particular for solving the linearized equation (5.44), (5.54).

   The set of parameters $(\omega, \kappa) \in \mathcal{C}^\gamma_\infty$ for which the invariant torus exists is the explicit Cantor set (4.25). We require that ω satisfies the diophantine property

$$|\omega \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau}, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \tag{1.32}$$

   and, in addition, the first and second Melnikov non-resonance conditions.

   Note that the Cantor like set $\mathcal{C}^\gamma_\infty$ is defined in terms of the "final torus" i_∞ (see (4.22)) and the "final eigenvalues" in (4.23) which are defined for *all* the values of the frequency ω by a Whitney-type extension argument, see the sentences after (1.40). This formulation completely decouples the Nash-Moser iteration (which provides the torus $i_\infty(\omega, \kappa, \varepsilon)$ and the constant $\alpha_\infty(\omega, \kappa, \varepsilon) \in \mathbb{R}^\nu$) from the discussion about the measure of the set of parameters where all the non-resonance conditions are indeed verified. This simplifies the measure estimates which are no longer imposed at each step but only once, see section 4.2. This formulation follows that of [14] (in a Lyapunov-Schmidt context) and [11] (in a KAM theorem) and [17] (in a Nash-Moser context). The measure estimates are done in section 4.2.

   In order to prove the existence of quasi-periodic solutions of the water waves equation (1.3), and not only of the system with modified Hamiltonian H_α with $\alpha := \alpha_\infty(\omega, \kappa, \varepsilon)$, we have then to prove that the curve of the unperturbed linear frequencies

$$[\kappa_1, \kappa_2] \ni \kappa \mapsto \vec{\omega}(\kappa) := (\sqrt{j(1 + \kappa j^2)})_{j \in \mathbb{S}^+} \in \mathbb{R}^\nu$$

   intersects the image $\alpha_\infty(\mathcal{C}^\gamma_\infty)$, under the map α_∞ of the Cantor set $\mathcal{C}^\gamma_\infty$, for "most" values of $\kappa \in [\kappa_1, \kappa_2]$. This is proved in Theorem 4.2 by degenerate KAM theory. For such values of κ we have found a quasi-periodic solution of (1.3) with diophantine frequency $\omega_\varepsilon(\kappa) := \alpha_\infty^{-1}(\vec{\omega}(\kappa), \kappa)$.

   The above functional setting perspective is in the spirit of the Théoréme de conjugaison hypothétique of Herman proved by Fejoz [25] for finite dimensional Hamiltonian systems, see also the discussion in [15]. A relevant difference is that in [25], in addition to α, also the normal frequencies are introduced as independent parameters, unlike in Theorem 4.1. Actually for PDEs it seems more convenient the present formulation: it is a major point of the work to know the asymptotic expansion (1.30) of the Floquet exponents.

3. *Degenerate KAM theory and measure estimates.* In Theorem 4.2 we prove that for all the values of $\kappa \in [\kappa_1, \kappa_2]$ except a set of small measure $O(\gamma^{1/k_0})$ (the value of $k_0 \in \mathbb{N}$ is fixed once for all in section 3) the frequency vector $\vec{\omega}(\kappa)$ belongs to the Cantor set $\alpha_\infty(\mathcal{C}^\gamma_\infty)$, see the set $\mathcal{G}_\varepsilon$ in (4.28). As already said, we use in an essential way that the unperturbed frequencies $\kappa \mapsto \omega_j(\kappa)$ are *analytic*, are simple (on the subspace of the even functions), grow asymptotically as $j^{3/2}$ and are *non-degenerate* in the sense of [10]. This is verified in Lemma 3.1 as in [10] by a Van der Monde determinant. Then we develop degenerate KAM theory which reduces this qualitative non-degeneracy condition into a

quantitative one, which is sufficient to estimate effectively the measure of the Cantor like set $\mathcal{G}_\varepsilon$ by the classical Rüssmann lemma. We deduce in Proposition 3.2 that $\exists k_0 > 0$, $\rho_0 > 0$ such that, for all $\kappa \in [\kappa_1, \kappa_2]$,

$$\max_{0 \le k \le k_0} \left| \partial_\kappa^k \left(\vec{\omega}(\kappa) \cdot \ell + \Omega_j(\kappa) - \Omega_{j'}(\kappa) \right) \right| \ge \rho_0 \langle \ell \rangle \, , \quad \forall (\ell, j, j') \neq (0, j, j), \; j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \, , \tag{1.33}$$

and similarly for the 0-th, 1-th and the 2-th order Melnikov non-resonance condition with the sign $+$. Note that the restriction to the subspace (1.8), see also (1.10), of functions with zero average in x eliminates the zero frequency $\omega_0 = 0$, which is trivially resonant (this is used also in [22]). Property (1.33) implies that for "most" parameters $\kappa \in [\kappa_1, \kappa_2]$ the unperturbed linear frequencies $(\vec{\omega}(\kappa), \vec{\Omega}(\kappa))$ satisfy the Melnikov conditions of $0, 1, 2$ order (but we do not use it explicitly). Actually, the condition (1.33) is stable under perturbations which are small in $\mathcal{C}^{k_0}$-norm, see Lemma 4.4. Since the perturbed Floquet exponents in (4.31) are small perturbations of the unperturbed linear frequencies $\sqrt{j(1 + \kappa j^2)}$ in $\mathcal{C}^{k_0}$-norm (see (4.30) and (4.33)) the property (1.33) still holds for the perturbed frequencies $\omega_\varepsilon(\kappa)$ defined in (4.29). As a consequence, by applying the classical Rüssmann lemma (Theorem 17.1 in [41]) we prove that the Cantor like set of non-resonant parameters $\mathcal{G}_\varepsilon$ has a large measure, see Lemma 4.5 and the end of the proof of Theorem 4.2.

Analysis of the linearized operators. The other crucial analysis for the Nash-Moser iterative scheme is to prove that the *linearized operator* obtained at any approximate solution is, for most values of the parameters, invertible, and that its inverse satisfies *tame* estimates in Sobolev spaces. We implement in section 5 the procedure developed in Berti-Bolle [15] and [8]-[9] for autonomous PDEs. It consists in introducing a convenient set of symplectic variables (see (5.27)) near the approximate torus such that the linearized equations become (approximately) decoupled in the action-angle components and the normal ones, see (5.44). As a consequence, the problem is reduced to "almost-approximately" invert the linearized operator $\mathcal{L}_\omega$ defined in (5.40). Actually, since the symplectic change of variables (5.27) modifies, up to a translation, only the finite dimensional action component, the linear operator $\mathcal{L}_\omega$ is nothing but the linearized water-waves operator $\mathcal{L}$ computed in (6.8) -in the original coordinates- up to a finite dimensional remainder and restricted to the normal directions. Thus the key part of the analysis consists in (almost) reducing the quasi-periodic linear operator $\mathcal{L}$ to constant coefficients, via linear changes of variables close to the identity, which map Sobolev spaces into itself and satisfy tame estimates, see Proposition 7.12.

This is achieved in sections 6 and 7 by making full use of pseudo-differential operator theory that we present in section 2.1 in a formulation convenient to our purposes.

Pseudo-differential operators. We underline that all the coefficients of the linearized operator $\mathcal{L}$ in (6.8) are $\mathcal{C}^\infty$ in (φ, x) because each approximate solution $(\eta(\varphi, x), \psi(\varphi, x))$ at which we linearize along the Nash-Moser iteration is a trigonometric polynomial in (φ, x) (at each step we apply the projector Π_n defined in (8.1)) and the water waves vector field is analytic. This allows to work in the usual framework of $\mathcal{C}^\infty$ pseudo-differential symbols.

In this paper we only use the class S^m of (classical) symbols introduced in Definition 2.6. We do not explicitly make use of pseudo-differential operators in the class $OPS^m_{\frac{1}{2},\frac{1}{2}}$ used by Alazard-Baldi in [1] (called semi-Fourier integral operators). Actually we shall produce similar transformations as *flows* of pseudo-PDEs (see (6.130)). The advantage is that the invertibility of such transformations, as well as the fact that they satisfy tame estimates in Sobolev spaces together with its inverses, follows easily by proving energy estimates for the flow, see Appendix 9.

For the Nash-Moser convergence we clearly need to perform quantitative estimates in Sobolev spaces. Then, given a pseudo-differential operator $A = \mathrm{Op}(a(\varphi, x, \xi)) \in OPS^m$, we introduce the norm $|A|_{m,s,\alpha}$ defined in (2.36) (more generally $|A|^{k_0,\gamma}_{m,s,\alpha}$ in Definition 2.7), which is inspired to the para-differential norm in Metivier [38], chapter 5. Note that $|A|_{m,s,\alpha}$ controls the regularity in (φ, x) of the symbol $a(\varphi, x, \xi) \in S^m$ only up to a limited smoothness.

We now explain the main steps for the reduction of the quasi-periodic linear operator $\mathcal{L}$ in (6.8).

1. *Reduction of $\mathcal{L}$ to constant coefficients in decreasing symbols.* The goal of section 6 (Proposition

6.31) is to reduce $\mathcal{L}$ to a quasi-periodic linear operator of the form

$$(h, \bar{h}) \mapsto \big(\omega\cdot\partial_\varphi + \mathrm{i}\mathtt{m}_3 T(D) + \mathrm{i}\mathtt{m}_1 |D|^{\frac12}\big)h + \mathcal{R}h + \mathcal{Q}\bar{h}\,, \quad h \in \mathbb{C}\,, \tag{1.34}$$

where $\mathtt{m}_3, \mathtt{m}_1 \in \mathbb{R}$ are constants satisfying $\mathtt{m}_3 \approx 1$, $\mathtt{m}_1 \approx 0$, the principal symbol operator is

$$T(D) := |D|^{1/2}(1 - \kappa\partial_{xx})^{1/2}\,,$$

and the remainders $\mathcal{R} := \mathcal{R}(\varphi)$, $\mathcal{Q} := \mathcal{Q}(\varphi)$ are small bounded operators acting in the Sobolev spaces H^s, which satisfy tame estimates. More precisely, in view of the KAM reducibility scheme of section 7, we need that all the operators in (1.38), together with its derivatives $\partial^k_{\omega,\kappa}\mathcal{R}$, $\partial^k_{\omega,\kappa}\mathcal{Q}$, $|k| \leq k_0$, satisfy tame estimates, see (6.248). We neglect in (1.34) smoothing operators which are supported on high Fourier frequencies (ultra-violet cut-off) and therefore satisfy (6.244)-(6.245). Note that (1.34) is an operator which acts on $(h, \bar{h})$. We shall deal in a quite different way the operator $h \mapsto (\omega\cdot\partial_\varphi + \mathrm{i}\mathtt{m}_3 T(D) + \mathrm{i}\mathtt{m}_1|D|^{\frac12})h + \mathcal{R}h$ and $\bar{h} \mapsto \mathcal{Q}\bar{h}$. We shall call the first operator, "diagonal", and the latter, 'off-diagonal", with respect to the variables $(h, \bar{h})$.

2. *Symmetrization and space-time reduction of $\mathcal{L}$ at the highest order.* The first part of the analysis (sections 6.1-6.2) is similar to Alazard-Baldi [1]. A difference is that we reduce the linear operator $\mathcal{L}$ in (6.8) to constant coefficients up to OPS^0 remainders (Lemma 6.7), while in [1] the remainders are $O(\partial_x^{-3/2})$. The reason of this difference is that we will not invert the linearized operator in (1.34) simply by a Neumann-argument, as done for the periodic solutions in [1], [32], [28], [29], [39]. This approach does not work in the quasi-periodic case. The key difference is that, in the periodic problem, a sufficiently regularizing operator in the space variable is also regularizing in the time variable, on the characteristic Fourier indices which correspond to the small divisors. This is clearly not true for quasi-periodic solutions.

   Our strategy will be to diagonalize (actually it is sufficient to "almost diagonalize") the linearized operator in (1.34) by the KAM scheme of section 7. This requires to analyze more in detail the pseudo-differential nature of the remainders after all the conjugation steps -a key difference concerns the nature of the block-off diagonal operators in $(h, \bar{h})$ with respect to the diagonal ones- and to be able to impose the second Melnikov non-resonance conditions.

   In section 6.3 we introduce complex coordinates $(h, \bar{h})$, which are convenient to reduce the off-diagonal blocks of the linear system to a very negative order (section 6.5). We could have introduced the complex variables $(h, \bar{h})$ right after section 6.1 performing the symmetrization procedure and the space reduction of the highest order (section 6.2) in the variables $(h, \bar{h})$. This way, however, would require to use an Egorov type argument to estimate the remainders unlike in section 6.2 we use (as in [1]) only the simple change of variables (6.22).

   Then in section 6.4, using a time-reparametrization as in [1], we obtain a quasi-periodic linear operator of the form (see (6.74))

$$(h, \bar{h}) \mapsto \big(\omega\cdot\partial_\varphi + \mathrm{i}\mathtt{m}_3 T(D) + a_{11}(\varphi, x)\partial_x + \mathrm{i}a_{12}(\varphi, x)\mathcal{H}|D|^{\frac12}\big)h + \mathrm{i}b(\varphi, x)\mathcal{H}|D|^{\frac12}\bar{h} + \dots\,. \tag{1.35}$$

   From this point we have to proceed quite differently with respect to [1].

3. *Block-decoupling.* In view of the transformations used in the next Egorov-step and the KAM reducibility scheme of section 7, we first reduce the order of the off-diagonal term $\mathrm{i}b(\varphi, x)\mathcal{H}|D|^{\frac12}\bar{h}$ to a very negative order OPS^{-M}. In section 6.5 we conjugate (1.35) to a quasi-periodic linear operator of the form (Proposition 6.11)

$$(h, \bar{h}) \mapsto \omega\cdot\partial_\varphi h + \mathrm{i}\mathtt{m}_3 T(D)h + a_{11}(\varphi, x)\partial_x h + \mathrm{i}a_{12}(\varphi, x)\mathcal{H}|D|^{\frac12}h + \mathcal{R}_M h + \mathcal{Q}_M\bar{h}$$

   where $\mathcal{R}_M \in OPS^0$ and $\mathcal{Q}_M \in OPS^{-M}$, for some M large enough which is fixed by the KAM reducibility scheme, see (7.9).

4. *Egorov analysis. Space reduction of the order ∂_x.* The goal of section 6.6 is to eliminate the first order vector field $a_{11}(\varphi, x)\partial_x$. For that Alazard-Baldi [1] used a semi-Fourier integral operator like $\mathrm{Op}(e^{\mathrm{i}a(\varphi,x)\sqrt{|\xi|}}) \in OPS^0_{\frac{1}{2},\frac{1}{2}}$. We shall use instead the flow $\Phi(\varphi) := \Phi(\varphi, \omega, \kappa)$ of the pseudo-PDE

$$u_t = \mathrm{i}a(\varphi, x, \omega, \kappa)|D|^{1/2}u\,. \tag{1.36}$$

The proof that Φ, as well as its inverse Φ^{-1}, is well posed in Sobolev spaces H^s and satisfies tame estimates, follow by the energy estimates of Appendix 9 (the vector field $\mathrm{i}a(\varphi, x, \omega, \kappa)|D|^{1/2}$ is skew-adjoint at the highest order). We think that this is conceptually simpler than proving directly the invertibility and the tame estimates of $\mathrm{Op}(e^{\mathrm{i}a(\varphi,x)\sqrt{|\xi|}})$ as in [1].

However the main advantage in order to use the present flow approach consists in the Egorov analysis of the pseudo-differential nature of the conjugated operator. The flow has a very different effect on the operator $h \mapsto (\mathrm{i}a_{12}(\varphi, x)\mathcal{H}|D|^{\frac{1}{2}} + \mathcal{R}_M)h$ and the off-diagonal one $\bar{h} \mapsto \mathcal{Q}_M\bar{h}$: the first remains a classical pseudo-differential operator in OPS^0 (Egorov analysis), but the off-diagonal one becomes a pseudo-differential operator in the class $OPS^{-M}_{\frac{1}{2},\frac{1}{2}}$.

Let us roughly explain why this is a relevant information. The flow $\Phi(\varphi) \sim \mathrm{Op}(e^{\mathrm{i}a(\varphi,x)\sqrt{|\xi|}})$ maps Sobolev spaces in itself. However each derivative

$$\partial_\varphi\Phi(\varphi) \sim \mathrm{Op}\big(e^{\mathrm{i}a(\varphi,x)\sqrt{|\xi|}}\,\mathrm{i}\partial_\varphi a(\varphi, x)\sqrt{|\xi|}\big)$$

is an unbounded operator which loses $|D|^{1/2}$ derivatives. In the Appendix we actually prove that $\partial^k_{\omega,\kappa}\partial^\beta_\varphi\Phi(\varphi)$ satisfies tame estimates with a loss of $|D|^{\frac{|\beta|+|k|}{2}}$ derivatives.

The main idea of the Egorov analysis in section 6.6 is that, given a scalar classical pseudo-differential operator $P_0 \in OPS^m$, the conjugated operator

$$P_+(\varphi) := \Phi(\varphi)P_0\Phi(\varphi)^{-1} = \mathrm{Op}(c(\varphi, x, \xi))\,, \quad c(\varphi, x, \xi) \in S^m\,, \tag{1.37}$$

remains as well a classical pseudo-differential operator. Therefore, the differentiated operator $\partial_\varphi P_+(\varphi) = \mathrm{Op}(\partial_\varphi c(\varphi, x, \xi)) \in OPS^m$ is a pseudo-differential operator of the same order of P_0 with a symbol $\partial_\varphi c$ which is just less regular in φ. Then the loss of regularity for $\partial_\varphi c$ is compensated by the usual Nash-Moser smoothing procedure in φ. The property (1.37) is due to the fact that P_+ is "transported" under the flow of (1.36) according to the Heisenberg equation (6.135).

This is the reason why we require that the diagonal remainder $\mathcal{R} \in OPS^0$ is just of order zero.

On the other hand, the off-diagonal term $\mathcal{Q}_M \in OPS^{-M}$ evolves, under the flow of (1.36), according to the "skew-Heisenberg" equation obtained replacing in (6.135) the commutator with the skew-commutator. As a consequence the symbol of $\mathcal{Q}^+_M := \Phi(\varphi)\mathcal{Q}_M\Phi(\varphi)^{-1}$ assumes the form $e^{\mathrm{i}a(\varphi,x)\sqrt{|\xi|}}q(\varphi, x, \xi)$ where $q(\varphi, x, \xi) \in S^{-M}$ is a classical symbol (actually we do not prove it explicitly because it is not needed). Thus the action of each ∂_φ on $\mathcal{Q}^+_M$ produces an operator which loses $|D|^{\frac{1}{2}}$ derivatives in space more than $\mathcal{Q}_M$. This is why we perform in section 6.5 a large number M of regularizing steps for the off-diagonal components $\mathcal{Q}$. The constant M is fixed later in (7.9). The precise tame estimates of $\partial^\beta_\varphi\mathcal{Q}^+_M$ are given in Proposition 6.26 for $M \geq \beta + k_0 + 4$. In section 7 we take $\beta \sim \mathtt{b}$, see (7.9).

5. *Space reduction of the order $|D|^{1/2}$.* In section 6.7 we reduce to constant coefficients also the diagonal operator term of order $|D|^{1/2}$. This concludes (section 6.8) the conjugation of $\mathcal{L}_\omega$ to a quasi-periodic linear operator like (1.34).

6. *KAM-reducibility scheme.* We apply the KAM diagonalization scheme of section 7 to a linear operator as in (1.34) where

$$\mathcal{R}\,,\ [\mathcal{R}, \partial_x]\,,\ \partial^{s_0}_{\varphi_m}\mathcal{R}\,,\ \partial^{s_0}_{\varphi_m}[\mathcal{R}, \partial_x]\,,\ \partial^{s_0+\mathtt{b}}_{\varphi_m}\mathcal{R}\,,\ \partial^{s_0+\mathtt{b}}_{\varphi_m}[\mathcal{R}, \partial_x]\,,\ m = 1, \ldots, \nu\,, \tag{1.38}$$

and similarly $\mathcal{Q}$, satisfy tame estimates for some $\mathtt{b} := \mathtt{b}(\tau, k_0) \in \mathbb{N}$ large enough, fixed in (7.6), see (7.4), (7.5), (7.7). Such condition is proved in Lemma 7.2, having assumed that M (= number of regularizing steps for the off-diagonal operators performed in section 6.5) is taken large as in (7.9) (essentially $M = O(\mathtt{b})$). It is the property which compensates, along the KAM iteration, the loss of derivatives in φ produced by the small divisors (this condition is strictly weaker than assuming a polynomial off-diagonal decay of $\mathcal{R}$, $\mathcal{Q}$, as in [7]-[8]).

The core of the KAM reducibility scheme of section 7 is to prove that the class of operators which are $\mathcal{D}^{k_0}$-modulo-tame (Definition 2.9) is closed under the operations involved by a KAM iteration, namely

(a) composition (Lemma 2.16),

(b) solution of the homological equation (Lemma 7.7),

(c) projections (Lemma 2.18).

We recall that we have to control that the KAM transformations (and all the operators) are k_0-times differentiable with respect to the parameters $(\omega, \kappa) \in \mathbb{R}^\nu \times [\kappa_1, \kappa_2]$ to prove that the Floquet exponents $(\omega, \kappa) \mapsto \mu_j^\infty(\omega, \kappa)$ in (4.23) are small perturbations of the linear frequencies $\sqrt{j(1+\kappa j^2)}$ in $\mathcal{C}^{k_0}$-norm.

The reason why we implement the KAM reducibility scheme for $\mathcal{D}^{k_0}$-modulo-tame operators and not only for $\mathcal{D}^{k_0}$-tame operators is that for a $\mathcal{D}^{k_0}$-tame operator the second estimate in Lemma 2.18 for the projector $\Pi_N^\perp$ does not hold (majorant like norms have been used also in [12]-[13]). The fact that the initial majorant operators $|\mathcal{R}|$, $|\mathcal{Q}|$ (see Definition 2.2) fulfill tame estimates (which is stronger that requiring tame estimates just for $\mathcal{R}$ and $\mathcal{Q}$) is verified in Lemma 7.6 thanks to the assumption that $[\partial_x, \mathcal{R}]$ and $\partial_{\varphi_m}^{s_0} \mathcal{R}$, as well as all the operators in (1.38), satisfy tame estimates, see Lemma 7.2. Note that the commutator $[\partial_x, r(x, D)] = r_x(x, D)$ is a pseudo-differential operator with the same order of $r(x, D)$ (this is used in particular in Proposition 6.26). This is another reason for which it is sufficient that the pseudo-differential remainder which acts on the diagonal (i.e. on h) is just in OPS^0.

The key (quadratic + super-exponentially small) inductive estimates required for the convergence of the iteration are provided by Lemma 7.9. More precisely (7.74) and (7.75) allow to prove the convergence of the scheme up to the Sobolev index s, by choosing $\mathtt{b} := \mathtt{b}(\tau)$ large enough as fixed in (7.6). The inductive relation (7.75) provides an a priori bound for the divergence of the modulo-tame constants $\mathfrak{M}_\nu^\sharp(s, \mathtt{b})$ of the operators $\langle \partial_\varphi \rangle^{\mathtt{b}} \mathcal{R}_{\nu+1}$ and $\langle \partial_\varphi \rangle^{\mathtt{b}} \mathcal{Q}_{\nu+1}$ along the iteration. Then (7.74) shows that $\mathfrak{M}_\nu^\sharp(s)$ converges very rapidly to 0 as $\nu \to +\infty$, see (7.22).

Note that the iterative KAM Theorem 7.3 requires only the smallness condition (7.14) which involves just the low norm $\| \ \|_{s_0+\mathtt{b}}$ but implies also tame estimates up to the Sobolev scale s, see (7.22). The important consequence is that, in Theorem 7.5, only the condition (7.33) in low norm, implies the tame estimates (7.37) for the transformations up to any $s \in [s_0, S]$. The smallness condition (7.33) will be verified inductively along the nonlinear Nash-Moser scheme of section 8. The tame property (7.37) (at any scale) is used in the convergence of the Nash-Moser iteration of section 8.

After the above analysis of the linearized operator, in section 8, we implement a differentiable Nash-Moser iterative scheme to find better and better approximate quasi-periodic solutions up to the scales

$$K_n := K_0^{\chi^n}, \quad \chi := 3/2, \tag{1.39}$$

which lead, at the limit, to an embedded torus invariant under the flow of the Hamiltonian PDE, see Theorem 8.2 and section 8.1.

We conclude the introduction with some other comment.

1. *Whitney extension.* At each iterative step of the Nash-Moser iteration -and correspondingly for the reduction of the linearized operator in sections 5, 6, 7- we only require that the frequency vector

$\omega \in \mathbb{R}^\nu$ satisfies finitely many non-resonance diophantine conditions. More precisely we assume at the n-th step that ω belongs to

$$\mathtt{DC}^{\gamma}_{K_n} := \left\{ \omega \in \Omega \subset \mathbb{R}^\nu \,:\, |\omega \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau} \,,\ \forall |\ell| \leq K_n \right\} \tag{1.40}$$

and similarly we require finitely many first and second order Melnikov non-resonance conditions, see (7.87) and (7.19) (the set Ω is the neighborhood (4.20) of the curve $\vec{\omega}([\kappa_1, \kappa_2])$ described by the unperturbed linear frequencies $\vec{\omega}$). This allows to perform a constructive Whitney extension of the solution, with respect to the parameters (ω, κ) in a way similar to [14]. We find this construction convenient in order to estimate the k-derivatives $\partial^k_{\omega,\kappa}$ of the approximate solutions (and of the eigenvalues) which, on an open subset (like $\mathtt{DC}^{\gamma}_{K_n}$) are well defined in the usual sense (instead of introducing the notion of Whitney derivatives on closed subsets). The quantitative estimates that we shall obtain (see for example (4.22) and (4.33)) are similar to those which are satisfied by the solution

$$h := (\omega \cdot \partial_\varphi)^{-1} g = \sum_{\ell \in \mathbb{Z}^\nu \setminus \{0\}} \frac{g_\ell}{\mathrm{i} \omega \cdot \ell} e^{\mathrm{i} \ell \cdot \varphi} \,, \qquad g := \sum_{\ell \in \mathbb{Z}^\nu \setminus \{0\}} g_\ell e^{\mathrm{i} \ell \cdot \varphi} \,,$$

of the basic linear equation of KAM theory $\omega \cdot \partial_\varphi h = g$, namely

$$\|\partial^k_\omega h\|_s \leq C \gamma^{-|k|} \|g\|_{s + \tau + |k| \tau} \,.$$

We note that each derivative ∂_ω produces a factor γ^{-1} and a loss of τ-derivatives in the Sobolev index. This is the phenomenon described by Pöschel in [40] as “anisotropic differentiability” of the Cantor families of KAM tori with respect to ω. Actually when solving the homological equations, see (7.58)-(7.59), we also have denominators which depend on both (ω, κ) and we have to estimate the regularity of the solution also with respect to κ, see Lemma 7.7.

2. *Dirichlet-Neumann operator.* In section 2.4 we use a self-contained proof of the representation of the Dirichlet-Neumann operator $G(\eta)$ as a pseudo-differential operator, due to Baldi [5]. The conformal change of variables (2.121)-(2.122) transforms the elliptic problem (1.2), which is defined in the variable fluid domain $\{y \leq \eta(x)\}$, into the elliptic problem (2.128) which is defined on the straight strip $\{Y \leq 0\}$ and can be solved by an explicit integration. By conjugating back such solution, it turns out that (Lemma 2.31) the principal symbol of $G(\eta)$ is just $|D|$ (see (2.118)) up to a small remainder $\mathcal{R}_G(\eta) \in OPS^{-\infty}$ (recall that the profile $\eta \in \mathcal{C}^\infty$). Actually $\psi \mapsto \mathcal{R}_G(\eta)[\psi]$ is a regularizing linear operator which satisfies tame estimates (with loss of derivatives) in η, see e.g. (2.132). For obtaining such quantitative estimates it is convenient to represent $\mathcal{R}_G$ as an integral operator (see (2.129) and Lemma 2.113) and to use the fact an integral operator transforms into another integral operator under changes of variable, see Lemma 2.25.

Acknowledgements. We thank P. Baldi, L. Biasco, W. Craig and J. M. Delort, for many useful discussions. This research was supported by PRIN 2012 “Variational and perturbative aspects of nonlinear differential problems”.

Notations. We denote by $\mathbb{N} := \{0, 1, 2, \ldots\}$ the natural numbers including $\{0\}$ and $\mathbb{N}^+ := \{1, 2, \ldots\}$. We use the multi-index notation $k = (k_1, \ldots, k_{\nu+1}) \in \mathbb{N}^{\nu+1}$ with $|k| := |k_1| + \ldots + |k_{\nu+1}|$ and we denote the derivative $\partial^k_\lambda := \partial^{k_1}_{\lambda_1} \ldots \partial^{k_{\nu+1}}_{\lambda_{\nu+1}}$. The parameter $\lambda = (\omega, \kappa) \in \mathbb{R}^\nu \times [\kappa_1, \kappa_2]$.

Given a set A we denote $\mathcal{N}(A, \eta)$ the open neighborhood of A of width η (which is empty if A is empty) in $\mathbb{R}^\nu \times [\kappa_1, \kappa_2]$, namely

$$\mathcal{N}(A, \eta) := \left\{ \lambda \in \mathbb{R}^\nu \times [\kappa_1, \kappa_2] : \operatorname{dist}(A, \lambda) \leq \eta \right\} . \tag{1.41}$$

We denote the tangential sites by

$$\mathbb{S}^+ \subset \mathbb{N}^+ \quad \text{and we set} \quad \mathbb{S} := \mathbb{S}^+ \cup (-\mathbb{S}^+) \,, \quad \mathbb{S}_0 := \mathbb{S}_+ \cup (-\mathbb{S}_+) \cup \{0\} \subseteq \mathbb{Z} \,. \tag{1.42}$$

In the paper we shall use for a linear pseudo-differential operator the norms $| \, |^{k_0, \gamma}_{m, s, \alpha}$ introduced in Definition 2.7 indexed by $k_0 \in \mathbb{N}$, $\gamma \in (0, 1)$, $m \in \mathbb{R}$, $s \geq s_0$, $\alpha \in \mathbb{N}$. In order to help the reading we recall here their meaning:

1. The index $k_0 \in \mathbb{N}$ is *fixed* in section 3. It depends only on properties of the linear frequencies $\omega_j(\kappa)$ in (1.17) and it does *not* vary along the whole paper. It denotes that the operators, functions, frequencies, etc. are k_0-times differentiable with respect to the parameters $\lambda = (\omega, \kappa)$.
2. The parameter $\gamma \in (0, 1)$ is the diophantine constant of the frequencies $|\omega \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau}$, $\forall \ell \in \mathbb{Z}^\nu \setminus \{0\}$, and similarly for the first and second order Melnikov non-resonance conditions. Along the paper $\gamma = O(\varepsilon^a)$ with $a > 0$ as small as wanted (actually we could take just $\gamma = o(1)$ as $\varepsilon \to 0$).
3. The parameter $m \in \mathbb{R}$ denotes the order of a pseudo-differential operator $A \in OPS^m$.
4. The index s denotes the Sobolev index. It will vary in a finite scale $s \in [s_0, S]$ where s_0 is fixed in (1.20). The largest possible value of S is fixed in the Nash-Moser iteration in section 8, see (8.12).
5. The constant $\alpha \in \mathbb{N}$ is the number of ∂_ξ derivatives that we estimate of a symbol $a(x, \xi)$, see (2.36). In section 6 we take $\alpha \approx M$ where M is the number of decoupling steps performed in section 6.5. The constant M is fixed in (7.9). The important point is that the largest values of α, M used along the paper do not depend on the Sobolev index s.

The notation $a \leq_{s,\alpha,M} b$ means that $a \leq C(s, \alpha, M) b$ for some constant $C(s, \alpha, M) > 0$ depending on the Sobolev index s, and the constants α, M. The notation $a \lessdot b$ means that $a \leq Cb$ for some absolute constant which depends only on the data of the problem. Sometimes, along the paper, we omit to write the dependence $\leq_{s_0,k_0}$ with respect to s_0, k_0, because s_0 (defined in (1.20)) and k_0 (fixed in section 3) are considered as fixed constants.

For scalar valued functions $\mu : \Lambda_0 \subset \mathbb{R}^{\nu+1} \to \mathbb{R}$ (for example the Floquet exponents) we denote

$$|\mu|^{k_0,\gamma} := |\mu|^{k_0,\gamma}_{\Lambda_0} := \sum\nolimits_{|k| \leq k_0} \gamma^{|k|} \sup_{\lambda \in \Lambda_0} |\partial_\lambda^k \mu(\lambda)| \,.$$

We will often not specify the domain $\Lambda_0 \subset \mathbb{R}^{\nu+1}$ which is understood from the context.

2 Functional setting

We regard a function $u(\varphi, x) \in L^2(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{C})$ of space-time also as a φ-dependent family of functions $u(\varphi, \cdot) \in L^2(\mathbb{T}_x, \mathbb{C})$ that we expand in Fourier series as

$$u(\varphi, x) = \sum_{j' \in \mathbb{Z}} u_{j'}(\varphi) e^{\mathrm{i} j' x} = \sum_{\ell' \in \mathbb{Z}^\nu, j' \in \mathbb{Z}} u_{\ell', j'} e^{\mathrm{i}(\ell' \cdot \varphi + j' x)} \,. \tag{2.1}$$

Along the paper we denote the Fourier coefficients $u_{\ell,j}$, $u_j(\varphi)$ of the function $u(\varphi, x)$ (with respect to the space variables (φ, x) or x, respectively) also as $\widehat{u}_{\ell,j}$, $\widehat{u}_j(\varphi)$. We also consider real valued functions $u(\varphi, x) \in \mathbb{R}$. When no confusion appears we will denote simply by L^2, $L^2(\mathbb{T}^\nu \times \mathbb{T})$, $L^2_x := L^2(\mathbb{T}_x)$ either the spaces of real or complex valued L^2-functions.

The Sobolev norm $\| \ \|_s$ defined in (1.19) is equivalent to

$$\|u\|_s \simeq \|u\|_{H^s_\varphi L^2_x} + \|u\|_{L^2_\varphi H^s_x} \,. \tag{2.2}$$

Definition 2.1. *Given a function $u \in L^2(\mathbb{T}^\nu \times \mathbb{T})$ as in* (2.1)*, we define the majorant function*

$$\underline{u}(\varphi, x) := \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}} |u_{\ell,j}| e^{\mathrm{i}(\ell \cdot \varphi + jx)} \,. \tag{2.3}$$

Note that the Sobolev norms of u and $\underline{u}$ are the same, i.e.

$$\|u\|_s = \|\underline{u}\|_s \,. \tag{2.4}$$

We consider also family of Sobolev functions $\lambda \mapsto u(\lambda) \in H^s$ which are k_0-times differentiable with respect to a parameter

$$\lambda := (\omega, \kappa) \in \Lambda_0 \subset \mathbb{R}^{\nu+1} \,.$$

For $\gamma \in (0,1)$ we define the weighted Sobolev norm

$$\|u\|_s^{k_0,\gamma} := \sum_{|k|\le k_0} \gamma^{|k|} \sup_{\lambda\in\Lambda_0} \|\partial_\lambda^k u(\lambda)\|_s \,. \tag{2.5}$$

For a function $u(\lambda,\cdot) : \mathbb{T}^d \to \mathbb{C}$, we define the $\mathcal{C}^s$-weighted norm

$$\|u\|_{\mathcal{C}^s}^{k_0,\gamma} := \sum_{|k|\le k_0} \gamma^{|k|} \sup_{\lambda\in\Lambda_0} \|\partial_\lambda^k u(\lambda)\|_{\mathcal{C}^s} \tag{2.6}$$

(we use it in section 2.3 to functions $K(\lambda,\cdot)$ with $d=\nu+1$). We also introduce the smoothing operators

$$(\Pi_K u)(\varphi,x) := \sum_{|(\ell,j)|\le K} u_{\ell j} e^{\mathrm{i}(\ell\cdot\varphi+jx)}\,, \qquad \Pi_K^\perp := \mathrm{Id} - \Pi_K\,, \tag{2.7}$$

which satisfy the usual smoothing properties

$$\|\Pi_K u\|_{s+b}^{k_0,\gamma} \le K^b \|u\|_s^{k_0,\gamma}\,, \qquad \|\Pi_K^\perp u\|_s^{k_0,\gamma} \le K^{-b}\|u\|_{s+b}^{k_0,\gamma}\,, \qquad \forall s,b\ge 0\,. \tag{2.8}$$

We have the following interpolation lemma.

Lemma 2.1. *Let $a_0, b_0 \ge 0$ and $p,q>0$. For all $\epsilon>0$ there exists a constant $C(\epsilon) := C(\epsilon,p,q) > 0$, which satisfies $C(1)<1$, such that*

$$\|u\|_{a_0+p}\|v\|_{b_0+q} \le \epsilon \|u\|_{a_0+p+q}\|v\|_{b_0} + C(\epsilon)\|u\|_{a_0}\|v\|_{b_0+p+q} \tag{2.9}$$

$$\|u\|_{a_0+p}^{k_0,\gamma}\|v\|_{b_0+q} \le \epsilon \|u\|_{a_0+p+q}^{k_0,\gamma}\|v\|_{b_0} + C(\epsilon)\|u\|_{a_0}^{k_0,\gamma}\|v\|_{b_0+p+q}\,. \tag{2.10}$$

Proof. By interpolation

$$\|u\|_{a_0+p} \le \|u\|_{a_0}^{\mu}\|u\|_{a_0+p+q}^{1-\mu}\,,\ \mu := \frac{q}{p+q}\,, \quad \|v\|_{b_0+q} \le \|v\|_{b_0}^{\eta}\|v\|_{b_0+p+q}^{1-\eta}\,,\ \eta := \frac{p}{p+q}\,.$$

Hence, noting that $\eta+\mu=1$, we have

$$\|u\|_{a_0+p}\|v\|_{b_0+q} \le (\|u\|_{a_0+p+q}\|v\|_{b_0})^{\eta}(\|u\|_{a_0}\|v\|_{b_0+p+q})^{\mu}\,.$$

By the asymmetric Young inequality we get, for any $\epsilon>0$,

$$\|u\|_{a_0+p}\|v\|_{b_0+q} \le \epsilon\|u\|_{a_0+p+q}\|v\|_{b_0} + C(\epsilon,p,q)\|u\|_{a_0}\|v\|_{b_0+p+q}$$

where $C(\epsilon,p,q) := \mu(\eta/\epsilon)^{\frac{\eta}{\mu}} = \frac{q}{p+q}\big(\frac{p}{\epsilon(p+q)}\big)^{p/q}$. Note that for $\epsilon=1$ the constant $C(1,p,q)<1$.

The estimate (2.10) follows by (2.9) recalling (2.5). ☐

Linear operators. Let $A : \mathbb{T}^\nu \mapsto \mathcal{L}(L^2(\mathbb{T}_x))$, $\varphi \mapsto A(\varphi)$, be a φ-dependent family of linear operators acting on $L^2(\mathbb{T}_x)$. We regard A also as an operator (that for simplicity we denote by A as well) which acts on functions $u(\varphi,x)$ of space-time, i.e. we consider the operator $A \in \mathcal{L}(L^2(\mathbb{T}^\nu\times\mathbb{T}))$ defined by

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

We say that an operator A is *real* if it maps real valued functions into real valued functions.

We represent a real operator acting on $(\eta,\psi)\in L^2(\mathbb{T}^{\nu+1},\mathbb{R}^2)$ by a matrix

$$\mathcal{R}\begin{pmatrix}\eta\\ \psi\end{pmatrix} = \begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}\eta\\ \psi\end{pmatrix} \tag{2.11}$$

where A,B,C,D are real operators acting on the scalar valued components $\eta,\psi\in L^2(\mathbb{T}^{\nu+1},\mathbb{R})$.

The action of an operator $A\in\mathcal{L}(L^2(\mathbb{T}^\nu\times\mathbb{T}))$ on a function u as in (2.1) is

$$Au(\varphi,x) = \sum_{j,j'\in\mathbb{Z}} A_j^{j'}(\varphi)u_{j'}(\varphi)e^{\mathrm{i}jx} = \sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}}\ \sum_{\ell'\in\mathbb{Z}^\nu, j'\in\mathbb{Z}} A_j^{j'}(\ell-\ell')u_{\ell',j'}e^{\mathrm{i}(\ell\cdot\varphi+jx)}\,. \tag{2.12}$$

We shall identify an operator A with the matrix $\big(A_j^{j'}(\ell-\ell')\big)_{j,j'\in\mathbb{Z},\ell,\ell'\in\mathbb{Z}^\nu}$.

Note that the differentiated operator $\partial_{\varphi_m}A(\varphi)$, $m=1,\dots,\nu$, is represented by the matrix elements $\mathrm{i}(\ell_m-\ell'_m)A_j^{j'}(\ell-\ell')$, and the commutator $[\partial_x,A]:=\partial_x\circ A-A\circ\partial_x$ is represented by the matrix with entries $\mathrm{i}(j-j')A_j^{j'}(\ell-\ell')$.

Definition 2.2. *Given a linear operator A as in* (2.12) *we define the operator*

1. $|A|$ **(majorant operator)** *whose matrix elements are* $|A_j^{j'}(\ell-\ell')|$,
2. $\Pi_N A$, $N\in\mathbb{N}$ **(smoothed operator)** *whose matrix elements are*

$$(\Pi_N A)_j^{j'}(\ell-\ell'):=\begin{cases}A_j^{j'}(\ell-\ell') & \text{if} \quad |\ell-\ell'|\le N\\ 0 & \text{otherwise}\,.\end{cases} \tag{2.13}$$

   We also denote $\Pi_N^\perp:=\mathrm{Id}-\Pi_N$,

3. $\langle\partial_\varphi\rangle^b A$, $b\in\mathbb{R}$, *whose matrix elements are* $\langle\ell-\ell'\rangle^b A_j^{j'}(\ell-\ell')$.

Lemma 2.2. *Given linear operators A, B we have*

$$\| |A+B|u\|_s\le \| |A|\, [\![u]\!]\|_s+\| |B|\, [\![u]\!]\|_s\,,\quad \| |AB|u\|_s\le \| |A||B|\, [\![u]\!]\|_s\,. \tag{2.14}$$

Proof. The first inequality in (2.14) follows by

$$\| |A+B|u\|_s^2\le\sum_{\ell,j}\langle\ell,j\rangle^{2s}\Big(\sum_{\ell',j'}|A_j^{j'}(\ell-\ell')||u_{\ell',j'}|+|B_j^{j'}(\ell-\ell')||u_{\ell',j'}|\Big)^2=\Big\||A|[\![u]\!]+|B|[\![u]\!]\Big\|_s^2\,.$$

The second inequality in (2.14) follows by

$$\begin{aligned}\| |AB|u\|_s^2&\le\sum_{\ell,j}\langle\ell,j\rangle^{2s}\Big(\sum_{\ell',j'}|(AB)_j^{j'}(\ell-\ell')||u_{\ell',j'}|\Big)^2\\&=\sum_{\ell,j}\langle\ell,j\rangle^{2s}\Big(\sum_{\ell',j'}\Big|\sum_{\ell_1,j_1}A_j^{j_1}(\ell-\ell_1)B_{j_1}^{j'}(\ell_1-\ell')\Big||u_{\ell',j'}|\Big)^2\\&\le\sum_{\ell,j}\langle\ell,j\rangle^{2s}\Big(\sum_{\ell_1,j_1}|A_j^{j_1}(\ell-\ell_1)|\sum_{\ell',j'}|B_{j_1}^{j'}(\ell_1-\ell')||u_{\ell',j'}|\Big)^2\\&=\sum_{\ell,j}\langle\ell,j\rangle^{2s}\Big(\sum_{\ell_1,j_1}|A_j^{j_1}(\ell-\ell_1)|\big(\widehat{|B|[\![u]\!]}\big)_{\ell_1,j_1}\Big)^2=\| |A|\big(|B|[\![u]\!]\big)\|_s^2\,.\end{aligned}$$

The lemma is proved. □

Definition 2.3. (Even operator) *A linear operator A as in* (2.12) *is* EVEN *if each $A(\varphi)$, $\varphi\in\mathbb{T}^\nu$, leaves invariant the space of functions even in x.*

Since the Fourier coefficients of an even function satisfy $u_{-j}=u_j$, $\forall j\in\mathbb{Z}$, we have that

$$A \text{ is even} \quad\Longleftrightarrow\quad \forall\varphi\in\mathbb{T}^\nu\,,\ A_j^{j'}(\varphi)=A_{-j}^{-j'}(\varphi),\ \forall j,j'\in\mathbb{Z}\,. \tag{2.15}$$

Definition 2.4. (Reversibility) *An operator $\mathcal{R}$ as in* (2.11) *is*

1. REVERSIBLE *if $\mathcal{R}(-\varphi)\circ\rho=-\rho\circ\mathcal{R}(\varphi)$, $\forall\varphi\in\mathbb{T}^\nu$, where the involution ρ is defined in* (1.11),
2. REVERSIBILITY PRESERVING *if $\mathcal{R}(-\varphi)\circ\rho=\rho\circ\mathcal{R}(\varphi)$, $\forall\varphi\in\mathbb{T}^\nu$.*

Conjugating the linear operator $\mathcal{L} := \omega \cdot \partial_\varphi + A(\varphi)$ by a family of invertible linear maps $\Phi(\varphi)$ we get the transformed operator

$$\mathcal{L}_+ := \Phi^{-1}(\varphi)\mathcal{L}\Phi(\varphi) = \omega \cdot \partial_\varphi + A_+(\varphi)\,, \quad A_+(\varphi) := \Phi^{-1}(\varphi)(\omega \cdot \partial_\varphi \Phi(\varphi)) + \Phi^{-1}(\varphi)A(\varphi)\Phi(\varphi)\,.$$

It results that the conjugation of an even and reversible operator with an operator $\Phi(\varphi)$ which is even and reversibility preserving is even and reversible. An operator $\mathcal{R}$ as in (2.11) is

1. reversible if and only if $\varphi \mapsto A(\varphi), D(\varphi)$ are odd and $\varphi \mapsto B(\varphi), C(\varphi)$ are even.
2. reversibility preserving if and only if $\varphi \mapsto A(\varphi), D(\varphi)$ are even and $\varphi \mapsto B(\varphi), C(\varphi)$ are odd.

From section 6.3 on, it is convenient to consider a real operator $\mathcal{R}$ as in (2.11), which acts on the real variables $(\eta, \psi) \in \mathbb{R}^2$, as a linear operator which acts on the complex variables

$$u := \eta + \mathrm{i}\psi\,, \quad \bar{u} := \eta - \mathrm{i}\psi\,, \qquad i.e.\;\; \eta = (u + \bar{u})/2\,, \quad \psi = (u - \bar{u})/(2\mathrm{i})\,. \tag{2.16}$$

We get that a *real* operator acting in the complex coordinates $(u, \bar{u})$ has the form

$$\mathbf{R} := \begin{pmatrix} \mathcal{R}_1 & \mathcal{R}_2 \\ \overline{\mathcal{R}}_2 & \overline{\mathcal{R}}_1 \end{pmatrix}\,, \quad \mathcal{R}_1 := \frac{1}{2}\big\{(A + D) - \mathrm{i}(B - C)\big\}\,, \quad \mathcal{R}_2 := \frac{1}{2}\big\{(A - D) + \mathrm{i}(B + C)\big\} \tag{2.17}$$

where the operator $\overline{A}$ is defined by

$$\overline{A}(u) := \overline{A(\bar{u})}\,. \tag{2.18}$$

It holds $\overline{AB} = \overline{A}\,\overline{B}$.

The composition of real operators is another real operator.

A real operator $\mathbf{R}$ as in (2.17) is even if the operators $\mathcal{R}_1$, $\mathcal{R}_2$ are even.

In the complex coordinates (2.16) the involution ρ defined in (1.11) is the map $u \mapsto \bar{u}$. Thus

Lemma 2.3. *The real operator* $\mathbf{R}$ *in* (2.17) *is*

1. *reversible if and only if* $\mathcal{R}_1(-\varphi) = -\overline{\mathcal{R}_1}(\varphi)$, $\mathcal{R}_2(-\varphi) = -\overline{\mathcal{R}_2}(\varphi)$, $\forall \varphi \in \mathbb{T}^\nu$,
2. *reversibility preserving if and only if* $\mathcal{R}_1(-\varphi) = \overline{\mathcal{R}_1}(\varphi)$, $\mathcal{R}_2(-\varphi) = \overline{\mathcal{R}_2}(\varphi)$, $\forall \varphi \in \mathbb{T}^\nu$.

2.1 Pseudo-differential operators and norms

Pseudo-differential operators on the torus may be seen as a particular case (see Definition 2.6) of pseudo-differential operators on $\mathbb{R}^n$, as developed for example in [27]. It is also convenient to define them also through Fourier series, see Definition 2.5, for which we refer to [42].

Given a function $a : \mathbb{Z} \to \mathbb{C}$ we denote the discrete derivative by $(\Delta_j a)(j) := a(j+1) - a(j)$. For $\beta \in \mathbb{N}$ we denote by $\Delta_j^\beta := \Delta_j \circ \ldots \circ \Delta_j$ the composition of β-discrete derivatives.

Definition 2.5. (ΨDO1) *Let* $u = \sum_{j \in \mathbb{Z}} u_j e^{\mathrm{i}jx}$. *A linear operator* A *defined by*

$$(Au)(x) := \sum\nolimits_{j \in \mathbb{Z}} a(x, j) u_j e^{\mathrm{i}jx} \tag{2.19}$$

is called pseudo-differential of order $\leq m$ *if its symbol* $a(x, j)$ *is* 2π*-periodic and* $\mathcal{C}^\infty$*-smooth in* x*, and satisfies the inequalities*

$$\big|\partial_x^\alpha \Delta_j^\beta a(x, j)\big| \leq C_{\alpha,\beta} \langle j \rangle^{m-\beta}\,, \quad \forall \alpha, \beta \in \mathbb{N}\,. \tag{2.20}$$

We also remark that, given an operator A, we recover its symbol by

$$a(x, j) = e^{-\mathrm{i}jx}(A[e^{\mathrm{i}jx}]). \tag{2.21}$$

When the symbol $a(x)$ is independent of j, the operator $A = \mathrm{Op}(a)$ is the multiplication operator for the function $a(x)$, i.e $A : u(x) \mapsto a(x)u(x)$. In such a case we shall also denote $A = \mathrm{Op}(a) = a(x)$.

Definition 2.6. (ΨDO2) *A linear operator A is called pseudo-differential of order $\leq m$ if its symbol $a(x,j)$ is the restriction to $\mathbb{R}\times\mathbb{Z}$ of a function $a(x,\xi)$ which is $\mathcal{C}^\infty$-smooth on $\mathbb{R}\times\mathbb{R}$, 2π-periodic in x, and satisfies the inequalities*

$$\big|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)\big| \leq C_{\alpha,\beta}\langle \xi\rangle^{m-\beta}\,, \quad \forall \alpha,\beta\in\mathbb{N}\,. \tag{2.22}$$

We call $a(x,\xi)$ the symbol of the operator A, that we denote

$$A = \mathrm{Op}(a) = a(x,D)\,, \quad D := D_x := \frac{1}{\mathrm{i}}\partial_x\,.$$

We denote by S^m the class of all the symbols $a(x,\xi)$ satisfying (2.22)*, and by OPS^m the set of pseudo-differential operators of order m. We set $OPS^{-\infty} := \cap_{m\in\mathbb{R}} OPS^m$.*

Definitions 2.5 and 2.6 are equivalent because any discrete symbol $a:\mathbb{R}\times\mathbb{Z}\to\mathbb{C}$ satisfying (2.20) can be extended to a $\mathcal{C}^\infty$-symbol $\widetilde{a}:\mathbb{R}\times\mathbb{R}\to\mathbb{C}$ satisfying (2.22), see section 7.2 in [42]. It is sufficient to proceed as follows. Given a function $\sigma:\mathbb{Z}\to\mathbb{C}$ we define the $\mathcal{C}^\infty$-extension

$$\widetilde{\sigma}:\mathbb{R}\to\mathbb{C}\,, \quad \widetilde{\sigma}(\xi) := \sum\nolimits_{j\in\mathbb{Z}} \sigma(j)\zeta(\xi-j)\,, \quad \forall\xi\in\mathbb{R}\,, \tag{2.23}$$

where $\zeta := \widehat{\theta}\in\mathcal{S}(\mathbb{R})$ (Schwartz class) is the Fourier transform of a function $\theta\in\mathcal{D}(\mathbb{R})$ (test functions) such that $\mathrm{supp}(\theta)\subset[-2/3,2/3]$, $\theta(x)+\theta(x-1)=1$, $\forall x\in[0,1]$, and $\sum_{j\in\mathbb{Z}}\theta(x+j)=1$. It results that $\zeta(k)=\delta_{0k}$, $\forall k\in\mathbb{Z}$, namely $\zeta(0)=1$ and $\zeta(k)=0$, $\forall k\neq 0$, so that $\widetilde{\sigma}(k)=\sigma(k)$, $\forall k\in\mathbb{Z}$. Moreover there are positive constants $c'_\beta>0$, independent of σ, such that (see Lemma 7.1.1 in [42])

$$|\Delta_j^\beta \sigma(j)| \leq c_\beta\langle j\rangle^{m-\beta} \quad\Longleftrightarrow\quad |\partial_\xi^\beta\widetilde{\sigma}(\xi)| \leq c'_\beta c_\beta\langle\xi\rangle^{m-\beta}\,. \tag{2.24}$$

Definition 2.6 is more convenient to get basic results concerning composition, asymptotic expansions, ... of pseudo-differential operators, that we recall below. We underline that, in the sequel, also when we use of the continuous symbol $a(x,\xi)$, we think $\mathrm{Op}(a)$ to act only on 2π-periodic functions $u(x)$ as in (2.19).

We shall use the following notation, used also in [1]. For any $m\in\mathbb{R}\setminus\{0\}$, we set

$$|D|^m := \mathrm{Op}(\chi(\xi)|\xi|^m)\,, \tag{2.25}$$

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

$$\chi(\xi) = \begin{cases} 0 & \text{if} \quad |\xi|\leq\frac13 \\ 1 & \text{if} \quad |\xi|\geq\frac23\,, \end{cases} \qquad \partial_\xi\chi(\xi)>0 \quad \forall\xi\in\Big(\frac13,\frac23\Big)\,. \tag{2.26}$$

Lemma 2.4. *The pseudo-differential operator $A:=\mathrm{Op}(a)$ is*

1. *even if and only if the symbol $a(-x,-\xi)=a(x,\xi)$ is even,*
2. *real if and only if the symbol $\overline{a(x,-\xi)}=a(x,\xi)$.*
3. *The operator $\overline{A}$ defined in* (2.18) *is pseudo-differential with symbol $\overline{a(x,-\xi)}$.*

We first recall some fundamental properties of pseudo-differential operators.

Composition of pseudo-differential operators. If $A=a(x,D)\in OPS^m$, $B=b(x,D)\in OPS^{m'}$, $m,m'\in\mathbb{R}$, are pseudo-differential operators with symbols $a\in S^m$, $b\in S^{m'}$ then the composition operator $AB := A\circ B = \sigma_{AB}(x,D)$ is a pseudo-differential operator with symbol

$$\sigma_{AB}(x,\xi) = \sum_{j\in\mathbb{Z}} a(x,\xi+j)\widehat{b}(j,\xi)e^{\mathrm{i}jx} = \sum_{j,j'\in\mathbb{Z}} \widehat{a}(j'-j,\xi+j)\widehat{b}(j,\xi)e^{\mathrm{i}j'x} \tag{2.27}$$

where $\widehat{\cdot}$ denotes the Fourier coefficients of the symbols $a(x,\xi)$ and $b(x,\xi)$ with respect to x. The symbol σ_{AB} has the following asymptotic expansion

$$\sigma_{AB}(x,\xi) \sim \sum\nolimits_{\beta\geq 0} \frac{1}{\mathrm{i}^\beta \beta!} \partial_\xi^\beta a(x,\xi) \partial_x^\beta b(x,\xi) \,, \tag{2.28}$$

that is, $\forall N \geq 1$,

$$\sigma_{AB}(x,\xi) = \sum_{\beta=0}^{N-1} \frac{1}{\beta! \mathrm{i}^\beta} \partial_\xi^\beta a(x,\xi)\, \partial_x^\beta b(x,\xi) + r_N(x,\xi) \qquad \text{where} \qquad r_N := r_{N,AB} \in S^{m+m'-N} \,. \tag{2.29}$$

The remainder r_N has the explicit formula

$$r_N(x,\xi) := \frac{1}{(N-1)!\, \mathrm{i}^N} \int_0^1 (1-\tau)^{N-1} \sum_{j\in\mathbb{Z}} (\partial_\xi^N a)(x, \xi+\tau j) \widehat{\partial_x^N b}(j,\xi) e^{\mathrm{i} jx} \,. \tag{2.30}$$

Adjoint of a pseudo-differential operator. If $A = a(x,D) \in OPS^m$ is a pseudo-differential operator with symbol $a \in S^m$, then its L^2-adjoint is the pseudo-differential operator

$$A^* = \mathrm{Op}(a^*) \qquad \text{with symbol} \qquad a^*(x,\xi) := \overline{\sum\nolimits_{j\in\mathbb{Z}} \widehat{a}(j, \xi - j) e^{\mathrm{i} jx}} \,. \tag{2.31}$$

Families of pseudo-differential operators. We consider φ-dependent families of pseudo-differential operators

$$(Au)(\varphi, x) = \sum\nolimits_{j\in\mathbb{Z}} a(\varphi, x, j) u_j(\varphi) e^{\mathrm{i} jx} \tag{2.32}$$

where the symbol $a(\varphi, x, \xi)$ is $\mathcal{C}^\infty$-smooth also in φ. We still denote $A := A(\varphi) = \mathrm{Op}(a(\varphi,\cdot)) = \mathrm{Op}(a)$.

By (2.27) and a Fourier expansion also in $\varphi \in \mathbb{T}^\nu$, the symbol of the composition operator AB is

$$\sigma_{AB}(\varphi,x,\xi) = \sum_{j\in\mathbb{Z}} a(\varphi,x,\xi+j) \widehat{b}(\varphi,j,\xi) e^{\mathrm{i} jx} = \sum_{\substack{j',j\in\mathbb{Z} \\ \ell,\ell_1\in\mathbb{Z}^\nu}} \widehat{a}(\ell-\ell_1, j'-j, \xi+j) \widehat{b}(\ell_1, j, \xi) e^{\mathrm{i}(\ell\cdot\varphi + j'x)} \,. \tag{2.33}$$

Similarly by (2.31) the symbol of the adjoint operator $A(\varphi)^* = \mathrm{Op}(a^*(\varphi,\cdot))$ is

$$a^*(\varphi,x,\xi) = \overline{\sum_{j\in\mathbb{Z}} \widehat{a}(\varphi, j, \xi-j) e^{\mathrm{i} jx}} = \overline{\sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \widehat{a}(\ell, j, \xi-j) e^{\mathrm{i}(\ell\cdot\varphi + jx)}} \,. \tag{2.34}$$

Along the paper we also consider families of pseudo-differential operators $A(\lambda) := \mathrm{Op}(a(\lambda,\varphi,x,\xi))$ which are k_0-times differentiable with respect to a parameter $\lambda := (\omega,\kappa) \in \Lambda_0 = \Omega_0 \times [\kappa_1,\kappa_2] \subset \mathbb{R}^\nu \times [\kappa_1,\kappa_2]$, where the regularity constant $k_0 \in \mathbb{N}$ is fixed once for all in section 3. Note that for any

$$\partial_\lambda^k A = \mathrm{Op}(\partial_\lambda^k a) \,, \quad \forall k \in \mathbb{N}^{\nu+1} \,, \ |k| \leq k_0 \,.$$

We now introduce a norm (inspired to Metivier [38], chapter 5) which controls the regularity in (φ,x), and the decay in ξ, of the symbol $a(\varphi,x,\xi) \in S^m$, together with its derivatives $\partial_\xi^\beta a \in S^{m-\beta}$, $0 \leq \beta \leq \alpha$, in the Sobolev norm $\| \ \|_s$.

Definition 2.7. (Weighted ΨDO norm) *Let $A(\lambda) := a(\lambda,\varphi,x,D) \in OPS^m$ be a family of pseudo-differential operators with symbol $a(\lambda,\varphi,x,\xi) \in S^m$, $m \in \mathbb{R}$, which are k_0-times differentiable with respect to $\lambda \in \Lambda_0 \subset \mathbb{R}^{\nu+1}$. For $\gamma \in (0,1)$, $\alpha \in \mathbb{N}$, $s \geq 0$, we define the weighted norm*

$$|A|_{m,s,\alpha}^{k_0,\gamma} := \sum_{|k|\leq k_0} \gamma^{|k|} \sup\nolimits_{\lambda\in\Lambda_0} |\partial_\lambda^k A(\lambda)|_{m,s,\alpha} \tag{2.35}$$

where we use the multi-index notation $k = (k_1,\ldots,k_{\nu+1}) \in \mathbb{N}^{\nu+1}$ with $|k| := |k_1| + \ldots + |k_{\nu+1}|$, and

$$|A|_{m,s,\alpha} := \max\nolimits_{0\leq\beta\leq\alpha} \sup_{\xi\in\mathbb{R}} \|\partial_\xi^\beta a(\lambda,\cdot,\cdot,\xi)\|_s \langle\xi\rangle^{-m+\beta} \,. \tag{2.36}$$

For each k_0, γ, m fixed, the norm (2.35) is non-decreasing both in s and α, namely

$$\forall s \leq s', \, \alpha \leq \alpha', \qquad | \, |_{m,s,\alpha}^{k_0,\gamma} \leq | \, |_{m,s',\alpha}^{k_0,\gamma}, \ | \, |_{m,s,\alpha}^{k_0,\gamma} \leq | \, |_{m,s,\alpha'}^{k_0,\gamma}. \tag{2.37}$$

Note also that the norm (2.35) is non-increasing in m, i.e.

$$m \leq m' \qquad \Longrightarrow \qquad | \, |_{m',s,\alpha}^{k_0,\gamma} \leq | \, |_{m,s,\alpha}^{k_0,\gamma}. \tag{2.38}$$

Given a function $a(\lambda, \varphi, x) \in \mathcal{C}^\infty$ which is k_0-times differentiable with respect to λ, the weighted norm of the corresponding multiplication operator is

$$|\mathrm{Op}(a)|_{0,s,\alpha}^{k_0,\gamma} = \|a\|_s^{k_0,\gamma}, \quad \forall \alpha \in \mathbb{N}, \tag{2.39}$$

where the weighted Sobolev norm $\|a\|_s^{k_0,\gamma}$ is defined in (2.5).

For a Fourier multiplier $g(D)$ with symbol $g \in S^m$, we simply have

$$|g(D)|_{m,s,\alpha} \leq C(m, \alpha, g), \quad \forall s \geq 0. \tag{2.40}$$

The norm $| \, |_{0,s,0}$ controls the action of a pseudo-differential operator on the Sobolev spaces H^s as we shall prove in Lemma 2.13.

Remark 2.5. The norm of Definition 2.7 is introduced in view of section 6.6 where we have to estimate the norm $|R_M|_{1-\frac{M}{2},s,0}^{k_0,\gamma}$ in (6.192). The remainder R_M depends on $|\mathrm{Op}(q_M)|_{1-\frac{M}{2},s,0}^{k_0,\gamma}$. The terms $q_1, \ldots, q_M$ are obtained iteratively, and each q_{k+1} depends on $\partial_\xi q_k$. Thus we need to control the Sobolev norm in (φ, x) of $\partial_\xi^M q_0$. This is made precise by estimating the norm $|\mathrm{Op}(q_0)|_{-\frac{3}{2},s,M}^{k_0,\gamma}$. □

The norm $| \, |_{m,s,\alpha}^{k_0,\gamma}$ is closed under composition and satisfies tame estimates.

Lemma 2.6. (Composition) *Let $A = a(\lambda, \varphi, x, D)$, $B = b(\lambda, \varphi, x, D)$ be pseudo-differential operators with symbols $a(\lambda, \varphi, x, \xi) \in S^m$, $b(\lambda, \varphi, x, \xi) \in S^{m'}$, $m, m' \in \mathbb{R}$. Then $A(\lambda) \circ B(\lambda) \in OPS^{m+m'}$ satisfies, for all $\alpha \in \mathbb{N}$, $s \geq s_0$,*

$$|AB|_{m+m',s,\alpha}^{k_0,\gamma} \leq_{m,\alpha,k_0} C(s)|A|_{m,s,\alpha}^{k_0,\gamma}|B|_{m',s_0+\alpha+|m|,\alpha}^{k_0,\gamma} + C(s_0)|A|_{m,s_0,\alpha}^{k_0,\gamma}|B|_{m',s+\alpha+|m|,\alpha}^{k_0,\gamma}. \tag{2.41}$$

Moreover, for any integer $N \geq 1$, the remainder $R_N := \mathrm{Op}(r_N)$ in (2.29) satisfies

$$\begin{aligned} |R_N|_{m+m'-N,s,\alpha}^{k_0,\gamma} \leq_{m,N,\alpha,k_0} \frac{1}{N!}\big(&C(s)|A|_{m,s,N+\alpha}^{k_0,\gamma}|B|_{m',s_0+2N+|m|+\alpha,\alpha}^{k_0,\gamma} \\ &+ C(s_0)|A|_{m,s_0,N+\alpha}^{k_0,\gamma}|B|_{m',s+2N+|m|+\alpha,\alpha}^{k_0,\gamma}\big). \end{aligned} \tag{2.42}$$

Both (2.41)-(2.42) *hold with the constant $C(s_0)$ interchanged with $C(s)$.*

Proof. As a first step we prove the estimates with no dependence on λ:

$$|AB|_{m+m',s,\alpha} \leq_{m,\alpha} C(s)|A|_{m,s,\alpha}|B|_{m',s_0+\alpha+|m|,\alpha} + C(s_0)|A|_{m,s_0,\alpha}|B|_{m',s+\alpha+|m|,\alpha} \tag{2.43}$$

$$|R_N|_{m+m'-N,s,\alpha} \leq_{m,N,s,\alpha} \frac{1}{N!}\big(|A|_{m,s,N+\alpha}|B|_{m',s_0+2N+|m|+\alpha,\alpha} + |A|_{m,s_0,N+\alpha}|B|_{m',s+2N+|m|+\alpha,\alpha}\big). \tag{2.44}$$

We first prove (2.43) for $\alpha = 0$. Denote by $\sigma := \sigma_{AB}$ the symbol in (2.33). For all $\xi \in \mathbb{R}$ we have

$$\|\sigma(\cdot, \xi)\|_s^2 \langle \xi \rangle^{-2(m+m')} = \sum_{j',\ell} \langle \ell, j' \rangle^{2s} \Big| \sum_{j,\ell_1} \widehat{a}(\ell - \ell_1, j' - j, \xi + j) \widehat{b}(\ell_1, j, \xi) \Big|^2 \langle \xi \rangle^{-2(m+m')} \leq S_1 + S_2 \tag{2.45}$$

where

$$S_1 := \sum_{j',\ell} \Big(\sum_{\langle \ell, j' \rangle \leq 2^{1/s} \langle \ell_1, j \rangle} |\widehat{a}(\ell - \ell_1, j' - j, \xi + j)| \langle \ell - \ell_1, j' - j \rangle^{s_0} |\widehat{b}(\ell_1, j, \xi)| \frac{\langle \ell_1, j \rangle^s \langle \ell, j' \rangle^s}{\langle \ell_1, j \rangle^s \langle \ell - \ell_1, j' - j \rangle^{s_0}} \Big)^2 \langle \xi \rangle^{-2(m+m')}$$

$$S_2 := \sum_{j',\ell} \Big(\sum_{\langle \ell, j' \rangle > 2^{1/s} \langle \ell_1, j \rangle} |\widehat{a}(\ell - \ell_1, j' - j, \xi + j)| \langle \ell - \ell_1, j' - j \rangle^{s} |\widehat{b}(\ell_1, j, \xi)| \frac{\langle \ell_1, j \rangle^{s_0} \langle \ell, j' \rangle^s}{\langle \ell_1, j \rangle^{s_0} \langle \ell - \ell_1, j' - j \rangle^{s}} \Big)^2 \langle \xi \rangle^{-2(m+m')}.$$

Now, by Cauchy-Schwartz inequality and denoting $\zeta(s_0) := \sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \frac{1}{\langle \ell, j\rangle^{2s_0}}$, we get

$$
\begin{aligned}
S_1 &\leq \sum_{j',\ell} \Big(\sum_{\langle \ell, j'\rangle \leq 2^{1/s}\langle \ell_1, j\rangle} |\widehat{a}(\ell-\ell_1, j'-j, \xi+j)| \langle \ell-\ell_1, j'-j\rangle^{s_0} |\widehat{b}(\ell_1, j, \xi)| \frac{\langle \ell_1, j\rangle^s \, 2}{\langle \ell-\ell_1, j'-j\rangle^{s_0}} \Big)^2 \langle \xi\rangle^{-2(m+m')} \\
&\leq 4\zeta(s_0) \sum_{j',\ell} \sum_{\ell_1, j} |\widehat{a}(\ell-\ell_1, j'-j, \xi+j)|^2 \langle \ell-\ell_1, j'-j\rangle^{2s_0} |\widehat{b}(\ell_1, j, \xi)|^2 \langle \ell_1, j\rangle^{2s} \ \langle \xi\rangle^{-2(m+m')} \\
&\leq 4\zeta(s_0) \sum_{\ell_1, j} |\widehat{b}(\ell_1, j, \xi)|^2 \langle \ell_1, j\rangle^{2s} \ \langle \xi\rangle^{-2m'} \sum_{j',\ell} |\widehat{a}(\ell-\ell_1, j'-j, \xi+j)|^2 \langle \ell-\ell_1, j'-j\rangle^{2s_0} \langle \xi\rangle^{-2m} . \quad (2.46)
\end{aligned}
$$

For each j, ℓ_1 fixed, we apply Peetre's inequality

$$
\langle \xi+\eta\rangle^m \leq C_m \langle \xi\rangle^m \langle \eta\rangle^{|m|}, \quad \forall m\in\mathbb{R},\ \eta\in\mathbb{R}, \xi\in\mathbb{R} \qquad (2.47)
$$

(where $C_m = 4^{|m|}$) with $\eta = j$, and we estimate, for any $s \geq s_0$,

$$
\begin{aligned}
\sup_\xi \sum_{j',\ell} |\widehat{a}(\ell-\ell_1, j'-j, \xi+j)|^2 \langle \ell-\ell_1, j'-j\rangle^{2s} \ \langle \xi\rangle^{-2m} &= \sup_\xi \|a(\cdot, \xi+j)\|_s^2 \langle \xi\rangle^{-2m} \\
= \big(\sup_\xi \|a(\cdot, \xi+j)\|_s^2 \langle \xi+j\rangle^{-2m}\big) \frac{\langle \xi+j\rangle^{2m}}{\langle \xi\rangle^{2m}} &\leq C_m^2 |A|_{m,s,0}^2 \langle j\rangle^{2|m|} \qquad (2.48)
\end{aligned}
$$

and therefore we get, by (2.46) and (2.48) for $s = s_0$,

$$
S_1 \leq 4\zeta(s_0) C_m^2 |A|_{m,s_0,0}^2 \sum_{\ell_1, j} |\widehat{b}(\ell_1, j, \xi)|^2 \langle \ell_1, j\rangle^{2s} \langle j\rangle^{2|m|} \ \langle \xi\rangle^{-2m'} \leq 4\zeta(s_0) C_m^2 |A|_{m,s_0,0}^2 |B|_{m',s+|m|,0}^2 . \quad (2.49)
$$

For the estimate of S_2 note that, since the indices satisfy $\langle \ell, j'\rangle > 2^{1/s}\langle \ell_1, j\rangle$ we have $\langle \ell, j'\rangle \leq \langle \ell_1, j\rangle + \langle \ell-\ell_1, j'-j\rangle \leq 2^{-1/s}\langle \ell, j'\rangle + \langle \ell-\ell_1, j'-j\rangle$ and therefore

$$
\langle \ell, j'\rangle \leq \big(1 - 2^{-1/s}\big)^{-1} \langle \ell-\ell_1, j'-j\rangle .
$$

As a consequence, arguing as above, we deduce that, for some constant $C(s) > 0$, we have

$$
S_2 \leq_m C(s) |A|_{m,s,0}^2 |B|_{m',s_0+|m|,0}^2 . \qquad (2.50)
$$

By (2.45) and (2.49), (2.50) we deduce the estimate (2.43) for $\alpha = 0$, i.e.

$$
|AB|_{m+m',s,0} \leq_m C(s) |A|_{m,s,0} |B|_{m',s_0+|m|,0} + C(s_0) |A|_{m,s_0,0} |B|_{m',s+|m|,0} . \qquad (2.51)
$$

Now we prove (2.43) for $\alpha \geq 1$. By differentiating (2.33) we get, for all $1 \leq \beta \leq \alpha$,

$$
\partial_\xi^\beta \sigma_{AB}(\varphi, x, \xi) = \sum_{\beta_1+\beta_2=\beta} C(\beta_1, \beta_2) \sum_{j\in\mathbb{Z}} \partial_\xi^{\beta_1} a(\varphi, x, \xi+j) \partial_\xi^{\beta_2} \widehat{b}(\varphi, j, \xi) e^{\mathrm{i}jx} .
$$

Therefore, since $\partial_\xi^{\beta_2} \widehat{b}(\varphi, j, \xi) = \widehat{\partial_\xi^{\beta_2} b}(\varphi, j, \xi)$ and, again by (2.33), we get

$$
\mathrm{Op}(\partial_\xi^\beta \sigma_{AB}) = \sum\nolimits_{\beta_1+\beta_2=\beta} C(\beta_1, \beta_2) \mathrm{Op}(\partial_\xi^{\beta_1} a) \circ \mathrm{Op}(\partial_\xi^{\beta_2} b) . \qquad (2.52)
$$

Since $\partial_\xi^{\beta_1} a \in S^{m-\beta_1}$, $\partial_\xi^{\beta_2} b \in S^{m'-\beta_2}$, $\beta_1 + \beta_2 = \beta$, the estimate (2.51) implies

$$
\begin{aligned}
|\mathrm{Op}(\partial_\xi^{\beta_1} a) \mathrm{Op}(\partial_\xi^{\beta_2} b)|_{m+m'-\beta,s,0} &\leq_{m,\beta} C(s) |\mathrm{Op}(\partial_\xi^{\beta_1} a)|_{m-\beta_1,s,0} |\mathrm{Op}(\partial_\xi^{\beta_2} b)|_{m'-\beta_2,s_0+\beta_1+|m|,0} \\
&\quad + C(s_0) |\mathrm{Op}(\partial_\xi^{\beta_1} a)|_{m-\beta_1,s_0,0} |\mathrm{Op}(\partial_\xi^{\beta_2} b)|_{m'-\beta_2,s+\beta_1+|m|,0} . \qquad (2.53)
\end{aligned}
$$

Therefore, for all $1 \le \beta \le \alpha$, by (2.52), (2.53) and the definition (2.36) we get

$$|\mathrm{Op}(\partial_\xi^\beta \sigma_{AB})|_{m+m'-\beta,s,0} \le_{m,\beta} C(s)|A|_{m,s,\alpha}|B|_{m',s_0+\alpha+|m|,\alpha} + C(s_0)|A|_{m,s_0,\alpha}|B|_{m',s+\alpha+|m|,\alpha}$$

which proves (2.43).

Now we prove (2.44). Recalling (2.30) it is sufficient to estimate each

$$r_{N,\tau}(\varphi,x,\xi) := \sum\nolimits_{j\in\mathbb{Z}} (\partial_\xi^N a)(\varphi,x,\xi+\tau j)\widehat{\partial_x^N b}(\varphi,j,\xi)e^{\mathrm{i}jx}\,, \quad \tau\in[0,1]\,. \tag{2.54}$$

Arguing as above (to prove (2.51)) we get

$$\begin{aligned}\|r_{N,\tau}(\cdot,\xi)\|_s\langle\xi\rangle^{N-(m+m')} &\le_{m,N} C(s)|\mathrm{Op}(\partial_\xi^N a)|_{m-N,s,0}|\mathrm{Op}(\partial_x^N b)|_{m',s_0+N+|m|,0}\\ &\quad + C(s_0)|\mathrm{Op}(\partial_\xi^N a)|_{m-N,s_0,0}|\mathrm{Op}(\partial_x^N b)|_{m',s+N+|m|,0}\\ &\le_{m,N} C(s)|\mathrm{Op}(\partial_\xi^N a)|_{m-N,s,0}|\mathrm{Op}(b)|_{m',s_0+2N+|m|,0}\\ &\quad + C(s_0)|\mathrm{Op}(\partial_\xi^N a)|_{m-N,s_0,0}|\mathrm{Op}(b)|_{m',s+2N+|m|,0}\end{aligned}$$

which gives (recall (2.30) and (2.36))

$$|R_N|_{m+m'-N,s,0} \le_{m,N} \frac{1}{N!}\big(C(s)|A|_{m,s,N}|B|_{m',s_0+2N+|m|,0} + C(s_0)|A|_{m,s_0,N}|B|_{m',s+2N+|m|,0}\big) \tag{2.55}$$

namely (2.44) for $\alpha=0$. We now prove (2.44) for $\alpha\ge1$. By differentiating (2.54) we get, $\forall 1\le\beta\le\alpha$,

$$\partial_\xi^\beta r_{N,\tau}(\varphi,x,\xi) = \sum_{\beta_1+\beta_2=\beta} C(\beta_1,\beta_2)\sum_{j\in\mathbb{Z}}(\partial_\xi^{N+\beta_1}a)(\varphi,x,\xi+\tau j)\widehat{\partial_x^N\partial_\xi^{\beta_2}b}(\varphi,j,\xi)e^{\mathrm{i}jx}$$

and so, arguing as for (2.53),

$$\begin{aligned}\|\partial_\xi^\beta r_{N,\tau}(\cdot,\xi)\|_s\langle\xi\rangle^{N+\beta-(m+m')} &\le_{m,N,\alpha}\sum_{\beta_1+\beta_2=\beta}\Big(C(s)|\mathrm{Op}(\partial_\xi^{N+\beta_1}a)|_{m-N-\beta_1,s,0}|\mathrm{Op}(\partial_\xi^{\beta_2}\partial_x^N b)|_{m'-\beta_2,s_0+N+|m|+\beta_1,0}\\ &\quad + C(s_0)|\mathrm{Op}(\partial_\xi^{N+\beta_1}a)|_{m-N-\beta_1,s_0,0}|\mathrm{Op}(\partial_\xi^{\beta_2}\partial_x^N b)|_{m'-\beta_2,s+N+|m|+\beta_1,0}\Big)\\ &\overset{(2.36)}{\le}_{m,N,\alpha} C(s)|A|_{m,s,N+\alpha}|B|_{m',s_0+2N+|m|+\alpha,\alpha} + C(s_0)|A|_{m,s_0,N+\alpha}|B|_{m',s+2N+|m|+\alpha,\alpha}\end{aligned}$$

and (2.44) is proved.

Finally we prove (2.41), (2.42) including the dependence on λ. For all $k\in\mathbb{N}^{\nu+1}$, $|k|\le k_0$, the derivative

$$\partial_\lambda^k\{A(\lambda)\circ B(\lambda)\} = \sum\nolimits_{k_1,k_2\in\mathbb{N}^{\nu+1},k_1+k_2=k} C(k_1,k_2)\partial_\lambda^{k_1}A(\lambda)\circ\partial_\lambda^{k_2}B(\lambda)\,.$$

Then (we have $|k|=|k_1|+|k_2|$)

$$\begin{aligned}&\gamma^{|k|}|\partial_\lambda^k\{A(\lambda)\circ B(\lambda)\}|_{m+m',s,\alpha} \le_{k_0} \sum_{k_1+k_2=k}\gamma^{|k_1|}\gamma^{|k_2|}|\partial_\lambda^{k_1}A(\lambda)\circ\partial_\lambda^{k_2}B(\lambda)|_{m+m',s,\alpha}\\ &\overset{(2.43)}{\le}_{k_0,m,\alpha}\sum_{k_1+k_2=k}\big(C(s)\gamma^{|k_1|}|\partial_\lambda^{k_1}A|_{m,s,\alpha}\gamma^{|k_2|}|\partial_\lambda^{k_2}B|_{m',s_0+\alpha+|m|,\alpha} + C(s_0)\gamma^{|k_1|}|\partial_\lambda^{k_1}A|_{m,s_0,\alpha}\gamma^{|k_2|}|\partial_\lambda^{k_2}B|_{m',s+\alpha+|m|,\alpha}\big)\end{aligned}$$

and (2.41) follows by the definition (2.35). The estimate (2.42) follows since for all $|k|\le k_0$

$$\begin{aligned}\gamma^{|k|}|\partial_\lambda^k\mathrm{Op}(r_{N,\tau})|_{m+m'-N,s,\alpha} &\le_{k_0,m,N,\alpha}\sum_{k_1+k_2=k}\big(C(s)\gamma^{|k_1|}|\partial_\lambda^{k_1}A|_{m,s,N+\alpha}\gamma^{|k_2|}|\partial_\lambda^{k_2}B|_{m',s_0+2N+|m|+\alpha,\alpha}\\ &\quad + C(s_0)\gamma^{|k_1|}|\partial_\lambda^{k_1}A|_{m,s_0,N+\alpha}\gamma^{|k_2|}|\partial_\lambda^{k_2}B|_{m',s+2N+|m|+\alpha,\alpha}\big)\,.\end{aligned}$$

The proof is complete. $\square$

When $B = g(D)$ is a Fourier multiplier, then $\mathrm{Op}(a) \circ g(D) = \mathrm{Op}(a(x,\xi)g(\xi))$ and we have a simpler estimate.

Lemma 2.7. *Let $A = a(\lambda, \varphi, x, \xi) \in OPS^m$, $m \in \mathbb{R}$, and let $g(D) \in OPS^{m'}$ be a Fourier multiplier (independent of λ). Then $|A \circ g(D)|^{k_0,\gamma}_{m+m',s,\alpha} \leq_{m,\alpha} |A|^{k_0,\gamma}_{m,s,\alpha}$.*

By (2.29) the commutator between two pseudo-differential operators $A = a(x,D) \in OPS^m$ and $B = b(x,D) \in OPS^{m'}$ is a pseudo-differential operator $[A,B] \in OPS^{m+m'-1}$ with symbol $a \star b$ (sometimes called the Moyal parenthesis of a and b), namely

$$[A,B] = \mathrm{Op}(a \star b) \,. \tag{2.56}$$

By (2.29) the symbol $a \star b \in S^{m+m'-1}$ admits the expansion

$$a \star b = -\mathrm{i}\{a,b\} + \mathtt{r}_2(a,b) \qquad \text{where} \quad \{a,b\} := \partial_\xi a \, \partial_x b - \partial_x a \, \partial_\xi b \tag{2.57}$$

is the Poisson bracket between $a(x,\xi)$ and $b(x,\xi)$, and $\mathtt{r}_2(a,b) := r_{2,AB} - r_{2,BA} \in S^{m+m'-2}$.

Lemma 2.8. (Commutators) *Let $A = a(\lambda, \varphi, x, D)$, $B = b(\lambda, \varphi, x, D)$ be pseudo-differential operators with symbols $a(\lambda, \varphi, x, \xi) \in S^m$, $b(\lambda, \varphi, x, \xi) \in S^{m'}$, $m, m' \in \mathbb{R}$. Then the commutator $[A,B] := AB - BA \in OPS^{m+m'-1}$ satisfies*

$$\begin{aligned} |[A,B]|^{k_0,\gamma}_{m+m'-1,s,\alpha} \leq_{m,m',\alpha,k_0} \big(& C(s) |A|^{k_0,\gamma}_{m,s+2+|m'|+\alpha,\alpha+1} |B|^{k_0,\gamma}_{m',s_0+2+|m|+\alpha,\alpha+1} \\ & + C(s_0) |A|^{k_0,\gamma}_{m,s_0+2+|m'|+\alpha,\alpha+1} |B|^{k_0,\gamma}_{m',s+2+|m|+\alpha,\alpha+1} \big) \,. \end{aligned} \tag{2.58}$$

Moreover the Poisson bracket $\{a,b\} \in S^{m+m'-1}$ satisfies

$$|\mathrm{Op}(\{a,b\})|^{k_0,\gamma}_{m+m'-1,s,\alpha} \leq_{\alpha,k_0} C(s) |A|^{k_0,\gamma}_{m,s+1,\alpha+1} |B|^{k_0,\gamma}_{m',s_0+1,\alpha+1} + C(s_0) |A|^{k_0,\gamma}_{m,s_0+1,\alpha+1} |B|^{k_0,\gamma}_{m',s+1,\alpha+1} \,. \tag{2.59}$$

Proof. The estimate (2.58) follows by (2.29), (2.42) for $N = 1$, and (2.37). The estimate (2.59) follows by (2.57), Definition 2.7, the tame estimates for the product of two functions (2.72) and (2.37). ☐

Note that in (2.59) the loss of regularity in s is smaller than in (2.58).

The adjoint A^* of a pseudo-differential operator $A = \mathrm{Op}(a) \in OPS^m$ is a pseudo-differential operator of the same order $A^* = \mathrm{Op}(a^*) \in OPS^m$ and the symbol a^* is defined in (2.31).

Lemma 2.9. (Adjoint) *Let $A = a(\lambda, \varphi, x, D)$ be a pseudo-differential operator with symbol $a(\lambda, \varphi, x, \xi) \in S^m, m \in \mathbb{R}$. Then the adjoint $A^* \in OPS^m$ satisfies*

$$|A^*|^{k_0,\gamma}_{m,s,0} \leq_m |A|^{k_0,\gamma}_{m,s+s_0+|m|,0} \,.$$

Proof. Recalling Definition 2.7 and (2.34) we have to estimate

$$|A^*|^2_{m,s,0} = \sup_{\xi \in \mathbb{R}} \|a^*(\cdot,\cdot,\xi)\|^2_s \langle \xi \rangle^{-2m} = \sum_{\ell,j} \langle \ell, j \rangle^{2s} |\widehat{a}(\ell, j, \xi - j)|^2 \langle \xi \rangle^{-2m} \,. \tag{2.60}$$

Since

$$|A|^2_{m,s+s_0+|m|,0} := \sup_{\xi \in \mathbb{R}} \|a(\cdot,\cdot,\xi)\|^2_{s+s_0+|m|} \langle \xi \rangle^{-2m} = \sup_{\xi \in \mathbb{R}} \sum\nolimits_{\ell,j} |\widehat{a}(\ell,j,\xi)|^2 \langle \ell, j \rangle^{2(s+s_0+|m|)} \langle \xi \rangle^{-2m}$$

we derive the bound, for all $\xi \in \mathbb{R}$, $\ell \in \mathbb{Z}^\nu$, $j \in \mathbb{Z}$,

$$|\widehat{a}(\ell, j, \xi - j)| \leq \frac{|A|_{m,s+s_0+|m|,0}}{\langle \ell, j \rangle^{s+s_0+|m|}} \langle \xi - j \rangle^m \,. \tag{2.61}$$

Then by (2.60), (2.61) and Peetre's inequality (2.47) we get

$$
\begin{aligned}
|A^*|^2_{m,s,0} &\leq \sum_{\ell,j} \frac{1}{\langle \ell, j\rangle^{2(s_0+|m|)}} \frac{\langle \xi - j\rangle^{2m}}{\langle \xi\rangle^{2m}} |A|^2_{m,s+s_0+|m|,0} \\
&\leq_m \sum_{\ell,j} \frac{\langle j\rangle^{2|m|}}{\langle \ell, j\rangle^{2(s_0+|m|)}} |A|^2_{m,s+s_0+|m|,0} \leq_m |A|^2_{m,s+s_0+|m|,0} \,. \qquad (2.62)
\end{aligned}
$$

The estimate for the derivatives with respect to λ follows analogously, since $\partial_\lambda^k A^* = \mathrm{Op}(\partial_\lambda^k a^*)$. □

Lemma 2.10. (Invertibility) *Let $\Phi := \mathrm{Id} + A$ where $A := \mathrm{Op}(a(\lambda, \varphi, x, j)) \in OPS^0$. There exist constants $C(s_0, \alpha, k_0)$, $C(s, \alpha, k_0) \geq 1$, $s \geq s_0$, such that, if*

$$
C(s_0, \alpha, k_0) |A|^{k_0,\gamma}_{0,s_0+\alpha,\alpha} \leq 1/2 \,, \qquad (2.63)
$$

then, for all ω, the operator Φ is invertible, $\Phi^{-1} \in OPS^0$ and, for all $s \geq s_0$,

$$
|\Phi^{-1} - \mathrm{Id}|^{k_0,\gamma}_{0,s,\alpha} \leq C(s, \alpha, k_0) |A|^{k_0,\gamma}_{0,s+\alpha,\alpha} \,.
$$

Proof. Iterating (2.41) (for $m = 0$) we deduce that there exist constants $C(s_0, \alpha, k_0), C(s, \alpha, k_0) \geq 1$ such that, $\forall n \in \mathbb{N}_+$,

$$
\begin{aligned}
&|A^n|^{k_0,\gamma}_{0,s_0,\alpha} \leq (C(s_0, \alpha, k_0))^{n-1} \big(|A|^{k_0,\gamma}_{0,s_0+\alpha,\alpha} \big)^n \,, \\
&|A^n|^{k_0,\gamma}_{0,s,\alpha} \leq n C(s, \alpha, k_0) \big(C(s_0, \alpha, k_0) |A|^{k_0,\gamma}_{0,s_0+\alpha,\alpha} \big)^{n-1} |A|^{k_0,\gamma}_{0,s+\alpha,\alpha} \,. \qquad (2.64)
\end{aligned}
$$

By (2.63) the operator Φ is invertible and the inverse Φ^{-1} may be expressed by the Neumann series $\Phi^{-1} = \mathrm{Id} + B$ with $B := \sum_{n \geq 1} (-1)^n A^n$. Moreover, since

$$
\|a(\cdot, j)\|_{L^\infty} \leq C(s_0) \|a(\cdot, j)\|_{s_0} \leq C(s_0) |A|_{0,s_0,0} \,, \quad \forall j \in \mathbb{Z} \,,
$$

the symbol of Φ satisfies $1 + a(\lambda, \varphi, x, j) \geq 1/2$, $\forall j \in \mathbb{Z}$, $\forall \lambda$, i.e it is elliptic. Hence the inverse operator B is pseudo-differential by the parametrix theorem (see [27]-Theorem 18.1.9). Moreover by (2.64)

$$
\begin{aligned}
|B|^{k_0,\gamma}_{0,s,\alpha} &\leq \sum\nolimits_{n \geq 1} |A^n|^{k_0,\gamma}_{0,s,\alpha} \leq \Big(\sum\nolimits_{n \geq 1} n (C(s_0, \alpha, k_0) |A|^{k_0,\gamma}_{0,s_0+\alpha,\alpha})^{n-1} \Big) C(s, \alpha, k_0) |A|^{k_0,\gamma}_{0,s+\alpha,\alpha} \\
&\leq C'(s, \alpha, k_0) |A|^{k_0,\gamma}_{0,s+\alpha,\alpha}
\end{aligned}
$$

by the smallness condition (2.63). □

2.2 $\mathcal{D}^{k_0}$-tame and $\mathcal{D}^{k_0}$-modulo-tame operators

Let $A := A(\lambda)$ be a linear operator k_0-times differentiable with respect to the parameter $\lambda \in \Lambda_0 \subset \mathbb{R}^{\nu+1}$.

Definition 2.8. ($\mathcal{D}^{k_0}$-σ-tame) *A linear operator $A := A(\lambda)$ is $\mathcal{D}^{k_0}$-σ-tame if the following weighted tame estimates hold: there exists $\sigma \geq 0$ such that, for all $s_0 \leq s \leq S$, with possibly $S = +\infty$, $\forall u \in H^{s+\sigma}$,*

$$
\sup_{|k| \leq k_0} \sup_{\lambda \in \Lambda_0} \gamma^{|k|} \| (\partial_\lambda^k A(\lambda)) u \|_s \leq \mathfrak{M}_A(s_0) \|u\|_{s+\sigma} + \mathfrak{M}_A(s) \|u\|_{s_0+\sigma} \qquad (2.65)
$$

where the functions $s \mapsto \mathfrak{M}_A(s) \geq 0$ are non-decreasing in s. We call $\mathfrak{M}_A(s)$ the TAME CONSTANT *of the operator A. The constant $\mathfrak{M}_A(s) := \mathfrak{M}_A(k_0, \sigma, s)$ depends also on k_0, σ but, since k_0, σ are considered in this paper absolute constants, we shall often omit to write them.*

When the "loss of derivatives" $\sigma = 0$ we simply call a $\mathcal{D}^{k_0}$-0-tame operator to be $\mathcal{D}^{k_0}$-tame.

Remark 2.11. In sections 6, 7 we work with $\mathcal{D}^{k_0}$-σ-tame operators with a finite $S < +\infty$, whose tame constants $\mathfrak{M}_A(s)$ may depend also on $C(S)$, for instance $\mathfrak{M}_A(s) \leq C(S)(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu})$, $\forall s_0 \leq s \leq S$. □

An immediate consequence of (2.65) (with $k=0$, $s=s_0$) is that

$$\|A\|_{\mathcal{L}(H^{s_0+\sigma},H^{s_0})} \leq 2\mathfrak{M}_A(s_0)\,. \tag{2.66}$$

Note also that representing the operator A by its matrix elements $\big(A_j^{j'}(\ell-\ell')\big)_{\ell,\ell'\in\mathbb{Z}^\nu,j,j'\in\mathbb{Z}}$ as in (2.12) we have, for all $|k|\leq k_0$, $j'\in\mathbb{Z}$, $\ell'\in\mathbb{Z}^\nu$,

$$\gamma^{2|k|}\sum\nolimits_{\ell,j}\langle \ell,j\rangle^{2s}|\partial_\lambda^k A_j^{j'}(\ell-\ell')|^2 \leq 2\big(\mathfrak{M}_A(s_0)\big)^2\langle \ell',j'\rangle^{2(s+\sigma)} + 2\big(\mathfrak{M}_A(s)\big)^2\langle \ell',j'\rangle^{2(s_0+\sigma)}\,. \tag{2.67}$$

The class of $\mathcal{D}^{k_0}$-σ-tame operators is closed under composition.

Lemma 2.12. **(Composition)** *Let A,B be respectively $\mathcal{D}^{k_0}$-σ_A-tame and $\mathcal{D}^{k_0}$-σ_B-tame operators with tame constants respectively $\mathfrak{M}_A(s)$ and $\mathfrak{M}_B(s)$. Then the composed operator $A\circ B$ is $\mathcal{D}^{k_0}$-$(\sigma_A+\sigma_B)$-tame with tame constant*

$$\mathfrak{M}_{AB}(s)\leq C(k_0)\big(\mathfrak{M}_A(s)\mathfrak{M}_B(s_0+\sigma_A)+\mathfrak{M}_A(s_0)\mathfrak{M}_B(s+\sigma_A)\big)\,.$$

Proof. As for the analogous inequality (2.75) below. □

Pseudo-differential operators are tame operators. We shall use in particular the following lemma.

Lemma 2.13. *Let $A=a(\lambda,\varphi,x,D)\in OPS^0$ be a family of pseudo-differential operators which are k_0-times differentiable with respect to λ. If $|A|_{0,s,0}^{k_0,\gamma}<+\infty$, $s\geq s_0$, then A is $\mathcal{D}^{k_0}$-tame with tame constant*

$$\mathfrak{M}_A(s)\leq C(s)|A|_{0,s,0}^{k_0,\gamma}\,. \tag{2.68}$$

Proof. By expanding (2.32) in Fourier, we have

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

Hence

$$\begin{aligned}\|Au\|_s^2 &= \sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\Big(\sum_{\ell'\in\mathbb{Z}^\nu,j'\in\mathbb{Z}}\widehat{a}(\ell-\ell',j-j',j')u_{\ell',j'}\Big)^2\langle \ell,j\rangle^{2s}\\ &\leq \sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\Big(\sum_{\ell'\in\mathbb{Z}^\nu,j'\in\mathbb{Z}}|\widehat{a}(\ell-\ell',j-j',j')||u_{\ell',j'}|\langle \ell,j\rangle^s\Big)^2 = S_1+S_2\end{aligned}$$

where

$$S_1:=\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\Big(\sum_{\langle \ell,j\rangle\langle \ell',j'\rangle^{-1}\leq 2^{1/s}}|\widehat{a}(\ell-\ell',j-j',j')|\langle \ell-\ell',j-j'\rangle^{s_0}|u_{\ell',j'}|\langle \ell',j'\rangle^s\frac{\langle \ell,j\rangle^s}{\langle \ell-\ell',j-j'\rangle^{s_0}\langle \ell',j'\rangle^s}\Big)^2$$

$$S_2:=\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\Big(\sum_{\langle \ell,j\rangle\langle \ell',j'\rangle^{-1}> 2^{1/s}}|\widehat{a}(\ell-\ell',j-j',j')|\langle \ell-\ell',j-j'\rangle^{s}|u_{\ell',j'}|\langle \ell',j'\rangle^{s_0}\frac{\langle \ell,j\rangle^s}{\langle \ell-\ell',j-j'\rangle^{s}\langle \ell',j'\rangle^{s_0}}\Big)^2.$$

By Cauchy Schwartz inequality, and denoting $\zeta(s_0):=\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\frac{1}{\langle \ell,j\rangle^{2s_0}}$ (which is $<+\infty$), we have

$$\begin{aligned}S_1 &\leq \sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\Big(\sum_{\langle \ell,j\rangle\langle \ell',j'\rangle^{-1}\leq 2^{1/s}}|\widehat{a}(\ell-\ell',j-j',j')|\langle \ell-\ell',j-j'\rangle^{s_0}|u_{\ell',j'}|\langle \ell',j'\rangle^s\frac{2}{\langle \ell-\ell',j-j'\rangle^{s_0}}\Big)^2\\ &\leq 4\zeta(s_0)\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\sum_{\ell'\in\mathbb{Z}^\nu,j'\in\mathbb{Z}}|\widehat{a}(\ell-\ell',j-j',j')|^2\langle \ell-\ell',j-j'\rangle^{2s_0}|u_{\ell',j'}|^2\langle \ell',j'\rangle^{2s}\\ &\leq 4\zeta(s_0)\sum_{\ell'\in\mathbb{Z}^\nu,j'\in\mathbb{Z}}|u_{\ell',j'}|^2\langle \ell',j'\rangle^{2s}\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}|\widehat{a}(\ell-\ell',j-j',j')|^2\langle \ell-\ell',j-j'\rangle^{2s_0}\\ &= 4\zeta(s_0)\sum_{\ell'\in\mathbb{Z}^\nu,j'\in\mathbb{Z}}|u_{\ell',j'}|^2\langle \ell',j'\rangle^{2s}\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}|\widehat{a}(\ell,j,j')|^2\langle \ell,j\rangle^{2s_0}\\ &= 4\zeta(s_0)\sum_{\ell'\in\mathbb{Z}^\nu,j'\in\mathbb{Z}}|u_{\ell',j'}|^2\langle \ell',j'\rangle^{2s}\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\|a(\cdot,\cdot,j')\|_{2s_0}^2\leq 4\zeta(s_0)\|u\|_s^2|A|_{0,s_0,0}^2\,.\end{aligned} \tag{2.69}$$

For the estimate of S_2 note that, since the indices satisfy $\langle \ell, j\rangle > 2^{1/s}\langle \ell', j'\rangle$ we have $\langle \ell, j\rangle \le \langle \ell', j'\rangle + \langle \ell' - \ell, j' - j\rangle \le 2^{-1/s}\langle \ell, j\rangle + \langle \ell - \ell', j - j'\rangle$ and therefore

$$\langle \ell, j\rangle \le \big(1 - 2^{-1/s}\big)^{-1}\langle \ell - \ell', j - j'\rangle\,.$$

As a consequence, repeating the same argument used for estimating S_1, we get

$$S_2 \le C(s)|A|^2_{0,s,0}\|u\|^2_{s_0}\,. \tag{2.70}$$

By (2.69), (2.70), we deduce that

$$\|Au\|_s \le 2(\zeta(s_0))^{1/2}|A|_{0,s_0,0}\,\|u\|_s + (C(s))^{1/2}|A|_{0,s,0}\|u\|_{s_0}$$

and therefore A is a tame operator with tame constant $\mathfrak{M}_A(s) \le C(s)|A|_{0,s,0}$ (for a different $C(s)$).

Since $\partial_\lambda^k A = \mathrm{Op}(\partial_\lambda^k a)$ for any $k \in \mathbb{N}^{\nu+1}$, $|k| \le k_0$, the general case of (2.68) follows. □

We now discuss the action of a $\mathcal{D}^{k_0}$-σ-tame operator $A(\omega)$ on Sobolev functions $u(\lambda) \in H^s$ which are k_0-times differentiable with respect to $\lambda \in \Lambda_0 \subset \mathbb{R}^{\nu+1}$. Recall the weighted norm $\|\ \|_s^{k_0,\gamma}$ in (2.5).

Lemma 2.14. *Let $A := A(\lambda)$ be a $\mathcal{D}^{k_0}$-σ-tame operator. Then, $\forall s \ge s_0$, for any family of Sobolev functions $u := u(\lambda) \in H^{s+\sigma}$ which is k_0-times differentiable with respect to λ, the following tame estimate holds*

$$\|Au\|_s^{k_0,\gamma} \le_{k_0} \mathfrak{M}_A(s_0)\|u\|^{k_0,\gamma}_{s+\sigma} + \mathfrak{M}_A(s)\|u\|^{k_0,\gamma}_{s_0+\sigma}\,.$$

Proof. For all $|k| \le k_0$, $\lambda \in \Lambda_0$, we have, by (2.65), (2.5)

$$\begin{aligned}
\|\partial_\lambda^k\big(A(\lambda)u(\lambda)\big)\|_s &\le_{k_0} \sum\nolimits_{k_1+k_2=k}\|(\partial_\lambda^{k_1}A(\lambda))[\partial_\lambda^{k_2}u(\lambda)]\|_s \\
&\le_{k_0} \sum\nolimits_{k_1+k_2=k}\gamma^{-|k_1|}\big(\mathfrak{M}_A(s_0)\|\partial_\lambda^{k_2}u\|_{s+\sigma} + \mathfrak{M}_A(s)\|\partial_\lambda^{k_2}u\|_{s_0+\sigma}\big) \\
&\le_{k_0} \gamma^{-|k|}\big(\mathfrak{M}_A(s_0)\|u\|^{k_0,\gamma}_{s+\sigma} + \mathfrak{M}_A(s)\|u\|^{k_0,\gamma}_{s_0+\sigma}\big)
\end{aligned}$$

and the lemma follows by the definition of the norm $\|\ \|_s^{k_0,\gamma}$ in (2.5). □

Lemma 2.14, (2.39) and (2.68) imply tame estimates for the product of two functions in weighted Sobolev norm: for all $s \ge s_0$,

$$\|uv\|_s \le C(s)\|u\|_s\|v\|_{s_0} + C(s_0)\|u\|_{s_0}\|v\|_s \tag{2.71}$$

$$\|uv\|_s^{k_0,\gamma} \le_{k_0} C(s)\|u\|_s^{k_0,\gamma}\|v\|_{s_0}^{k_0,\gamma} + C(s_0)\|u\|_{s_0}^{k_0,\gamma}\|v\|_s^{k_0,\gamma}\,, \tag{2.72}$$

as well as the algebra estimate $\|uv\|_s^{k_0,\gamma} \le_{k_0} C(s)\|u\|_s^{k_0,\gamma}\|v\|_s^{k_0,\gamma}$. In view of the KAM reducibility scheme of section 7 we also consider the stronger notion of $\mathcal{D}^{k_0}$-modulo-tame operator, that we need only for operators with loss of derivatives $\sigma = 0$.

Definition 2.9. **($\mathcal{D}^{k_0}$-modulo-tame)** *A linear operator $A := A(\lambda)$ is $\mathcal{D}^{k_0}$-modulo-tame if, for all $k \in \mathbb{N}^{\nu+1}$, $|k| \le k_0$, the majorant operators $|\partial_\lambda^k A|$ (Definition 2.2) satisfy the following weighted tame estimates: for all $s \ge s_0$, $u \in H^s$,*

$$\sup_{|k|\le k_0}\gamma^{|k|}\||\partial_\lambda^k A|u\|_s \le \mathfrak{M}^\sharp_A(s_0)\|u\|_s + \mathfrak{M}^\sharp_A(s)\|u\|_{s_0} \tag{2.73}$$

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

Lemma 2.15. *An operator A which is $\mathcal{D}^{k_0}$-modulo-tame is also $\mathcal{D}^{k_0}$-tame and $\mathfrak{M}_A(s) \le \mathfrak{M}^\sharp_A(s)$.*

Proof. For all $|k| \le k_0$ one has

$$\|(\partial_\lambda^k A)u\|_s^2 = \sum_{\ell,j} \langle \ell, j\rangle^{2s} \Big|\sum_{\ell',j'} \partial_\lambda^k A_j^{j'}(\ell-\ell') u_{\ell',j'}\Big|^2$$
$$\le \sum_{\ell,j} \langle \ell, j\rangle^{2s} \Big(\sum_{\ell',j'} |\partial_\lambda^k A_j^{j'}(\ell-\ell')||u_{\ell',j'}|\Big)^2 = \| |\partial_\lambda^k A| [\![u]\!] \|_s^2$$

where $[\![u]\!]$ is the function defined in (2.3). Then the lemma follows by (2.73), (2.4) and Definition 2.8. □

The class of operators which are $\mathcal{D}^{k_0}$-modulo-tame is closed under sum and composition.

Lemma 2.16. (Sum and composition) *Let A, B be $\mathcal{D}^{k_0}$-modulo-tame operators with modulo-tame constants respectively $\mathfrak{M}_A^\sharp(s)$ and $\mathfrak{M}_B^\sharp(s)$. Then $A+B$ is $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constant*

$$\mathfrak{M}_{A+B}^\sharp(s) \le \mathfrak{M}_A^\sharp(s) + \mathfrak{M}_B^\sharp(s) . \tag{2.74}$$

The composed operator $A \circ B$ is $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constant

$$\mathfrak{M}_{AB}^\sharp(s) \le C(k_0)\big(\mathfrak{M}_A^\sharp(s)\mathfrak{M}_B^\sharp(s_0) + \mathfrak{M}_A^\sharp(s_0)\mathfrak{M}_B^\sharp(s)\big) . \tag{2.75}$$

Assume in addition that $\langle \partial_\varphi\rangle^{\mathtt{b}} A$, $\langle \partial_\varphi\rangle^{\mathtt{b}} B$ are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constant respectively $\mathfrak{M}_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}^\sharp(s)$ and $\mathfrak{M}_{\langle \partial_\varphi\rangle^{\mathtt{b}} B}^\sharp(s)$, then $\langle \partial_\varphi\rangle^{\mathtt{b}}(AB)$ is $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constant satisfsying

$$\begin{aligned}\mathfrak{M}_{\langle \partial_\varphi\rangle^{\mathtt{b}}(AB)}^\sharp(s) \le C(\mathtt{b})C(k_0)\Big(&\mathfrak{M}_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}^\sharp(s)\mathfrak{M}_B^\sharp(s_0) + \mathfrak{M}_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}^\sharp(s_0)\mathfrak{M}_B^\sharp(s)\\ &+ \mathfrak{M}_A^\sharp(s)\mathfrak{M}_{\langle \partial_\varphi\rangle^{\mathtt{b}} B}^\sharp(s_0) + \mathfrak{M}_A^\sharp(s_0)\mathfrak{M}_{\langle \partial_\varphi\rangle^{\mathtt{b}} B}^\sharp(s)\Big) .\end{aligned} \tag{2.76}$$

The constants $C(k_0), C(\mathtt{b}) \ge 1$.

Proof. The bound (2.74) follows by (2.14) and (2.4).
Proof of (2.75). For all $|k| \le k_0$ we have

$$\begin{aligned}\gamma^{|k|}\| |\partial_\lambda^k (AB)| u\|_s &\le C(k_0)\gamma^{|k|}\sum_{k_1+k_2=k} \| |(\partial_\lambda^{k_1} A)(\partial_\lambda^{k_2} B)| u\|_s\\ &\overset{(2.14)}{\le} C(k_0)\sum_{k_1+k_2=k} \gamma^{|k_1|}\gamma^{|k_2|}\| |\partial_\lambda^{k_1} A||\partial_\lambda^{k_2} B| [\![u]\!]\|_s\\ &\overset{(2.73)}{\le} C(k_0)\sum_{|k_2|\le|k|} \mathfrak{M}_A^\sharp(s_0)\gamma^{|k_2|}\| |\partial_\lambda^{k_2} B| [\![u]\!]\|_s + \mathfrak{M}_A^\sharp(s)\gamma^{|k_2|}\| |\partial_\lambda^{k_2} B| [\![u]\!]\|_{s_0}\\ &\overset{(2.73),(2.4)}{\le} C(k_0)\big(\mathfrak{M}_A^\sharp(s_0)\mathfrak{M}_B^\sharp(s_0)\|u\|_s + \big(\mathfrak{M}_A^\sharp(s)\mathfrak{M}_B^\sharp(s_0) + \mathfrak{M}_A^\sharp(s_0)\mathfrak{M}_B^\sharp(s)\big)\|u\|_{s_0}\big)\end{aligned}$$

and (2.75) follows by recalling Definition 2.9.
Proof of (2.76). For all $|k| \le k_0$ we have (use the first inequality in (2.14))

$$\Big\| |\langle \partial_\varphi\rangle^{\mathtt{b}}\big[\partial_\lambda^k (AB)\big]| u\Big\|_s \le C(k_0)\sum_{k_1+k_2=k} \Big\| |\langle \partial_\varphi\rangle^{\mathtt{b}}\big[(\partial_\lambda^{k_1} A)(\partial_\lambda^{k_2} B)\big]| [\![u]\!]\Big\|_s . \tag{2.77}$$

Next, recalling the Definition 2.2 of the operator $\langle \partial_\varphi\rangle^{\mathtt{b}}$ and (2.3), we have

$$\begin{aligned}\Big\| |\langle \partial_\varphi\rangle^{\mathtt{b}}\big[(\partial_\lambda^{k_1} A)(\partial_\lambda^{k_2} B)\big]| [\![u]\!]\Big\|_s^2 &= \sum_{\ell,j}\langle \ell, j\rangle^{2s}\Big(\sum_{\ell',j'} |\langle \ell-\ell'\rangle^{\mathtt{b}}[(\partial_\lambda^{k_1} A)(\partial_\lambda^{k_2} B)]_j^{j'}(\ell-\ell')||u_{\ell',j'}|\Big)^2\\ &\le \sum_{\ell,j}\langle \ell, j\rangle^{2s}\Big(\sum_{\ell',j',\ell_1,j_1} \langle \ell-\ell'\rangle^{\mathtt{b}}|(\partial_\lambda^{k_1} A)_j^{j_1}(\ell-\ell_1)||(\partial_\lambda^{k_2} B)_{j_1}^{j'}(\ell_1-\ell')||u_{\ell',j'}|\Big)^2 .\end{aligned} \tag{2.78}$$

Since $\langle \ell-\ell'\rangle^{\mathtt{b}} \leq C(\mathtt{b})(\langle \ell-\ell_1\rangle^{\mathtt{b}} + \langle \ell_1-\ell'\rangle^{\mathtt{b}})$, we deduce that

$$
\begin{aligned}
(2.78) \leq\ & C(\mathtt{b})^2 \sum_{\ell,j} \langle \ell,j\rangle^{2s}\Big(\sum_{\ell',j',\ell_1,j_1} |\langle \ell-\ell_1\rangle^{\mathtt{b}} (\partial_\lambda^{k_1} A)_j^{j_1}(\ell-\ell_1)||(\partial_\lambda^{k_2} B)_{j_1}^{j'}(\ell_1-\ell')||u_{\ell',j'}|\Big)^2 \\
& + C(\mathtt{b})^2 \sum_{\ell,j} \langle \ell,j\rangle^{2s}\Big(\sum_{\ell',j',\ell_1,j_1} |(\partial_\lambda^{k_1} A)_j^{j_1}(\ell-\ell_1)||\langle \ell_1-\ell'\rangle^{\mathtt{b}}(\partial_\lambda^{k_2} B)_{j_1}^{j'}(\ell_1-\ell')||u_{\ell',j'}|\Big)^2 \\
\leq\ & C(\mathtt{b})^2 \Big(\Big\| \big|\langle \partial_\varphi\rangle^{\mathtt{b}} (\partial_\lambda^{k_1} A)\big| \big[|\partial_\lambda^{k_2} B| \|u\|\big] \Big\|_s^2 + \Big\| \big|\partial_\lambda^{k_1} A\big| \big[|\langle \partial_\varphi\rangle^{\mathtt{b}}(\partial_\lambda^{k_2} B)| \|u\|\big] \Big\|_s^2 \Big) . \qquad (2.79)
\end{aligned}
$$

Hence (2.77)-(2.79), (2.73) and (2.4) imply

$$
\begin{aligned}
\big\| |\langle \partial_\varphi\rangle^{\mathtt{b}} [\partial_\lambda^k (AB)]| u \big\|_s \leq\ & C(\mathtt{b}) C(k_0) \gamma^{-|k|} \big(\mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}(s_0) \mathfrak{M}^\sharp_B(s_0) + \mathfrak{M}^\sharp_A(s_0) \mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} B}(s_0)\big) \|u\|_s \\
& + C(\mathtt{b}) C(k_0) \gamma^{-|k|} \Big(\mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}(s) \mathfrak{M}^\sharp_B(s_0) + \mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}(s_0) \mathfrak{M}^\sharp_B(s) \\
& + \mathfrak{M}^\sharp_A(s) \mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} B}(s_0) + \mathfrak{M}^\sharp_A(s_0) \mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} B}(s) \Big) \|u\|_{s_0}
\end{aligned}
$$

which proves (2.76). □

As a consequence of (2.75), if A is $\mathcal{D}^{k_0}$-modulo-tame, then, for all $n \geq 1$, each A^n is $\mathcal{D}^{k_0}$-modulo-tame and

$$
\mathfrak{M}^\sharp_{A^n}(s) \leq \big(2C(k_0)\mathfrak{M}^\sharp_A(s_0)\big)^{n-1} \mathfrak{M}^\sharp_A(s) . \qquad (2.80)
$$

Moreover, by (2.76), if $\langle \partial_\varphi\rangle^{\mathtt{b}} A$ is $\mathcal{D}^{k_0}$-modulo-tame, then, for all $n \geq 2$, each $\langle \partial_\varphi\rangle^{\mathtt{b}} A^n$ is $\mathcal{D}^{k_0}$-modulo-tame with

$$
\mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} A^n}(s) \leq (4C(\mathtt{b})C(k_0))^{n-1} \Big(\mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}(s) \big[\mathfrak{M}^\sharp_A(s_0)\big]^{n-1} + \mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}(s_0) \mathfrak{M}^\sharp_A(s) \big[\mathfrak{M}^\sharp_A(s_0)\big]^{n-2} \Big) . \quad (2.81)
$$

Lemma 2.17 (Invertibility). *Let $\Phi := \mathrm{Id} + A$ where $A := A(\lambda)$ is $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constant $\mathfrak{M}^\sharp_A(s)$. Assume the smallness condition*

$$
4C(\mathtt{b})C(k_0)\mathfrak{M}^\sharp_A(s_0) \leq 1/2 . \qquad (2.82)
$$

Then the operator Φ is invertible, $\breve{A} := \Phi^{-1} - \mathrm{Id}$ is $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constant

$$
\mathfrak{M}^\sharp_{\breve{A}}(s) \leq 2\mathfrak{M}^\sharp_A(s) . \qquad (2.83)
$$

Moreover $\langle \partial_\varphi\rangle^{\mathtt{b}} \breve{A}$ is $\mathcal{D}^{k_0}$-modulo-tame with tame-constant

$$
\mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} \breve{A}}(s) \leq 2\mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}(s) + 8C(\mathtt{b})C(k_0)\mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}(s_0)\, \mathfrak{M}^\sharp_A(s) . \qquad (2.84)
$$

Proof. By (2.66) and (2.82) the operatorial norm $\|A\|_{\mathcal{L}(H^{s_0})} \leq 2\mathfrak{M}^\sharp_A(s_0) \leq 1/2$. Then Φ is invertible and the inverse operator $\Phi^{-1} = \mathrm{Id} + \breve{A}$ with $\breve{A} := \sum_{n\geq 1}(-1)^n A^n$ satisfy the estimate (2.83) by (2.74), (2.80), (2.82). Similarly (2.84) follows by (2.74), (2.81) and (2.82). □

Lemma 2.18. (Smoothing) *Suppose that $\langle \partial_\varphi\rangle^{\mathtt{b}} A$, $\mathtt{b} \geq 0$, is $\mathcal{D}^{k_0}$-modulo-tame. Then the operator $\Pi_N^\perp A$ is $\mathcal{D}^{k_0}$-modulo-tame with tame constant*

$$
\mathfrak{M}^\sharp_{\Pi_N^\perp A}(s) \leq N^{-\mathtt{b}} \mathfrak{M}^\sharp_{\langle \partial_\varphi\rangle^{\mathtt{b}} A}(s) , \quad \mathfrak{M}^\sharp_{\Pi_N^\perp A}(s) \leq \mathfrak{M}^\sharp_A(s) . \qquad (2.85)
$$

Proof. For all $|k| \leq k_0$ one has, recalling (2.13),

$$
\begin{aligned}
\| |\Pi_N^\perp \partial_\lambda^k A| u\|_s^2 &= \sum_{\ell,j} \langle \ell,j\rangle^{2s} \Big(\sum_{j',|\ell-\ell'|>N} |\partial_\lambda^k A_j^{j'}(\ell-\ell')||u_{\ell'j'}| \Big)^2 \\
&\leq N^{-2\mathtt{b}} \sum\nolimits_{\ell,j} \langle \ell,j\rangle^{2s} \big(\sum\nolimits_{j',\ell'} |\langle \ell-\ell'\rangle^{\mathtt{b}} \partial_\lambda^k A_j^{j'}(\ell-\ell')||u_{\ell'j'}| \big)^2 \\
&= N^{-2\mathtt{b}} \| |\langle \partial_\varphi\rangle^{\mathtt{b}} (\partial_\lambda^k A)| \, [\|u\|] \|_s^2
\end{aligned}
$$

and, using (2.73), (2.4), we deduce the first inequality in (2.85). Similarly we get $\| |\Pi_N^\perp \partial_\lambda^k A| u\|_s^2 \leq \| |\partial_\lambda^k A| \, \|u\| \, \|_s^2$ which implies the second inequality in (2.85). □

The next two lemmata will be used in the proof of Theorem 7.3-$(\mathbf{S3})_\nu$.

Lemma 2.19. *Let A and B be linear operators such that $|A|, |\langle\partial_\varphi\rangle^{\mathtt b} A|, |B|, |\langle\partial_\varphi\rangle^{\mathtt b} B| \in \mathcal{L}(H^{s_0})$. Then*

1. $\||A+B|\|_{\mathcal{L}(H^{s_0})} \leq \||A|\|_{\mathcal{L}(H^{s_0})} + \||B|\|_{\mathcal{L}(H^{s_0})}$, $\qquad \||AB|\|_{\mathcal{L}(H^{s_0})} \leq \||A|\|_{\mathcal{L}(H^{s_0})} \||B|\|_{\mathcal{L}(H^{s_0})}$,
2. $\||\langle\partial_\varphi\rangle^{\mathtt b}(AB)|\|_{\mathcal{L}(H^{s_0})} \lesssim_{\mathtt b} \||\langle\partial_\varphi\rangle^{\mathtt b} A|\|_{\mathcal{L}(H^{s_0})} \||B|\|_{\mathcal{L}(H^{s_0})} + \||A|\|_{\mathcal{L}(H^{s_0})} \||\langle\partial_\varphi\rangle^{\mathtt b} B|\|_{\mathcal{L}(H^{s_0})}$,
3. $\||\Pi_N^\perp A|\|_{\mathcal{L}(H^{s_0})} \leq N^{-\mathtt b} \||\langle\partial_\varphi\rangle^{\mathtt b} A|\|_{\mathcal{L}(H^{s_0})}$, $\qquad \||\Pi_N^\perp A|\|_{\mathcal{L}(H^{s_0})} \leq \||A|\|_{\mathcal{L}(H^{s_0})}$.

Proof. Item 1 is a direct consequence of (2.14) and (2.4). Items 2-3 are proved arguing as in Lemmata 2.16 and 2.18. □

Lemma 2.20. *Let $\Phi_i := \mathrm{Id} + \Psi_i$, $i = 1, 2$, satisfy,*

$$\||\Psi_i|\|_{\mathcal{L}(H^{s_0})} \leq 1/2\,, \quad i = 1, 2\,. \tag{2.86}$$

Then $\Phi_i^{-1} = \mathrm{Id} + \check\Psi_i$, $i = 1, 2$, satisfy $\||\check\Psi_1 - \check\Psi_2|\|_{\mathcal{L}(H^{s_0})} \leq 4 \||\Psi_1 - \Psi_2|\|_{\mathcal{L}(H^{s_0})}$ and

$$\begin{aligned} \||\langle\partial_\varphi\rangle^{\mathtt b} |\check\Psi_1 - \check\Psi_2|\|_{\mathcal{L}(H^{s_0})} &\lesssim_{\mathtt b} \||\langle\partial_\varphi\rangle^{\mathtt b} |\Psi_1 - \Psi_2|\|_{\mathcal{L}(H^{s_0})} \\ &+ \big(1 + \||\langle\partial_\varphi\rangle^{\mathtt b} \check\Psi_1|\|_{\mathcal{L}(H^{s_0})} + \||\langle\partial_\varphi\rangle^{\mathtt b} \check\Psi_2|\|_{\mathcal{L}(H^{s_0})}\big) \||\Psi_1 - \Psi_2|\|_{\mathcal{L}(H^{s_0})}\,. \end{aligned}$$

Proof. Use $\check\Psi_1 - \check\Psi_2 = \Phi_1^{-1} - \Phi_2^{-2} = \Phi_1^{-1}(\Psi_2 - \Psi_1)\Phi_2^{-1}$ and apply Lemma 2.19-1-2, using (2.86). □

The composition operator $u(y) \mapsto u(y + p(y))$ induced by a diffeomorphism of the torus $\mathbb{T}^d$ is tame.

Lemma 2.21. (Change of variable) *Let $p := p(\lambda, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$ be a family of 2π-periodic functions which is k_0-times differentiable with respect to $\lambda \in \Lambda_0 \subset \mathbb{R}^{\nu+1}$, satisfying*

$$\|p\|_{\mathcal{C}^{s_0+1}} \leq 1/2\,, \quad \|p\|_{s_0}^{k_0,\gamma} \leq 1\,. \tag{2.87}$$

Let $g(y) := y + p(y)$, $y \in \mathbb{T}^d$. Then the composition operator

$$A : u(y) \mapsto (u \circ g)(y) = u(y + p(y))$$

satisfies the tame estimates

$$\|Au\|_{s_0} \lesssim_{s_0} \|u\|_{s_0}\,, \qquad \|Au\|_s \leq C(s)\|u\|_s + C(s_0)\|p\|_s \|u\|_{s_0+1}\,, \ \forall s \geq s_0 + 1\,, \tag{2.88}$$

and for any $|k| \leq k_0$,

$$\|(\partial_\lambda^k A)u\|_{s_0} \lesssim_{s_0,k} \gamma^{-|k|} \|u\|_{s_0+|k|}\,, \tag{2.89}$$

$$\|(\partial_\lambda^k A)u\|_s \lesssim_{s,k} \gamma^{-|k|} \big(\|u\|_{s+|k|} + \|p\|_s^{|k|,\gamma} \|u\|_{s_0+|k|+1}\big)\,, \quad \forall s \geq s_0 + 1\,. \tag{2.90}$$

The map g is invertible with inverse $g^{-1}(z) = z + q(z)$. Suppose $\partial_\lambda^k p(\lambda, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$ for all $|k| \leq k_0$. There exists a constant $\delta := \delta(s_0, k_0) \in (0, 1)$ such that, if $\|p\|_{2s_0+k_0+1}^{k_0,\gamma} \leq \delta$, then

$$\|q\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|p\|_{s+k_0}^{k_0,\gamma}\,, \quad \forall s \geq s_0\,. \tag{2.91}$$

The composition operators A and A^{-1} are $\mathcal{D}^{k_0}$-(k_0+1)-tame with tame constants satisfying for any $S > s_0$,

$$\mathfrak{M}_A(s) \lesssim_{S,k_0} 1 + \|p\|_s^{k_0,\gamma}\,, \quad \mathfrak{M}_{A^{-1}}(s) \lesssim_{S,k_0} 1 + \|p\|_{s+k_0}^{k_0,\gamma}\,, \quad \forall s_0 \leq s \leq S\,. \tag{2.92}$$

Proof. PROOF OF (2.88). By Lemma B.4-(*ii*) in [6] and (2.87), we have

$$\|Au\|_{s_0} \leq_{s_0} \|u\|_{s_0} + \|p\|_{\mathcal{C}^{s_0}}\|u\|_1 \leq_{s_0} \|u\|_{s_0} \qquad \text{and} \qquad \|Au\|_{s_0+1} \leq_{s_0} \|u\|_{s_0+1} \,. \tag{2.93}$$

Thus the the first inequality in (2.88), and the second one for $s = s_0 + 1$, are proved. Now we prove the second inequality in (2.88), arguing by induction on s. We assume that it holds for $s \geq s_0 + 1$ and we prove it for $s + 1$. As a notation we denote by $\nabla u := (u_{x_1}, \ldots, u_{x_d})$ the gradient of the function u and $A(\nabla u) := (Au_{x_1}, \ldots, Au_{x_d})$. By the definition of the $\| \ \|_{s+1}$ norm and (2.71) we have

$$\begin{aligned}\|Au\|_{s+1} &\leq \|Au\|_{L^2} + \max_{|\alpha|=1} \|\partial_x^\alpha (Au)\|_s \\ &\leq \|Au\|_{L^2} + C(s)\|A(\nabla u)\|_s + C(s)\|A(\nabla u)\|_s\|p\|_{s_0+1} + C(s_0)\|A(\nabla u)\|_{s_0}\|p\|_{s+1} \,.\end{aligned}$$

Hence, by the inductive hyphothesis and using (2.87), (2.93), we get

$$\|Au\|_{s+1} \leq C_1(s)\|u\|_{s+1} + C_1(s)\|p\|_s\|u\|_{s_0+2} + C_0(s_0)\|p\|_{s+1}\|u\|_{s_0+1} \tag{2.94}$$

for some constants $C_1(s), C_0(s_0) > 0$. Applying (2.9) with $a_0 = b_0 = s_0 + 1$, $q = 1$, $p = s - s_0 - 1$, $\epsilon = 1/C_1(s)$, we estimate

$$C_1(s)\|p\|_s\|u\|_{s_0+2} \leq \|p\|_{s+1}\|u\|_{s_0+1} + C_2(s)\|p\|_{s_0+1}\|u\|_{s+1} \,,$$

and, by (2.94), using again that $\|p\|_{s_0+1} \leq 1$, we get

$$\|Au\|_{s+1} \leq C(s+1)\|u\|_{s+1} + C(s_0)\|p\|_{s+1}\|u\|_{s_0+1} \,,$$

with $C(s+1) = C_1(s) + C_2(s)$ and $C(s_0) = 1 + C_0(s_0)$. This is (2.88) for the Sobolev index $s + 1$.
PROOF OF (2.89)-(2.90). We prove the estimate (2.90). We argue by induction on $|k| \leq k_0$. For $k = 0$, the estimate (2.90) follows by (2.88). Now we assume that (2.90) holds for any $|k| \leq n < k_0$ and we prove it for $n + 1$. Let $\alpha \in \mathbb{N}^{\nu+1}$ such that $|\alpha| = 1$. One has

$$(\partial_\lambda^{k+\alpha} A)u = \partial_\lambda^k (A(\nabla u) \cdot \partial_\lambda^\alpha p) = \sum\nolimits_{k_1+k_2=k} C(k_1, k_2)(\partial_\lambda^{k_1} A)(\nabla u) \cdot \partial_\lambda^{k_2+\alpha} p \,. \tag{2.95}$$

For any $k_1, k_2 \in \mathbb{N}^{\nu+1}$, with $k_1 + k_2 = k$, we have, using (2.71),

$$\begin{aligned}\|(\partial_\lambda^{k_1} A)(\nabla u) \cdot \partial_\lambda^{k_2+\alpha} p\|_s &\leq_s \|(\partial_\lambda^{k_1} A)(\nabla u)\|_s\|\partial_\lambda^{k_2+\alpha} p\|_{s_0} + \|(\partial_\lambda^{k_1} A)(\nabla u)\|_{s_0}\|\partial_\lambda^{k_2+\alpha} p\|_s \\ &\overset{(2.89),(2.90)}{\leq_{s,k_1}} \gamma^{-|k_1|}\big(\|u\|_{s+|k_1|+1} + \|p\|_s^{|k_1|,\gamma}\|u\|_{s_0+|k_1|+2}\big)\gamma^{-(|k_2|+1)}\|p\|_{s_0}^{|k_2|+1,\gamma} \\ &\quad + \gamma^{-|k_1|}\|u\|_{s_0+|k_1|+2}\gamma^{-(|k_2+1)}\|p\|_s^{|k_2|+1,\gamma} \\ &\overset{(2.87)}{\leq_{s,k_1}} \gamma^{-(|k|+1)}\big(\|u\|_{s+|k|+1} + \|p\|_s^{|k|+1,\gamma}\|u\|_{s_0+|k|+2}\big)\end{aligned}$$

and recalling (2.95) we get the estimate (2.90) for $|k| + 1$.
PROOF OF (2.91). Since $y + p(\lambda, y) = z \iff z + q(\lambda, z) = y$ the function $q(\lambda, z)$ satisfies

$$q(\lambda, z) + p(\lambda, z + q(\lambda, z)) = 0. \tag{2.96}$$

If $p \in \mathcal{C}^1$ with respect to (λ, y), then, by the standard implicit function theorem, q is $\mathcal{C}^1$ with respect to (λ, z) and by differentiating the identity (2.96) one gets, denoting by D_λ, D_y, D_z the Fréchet derivatives with respect to the variables λ, y, z,

$$\begin{aligned}D_\lambda q(\lambda, z) &= -\big(\mathrm{Id} + D_y p(\lambda, z + q(\lambda, z))\big)^{-1} D_\lambda p(\lambda, z + q(\lambda, z)) \,, \\ D_z q(\lambda, z) &= -\big(\mathrm{Id} + D_y p(\lambda, z + q(\lambda, z))\big)^{-1} D_x p(\lambda, z + q(\lambda, z)) \,.\end{aligned}$$

It then follows by usual bootstrap arguments that if p is k_0-times differentiable with respect to λ and $\partial_\lambda^k p(\lambda, \cdot) \in \mathcal{C}^\infty$ for any $|k| \leq k_0$, then q is k_0-times differentiable with respect to λ and $\partial_\lambda^k q(\lambda, \cdot) \in \mathcal{C}^\infty$ for any $|k| \leq k_0$. We now prove

$$\|\partial_\lambda^k q\|_s \leq_s \gamma^{-|k|}\|p\|_{s+|k|}^{|k|,\gamma} \,, \qquad \forall k \in \mathbb{N}^{\nu+1} \,, \ |k| \leq k_0 \,. \tag{2.97}$$

which, recalling (2.5), implies (2.91). Denote by A_q the composition operator

$$A_q : h(x) \mapsto h(x + q(x))$$

so that $q = -A_q[p]$. By differentiating the equation $q(\lambda, z) + p(\lambda, z + q(\lambda, z)) = 0$, $(s_0 + 1)$-times, one gets that $\|q\|_{\mathcal{C}^{s_0+1}} \leq C(s_0)\|p\|_{\mathcal{C}^{s_0+1}} \leq 1/2$, provided $\|p\|_{\mathcal{C}^{s_0+1}}$ is small enough and $\|q\|_{s_0}^{k_0,\gamma} \leq C(s_0)\|p\|_{s_0+k_0}^{k_0,\gamma} \leq 1/2$, provided $\|p\|_{s_0+k_0}^{k_0,\gamma}$ small enough. Therefore, we can apply the estimates (2.88)-(2.90) to the operator A_q. By (2.88), one has

$$\|q\|_s = \|A_q(p)\|_s \leq C(s)\|p\|_s + C(s_0)\|q\|_s\|p\|_{s_0+1}\,,$$

which, for $C(s_0)\|p\|_{s_0+1} \leq 1/2$, implies (2.97) for $k = 0$. Now we assume that (2.97) holds up to $|k| = n$ and we prove it for $n + 1$. Let $\alpha \in \mathbb{N}^{\nu+1}$ such that $|\alpha| = 1$. We have

$$\begin{aligned}\partial_\lambda^{k+\alpha} q &= -\partial_\lambda^{k+\alpha}(A_q(p)) = -\partial_\lambda^{k}\big(A_q(\nabla p)\cdot\partial_\lambda^\alpha q + A_q(\partial_\lambda^\alpha p)\big)\\ &= -A_q(\nabla p)\cdot\partial_\lambda^{k+\alpha} q - \sum_{k_1+k_2=k, |k_2|<|k|} C_{k_1,k_2}\partial_\lambda^{k_1}\big(A_q(\nabla p)\big)\cdot\partial_\lambda^{k_2+\alpha} q - \partial_\lambda^k\big(A_q(\partial_\lambda^\alpha p)\big)\,.\end{aligned}$$

Using (2.71) we get

$$\begin{aligned}\|\partial_\lambda^{k+\alpha} q\|_s &\leq C(s_0)\|A_q(\nabla p)\|_{s_0}\|\partial_\lambda^{k+\alpha} q\|_s + C(s)\|A_q(\nabla p)\|_s\|\partial_\lambda^{k+\alpha} q\|_{s_0} + \|\partial_\lambda^k(A_q(\partial_\lambda^\alpha p))\|_s\\ &\quad + C(k,s)\sum_{k_1+k_2=k, |k_2|<|k|}\|\partial_\lambda^{k_1}(A_q(\nabla p))\|_s\|\partial_\lambda^{k_2+\alpha} q\|_{s_0} + \|\partial_\lambda^{k_1}(A_q(\nabla p))\|_s\|\partial_\lambda^{k_2+\alpha} q\|_{s_0}\\ &\overset{(2.88),(2.97),\|p\|_{s_0+2}\leq 1}{\leq} C_1(s_0)\|p\|_{s_0+1}\|\partial_\lambda^{k+\alpha} q\|_s + C_1(s)\|p\|_{s+1}\|\partial_\lambda^{k+\alpha} q\|_{s_0} + \gamma^{-|k|}\|A_q(\partial_\lambda^\alpha p)\|_s^{|k|,\gamma}\\ &\quad + \gamma^{-(|k|+1)}C_1(k,s)\sum_{\substack{k_1+k_2=k\\ |k_2|<|k|}}\|A_q(\nabla p)\|_s^{|k_1|,\gamma}\|p\|_{s_0+|k_2|+1}^{|k_2|+1,\gamma} + \|A_q(\nabla p)\|_{s_0}^{|k_1|,\gamma}\|p\|_{s+|k_2|+1}^{|k_2|+1,\gamma}\\ &\leq C_1(s_0)\|p\|_{s_0+1}\|\partial_\lambda^{k+\alpha} q\|_s + C_1(s)\|p\|_{s+1}\|\partial_\lambda^{k+\alpha} q\|_{s_0} + C_2(s,k)\gamma^{-(|k|+1)}\|p\|_{s+|k|+1}^{|k|+1,\gamma}\end{aligned}\tag{2.98}$$

using (2.89), (2.90), (2.97), Lemma (2.14) and $\|p\|_{s_0+k_0+1}^{k_0,\gamma} \leq 1$. Then, for $s = s_0$, one has

$$\|\partial_\lambda^{k+\alpha} q\|_{s_0} \leq 2C_1(s_0)\|p\|_{s_0+1}\|\partial_\lambda^{k+\alpha} q\|_{s_0} + C_2(s_0,k)\gamma^{-(|k|+1)}\|p\|_{s_0+|k|+1}^{|k|+1,\gamma}\,,\tag{2.99}$$

implying (2.97) for $k + \alpha$ and $s = s_0$, by taking $2C_1(s_0)\|p\|_{s_0+1} \leq 1/2$. Then the estimate for $s > s_0$, follows by (2.98), (2.99), (2.87). Finally (2.92) follows by (2.88)-(2.90), (2.91). ☐

We finally state the following generalized Moser tame estimates for the composition operator

$$u(\varphi, x) \mapsto \mathtt{f}(u)(\varphi, x) := f(\varphi, x, u(\varphi, x))$$

which can be proved arguing as in the previous lemma. Since the variables $(\varphi, x) := y$ have the same role, we present it for a generic Sobolev space $H^s(\mathbb{T}^d)$.

Lemma 2.22. (Composition operator) *Let $f \in \mathcal{C}^\infty(\mathbb{T}^d \times \mathbb{R}, \mathbb{R})$. If $u(\lambda) \in H^s(\mathbb{T}^d)$ is a family of Sobolev functions satisfying $\|u\|_{s_0}^{k_0,\gamma} \leq 1$, then, $\forall s > s_0 := (d+1)/2$, $\|\mathtt{f}(u)\|_s \leq C(s,f)(1 + \|u\|_s)$ and $\|\mathtt{f}(u)\|_s^{k_0,\gamma} \leq C(s,k_0,f)(1 + \|u\|_s^{k_0,\gamma})$.*

2.3 Integral operators and Hilbert transform

We now consider integral operators with a $\mathcal{C}^\infty$ Kernel.

Lemma 2.23. (Integral operators) *Let $K := K(\lambda, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^\nu \times \mathbb{T} \times \mathbb{T})$. Then the integral operator*

$$(\mathcal{R}u)(\varphi, x) := \int_{\mathbb{T}} K(\lambda, \varphi, x, y) u(\varphi, y)\, dy \tag{2.100}$$

is in $OPS^{-\infty}$ and, for all $m, s, \alpha \in \mathbb{N}$,

$$|\mathcal{R}|^{k_0,\gamma}_{-m,s,\alpha} \le C(m, s, \alpha, k_0) \|K\|^{k_0,\gamma}_{\mathcal{C}^{s+m+\alpha}}\,. \tag{2.101}$$

Proof. By (2.21) the symbol associated to the integral operator $\mathcal{R}$ is

$$a(\lambda, \varphi, x, j) = \int_{\mathbb{T}} K(\lambda, \varphi, x, y) e^{\mathrm{i}(y-x)j}\, dy\,, \quad \forall j \in \mathbb{Z}\,. \tag{2.102}$$

The function a is $\mathcal{C}^\infty$ in (φ, x) and k_0-times differentiable with respect to λ. For all $m, \beta, p \in \mathbb{N}$, $n \in \mathbb{N}^\nu$, $k \in \mathbb{N}^{\nu+1}$, one has

$$\begin{aligned}\partial_\lambda^k \partial_\varphi^n \partial_x^p \Delta_j^\beta a(\lambda, \varphi, x, \xi)(\mathrm{i}j)^{m+\beta} &= \sum\nolimits_{p_1+p_2=p} C_{p_1,p_2} \Delta_j^\beta \int_{\mathbb{T}} (\partial_\lambda^k \partial_\varphi^n \partial_x^{s_1} K)(\lambda, \varphi, x, y) \partial_y^{p_2+m+\beta}(e^{\mathrm{i}(y-x)j})\, dy \\ &= \sum\nolimits_{p_1+p_2=p} C_{p_1,p_2,m,\beta} \int_{\mathbb{T}} (\partial_\lambda^k \partial_\varphi^n \partial_x^{p_1} \partial_y^{p_2+m+\beta} K)(\lambda, \varphi, x, y) \Delta_j^\beta (e^{\mathrm{i}(y-x)j})\, dy\end{aligned}$$

integrating by parts. Using that $|\Delta_j^\beta(e^{\mathrm{i}xj})| = |e^{\mathrm{i}x\beta}(e^{\mathrm{i}x} - 1)^\beta)| \le 2^\beta$, $\forall \beta \in \mathbb{N}$, $x \in \mathbb{R}$, and recalling (2.6), we deduce that, for all $|k| \le k_0$,

$$|\partial_\lambda^k \partial_\varphi^n \partial_x^p \Delta_j^\beta a(\lambda, \varphi, x, j)| \le C(p, m, \beta) \gamma^{-|k|} \|K\|^{k_0,\gamma}_{\mathcal{C}^{p+m+\beta+|n|}} \langle j \rangle^{-m-\beta}\,. \tag{2.103}$$

Now we construct an extension $\widetilde{a}(\lambda, \varphi, x, \xi)$ of the symbol $a(\lambda, \varphi, x, j)$ as in (2.23), namely we define

$$\widetilde{a}(\lambda, \varphi, x, \xi) := \sum\nolimits_{j \in \mathbb{Z}} a(\lambda, \varphi, x, j) \zeta(\xi - j)\,, \quad \forall \xi \in \mathbb{R}\,. \tag{2.104}$$

Since $\widetilde{a}(\cdot, j) = a(\cdot, j)$ for all $j \in \mathbb{Z}$ one has that $\mathrm{Op}(\widetilde{a}) = \mathrm{Op}(a) = \mathcal{R}$. By (2.24) and (2.103) it results that for all $m, \beta, p \in \mathbb{N}$, $n \in \mathbb{N}^\nu$, $k \in \mathbb{N}^{\nu+1}$ with $|k| \le k_0$, there exist constants $C'(p, m, \beta) > 0$ such that

$$|\partial_\lambda^k \partial_\varphi^n \partial_x^p \partial_\xi^\beta \widetilde{a}(\lambda, \varphi, x, \xi)| \le C'(p, m, \beta) \gamma^{-|k|} \|K\|^{k_0,\gamma}_{\mathcal{C}^{p+m+\beta+|n|}} \langle \xi \rangle^{-m-\beta}\,. \tag{2.105}$$

By (2.2) and (2.105) we get: for all $m, s, \beta \in \mathbb{N}$, $|k| \le k_0$,

$$\begin{aligned}\|\partial_\xi^\beta \partial_\lambda^k \widetilde{a}(\lambda, \cdot, \xi)\|_s \langle \xi \rangle^{m+\beta} &\simeq \Big(\|\partial_\xi^\beta \partial_\lambda^k \widetilde{a}(\lambda, \cdot, \xi)\|_{L^2_\varphi L^2_x} + \|\partial_x^s \partial_\xi^\beta \partial_\lambda^k \widetilde{a}(\lambda, \cdot, \xi)\|_{L^2_\varphi L^2_x} \\ &\quad + \sup_{n \in \mathbb{Z}^\nu, |n|=s} \|\partial_\varphi^n \partial_\xi^\beta \partial_\lambda^k \widetilde{a}(\lambda, \cdot, \xi)\|_{L^2_\varphi L^2_x} \Big) \langle \xi \rangle^{m+\beta} \le_{m,s,\beta} \gamma^{-|k|} \|K\|^{k_0,\gamma}_{\mathcal{C}^{s+m+\beta}}\end{aligned}$$

that, recalling (2.36) and (2.35), proves (2.101). □

Remark 2.24. The extended symbol $\widetilde{a}$ in (2.104) can be explicitly written, using (2.102) and the Poisson summation formula, as $\widetilde{a}(\lambda, \varphi, x, \xi) = \int_{\mathbb{R}} K(\lambda, \varphi, x, y) \theta(y) e^{\mathrm{i}\xi y} dy$ where the test function $\theta \in \mathcal{D}(\mathbb{R})$ is defined after (2.23). This expression can be used as well to prove the estimate (2.101). □

An integral operator transforms into another integral operator under a changes of variables

$$Pu(\varphi, x) := u(\varphi, x + p(\varphi, x))\,. \tag{2.106}$$

Lemma 2.25. *Let $K(\lambda, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^\nu \times \mathbb{T} \times \mathbb{T})$ and $p(\lambda, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$. There exists $\delta := \delta(s_0, k_0) > 0$ such that if $\|p\|^{k_0,\gamma}_{2s_0+k_0+1} \le \delta$, then the integral operator $\mathcal{R}$ as in (2.100) transforms into the integral operator*

$$\big(P^{-1} \mathcal{R} P\big) u(\varphi, x) = \int_{\mathbb{T}} \tilde{K}(\lambda, \varphi, x, y) u(\varphi, y)\, dy \tag{2.107}$$

with a $\mathcal{C}^\infty$ Kernel $\widetilde{K}(\lambda, \cdot, \cdot, \cdot)$ which satisfies

$$\|\tilde{K}\|^{k_0,\gamma}_s \le C(s, k_0) \big(\|K\|^{k_0,\gamma}_{s+k_0} + \|p\|^{k_0,\gamma}_{s+k_0+1} \|K\|^{k_0,\gamma}_{s_0+k_0+1} \big) \qquad \forall s \ge s_0\,. \tag{2.108}$$

Proof. We denote by $z \mapsto z + q(\lambda, \varphi, z)$ the inverse diffeomorphism of $x \mapsto x + p(\lambda, \varphi, x)$, for all $\varphi \in \mathbb{T}^\nu$, $\lambda \in \Lambda_0$. We have $(\mathcal{R}P)u(\varphi, x) = \int_{\mathbb{T}} K(\lambda, \varphi, x, y)u(\varphi, y + p(\lambda, \varphi, y))\, dy$ and making the change of variable $z = y + p(\lambda, \varphi, y)$ we get (2.107) with Kernel

$$\tilde{K}(\lambda, \varphi, x, z) := \big(1 + \partial_z q(\lambda, \varphi, z)\big) K(\lambda, \varphi, x + q(\lambda, \varphi, x), z + q(\lambda, \varphi, z)) \,.$$

Since $p \in \mathcal{C}^\infty$, by Lemma 2.21 also $q \in \mathcal{C}^\infty$, therefore $\tilde{K}$ is $\mathcal{C}^\infty$. The estimate (2.108) for $\tilde{K}$ then follows by (2.72), (2.89), (2.90), (2.91) and by Lemma 2.14. $\square$

We now study the properties of the Hilbert transform $\mathcal{H}$. It can be defined through Fourier series by

$$\mathcal{H}\cos(jx) := \mathrm{sign}(j)\sin(jx), \quad \mathcal{H}\sin(jx) := -\,\mathrm{sign}(j)\cos(jx)\,, \ \ \forall j \in \mathbb{Z}\setminus\{0\}, \quad \mathcal{H}(1) := 0\,, \tag{2.109}$$

or in exponential basis

$$\mathcal{H}e^{\mathrm{i}jx} := -\mathrm{i}\,\mathrm{sign}(j) e^{\mathrm{i}jx}\,, \ \forall j \neq 0\,, \quad \mathcal{H}(1) := 0\,. \tag{2.110}$$

The Hilbert transform admits also an integral representation. Given a 2π-periodic function u its Hilbert transform is

$$\mathcal{H}u(x) := \frac{1}{2\pi}\,\text{p.v.}\int \frac{u(y)}{\tan(\frac12\,(x-y))}\,dy := \lim_{\varepsilon\to 0}\frac{1}{2\pi}\Big\{\int_{x-\pi}^{x-\varepsilon} + \int_{x+\varepsilon}^{x+\pi}\Big\}\frac{u(y)}{\tan(\frac12\,(x-y))}\,dy. \tag{2.111}$$

The commutator between the Hilbert transform $\mathcal{H}$ and the multiplication operator for a smooth function a is a regularizing operator in $OPS^{-\infty}$.

Lemma 2.26. *Let $a(\lambda, \cdot, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$. Then the commutator $[a, \mathcal{H}] \in OPS^{-\infty}$ and, for all $m, s, \alpha \in \mathbb{N}$,*

$$|\!|\!|[a, \mathcal{H}]|\!|\!|_{-m,s,\alpha}^{k_0,\gamma} \leq C(m, s, \alpha, k_0)\|a\|_{s+s_0+1+m+\alpha}^{k_0,\gamma}\,. \tag{2.112}$$

Proof. By (2.111) the commutator

$$(\mathcal{H}a - a\mathcal{H})u = \frac{1}{2\pi}\text{p.v.}\int \frac{(a(y)-a(x))u(y)}{\tan(\frac12\,(x-y))}\,dy = \frac{1}{2\pi}\int_{\mathbb{T}} K(x,y)u(y)\,dy$$

is an integral operator with $\mathcal{C}^\infty$ Kernel (note that the integral is no longer a principal value)

$$K(\lambda, \varphi, x, y) := \frac{a(\lambda, \varphi, y) - a(\lambda, \varphi, x)}{\tan((x-y)/2)} = \Big(\int_0^1 a_x(\lambda, \varphi, x + t(y-x))dt\Big)\frac{y-x}{\tan((x-y)/2)}\,.$$

Then (2.112) follows by Lemma 2.23 and the bound $\|K\|_{\mathcal{C}^s}^{k_0,\gamma} \leq_s \|K\|_{s+s_0}^{k_0,\gamma} \leq_s \|a\|_{s+s_0+1}^{k_0,\gamma}$ for all $s \geq 0$. $\square$

We now conjugate the Hilbert transform by a family of changes of variables as in (2.106), see also the Appendices H and I in [32] and [6]-Lemma B.5.

Lemma 2.27. *Let $p = p(\lambda, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$. There exists $\delta(s_0, k_0) > 0$ such that, if $\|p\|_{2s_0+k_0+1}^{k_0,\gamma} \leq \delta(s_0, k_0)$, then the operator $P^{-1}\mathcal{H}P - \mathcal{H}$ is an integral operator of the form*

$$(P^{-1}\mathcal{H}P - \mathcal{H})u(\varphi, x) = \int_{\mathbb{T}} K(\lambda, \varphi, x, z)u(\varphi, z)\,dz \tag{2.113}$$

where $K = K(\lambda, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^\nu \times \mathbb{T} \times \mathbb{T})$ satisfies

$$\|K\|_s^{k_0,\gamma} \leq C(s, k_0)\|p\|_{s+k_0+2}^{k_0,\gamma}\,, \quad \forall s \geq s_0\,. \tag{2.114}$$

Proof. The inverse diffeomorphism of $x \mapsto x+p(\varphi,x)$ has the form $z \mapsto z+q(\varphi,z)$. Changing the variable $z = y + p(\varphi, y)$ in the integral (2.111) gives

$$P^{-1}\mathcal{H}Pu(\varphi,x) = \frac{1}{2\pi}\,\mathrm{p.v.}\int \frac{u(\varphi,z)(1+\partial_z q(\lambda,\varphi,z))}{\tan(\frac{1}{2}\,[x-z+q(\lambda,\varphi,x)-q(\lambda,\varphi,z)])}\,dz\,.$$

As a consequence we get (2.113) (which is no longer a principal value) with Kernel

$$\begin{aligned} K(\lambda,\varphi,x,z) &:= \frac{1}{2\pi}\Big(\frac{1+\partial_z q(\lambda,\varphi,z)}{\tan(\frac{1}{2}[x-z+q(\lambda,\varphi,x)-q(\lambda,\varphi,z)])} - \frac{1}{\tan(\frac{1}{2}[x-z])}\Big) \\ &= -\frac{1}{\pi}\,\partial_z \log\Big(\frac{\sin(\frac{1}{2}[x-z+q(\lambda,\varphi,x)-q(\lambda,\varphi,z)])}{\sin(\frac{1}{2}[x-z])}\Big) \\ &= -\frac{1}{\pi}\partial_z \log\big(1+g(\lambda,\varphi,x,z)\big) \end{aligned} \tag{2.115}$$

(note that q is small) where the family of $\mathcal{C}^\infty$ functions

$$g(\lambda,\varphi,x,z) := \cos\Big(\frac{q(\lambda,\varphi,x)-q(\lambda,\varphi,z)}{2}\Big) - 1 + \cos\Big(\frac{x-z}{2}\Big)\frac{\sin(\frac{1}{2}[q(\lambda,\varphi,x)-q(\lambda,\varphi,z)])}{\sin(\frac{1}{2}[x-z])}$$

satisfies the estimate $\|g\|_s^{k_0,\gamma} \le_{s,k_0} \|q\|_{s+1}^{k_0,\gamma} \le_{s,k_0} \|p\|_{s+k_0+1}^{k_0,\gamma}$ using (2.91). Lemma 2.22 implies (2.114). □

2.4 Dirichlet-Neumann operator

We now present some fundamental properties of the Dirichlet-Neumann operator G defined in (1.4) that are used in the paper. There is a huge literature about it for which we refer to the recent work of Alazard-Delort [3]-[4] and the book of Lannes [33], and references therein. We remark that for our purposes it is sufficient to work in the class of smooth $\mathcal{C}^\infty$ profiles $\eta(x)$ because at each step of the Nash-Moser iteration we perform a $\mathcal{C}^\infty$-regularization.

The mapping $(\eta,\psi) \to G(\eta)\psi$ is linear with respect to ψ and nonlinear with respect to η. The derivative with respect to η ("shape derivative") is given by (see e.g. [33])

$$G'(\eta)[\hat\eta]\psi = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\{G(\eta+\varepsilon\hat\eta)\psi - G(\eta)\psi\} = -G(\eta)(B\hat\eta) - \partial_x(V\hat\eta) \tag{2.116}$$

where

$$B := B(\eta,\psi) := \frac{\eta_x\psi_x + G(\eta)\psi}{1+\eta_x^2}\,, \qquad V := V(\eta,\psi) := \psi_x - B\eta_x. \tag{2.117}$$

The vector $(V,B) = \nabla_{x,y}\Phi$ is the velocity potential evaluated at the free surface $(x,\eta(x))$.

Note also that $G(\eta)$ is an even operator according to Definition 2.3.

The Dirichlet-Neumann operator is a *pseudo-differential* operator of the form

$$G(\eta) = |D| + \mathcal{R}_G(\eta) \tag{2.118}$$

where $G(0) = |D|$ and the remainder $\mathcal{R}_G(\eta) \in OPS^{-\infty}$. The explicit representation of the integral Kernel of $\mathcal{R}_G(\eta)$ given by (2.129), (2.113), (2.115), has been taught to us by Baldi [5]. We use it to estimate the pseudo-differential norm $|\mathcal{R}_G(\eta)|_{-m,s,\alpha}^{k_0,\gamma}$. Note that the free profile $\eta(x) := \eta(\omega,\kappa,\varphi,x)$ as well as the potential $\psi(\omega,\kappa,\varphi,x)$ may depend also on the angles $\varphi \in \mathbb{T}^\nu$ and the parameters $\lambda := (\omega,\kappa) \in \mathbb{R}^\nu \times [\kappa_1,\kappa_2]$. For simplicity of notations we sometimes omit to write the dependence on φ, ω, κ.

Proposition 2.28. *Assume that $\partial_\lambda^k \eta(\lambda,\cdot,\cdot)$ is $\mathcal{C}^\infty$ for all $|k| \le k_0$. There exists $\delta := \delta(s_0,k_0) > 0$ such that, if*

$$\|\eta\|_{2s_0+2k_0+1}^{k_0,\gamma} \le \delta\,, \tag{2.119}$$

then the Dirichlet-Neumann operator $G(\eta)$ may be written as in (2.118) *where $\mathcal{R}_G(\eta)$ is an integral operator with $\mathcal{C}^\infty$ Kernel K_G (see* (2.100)*) which satisfies, for all $m, s, \alpha \in \mathbb{N}$, the estimate*

$$|\mathcal{R}_G(\eta)|_{-m,s,\alpha}^{k_0,\gamma} \le C(s,m,\alpha,k_0)\|K_G\|_{\mathcal{C}^{s+m+\alpha}}^{k_0,\gamma} \le C(s,m,\alpha,k_0)\|\eta\|_{s+s_0+2k_0+m+\alpha+3}^{k_0,\gamma} \,. \tag{2.120}$$

Let $s_1 \ge 2s_0+1$. There exists $\delta(s_1) > 0$ such that, the map $\{\|\eta\|_{s_1+6} < \delta(s_1)\} \to H^{s_1}(\mathbb{T}^\nu \times \mathbb{T} \times \mathbb{T})$, $\eta \mapsto K_G(\eta)$, is $\mathcal{C}^1$.

Remark 2.29. Note that the assumption (2.119) in low norm $\| \ \|_{2s_0+2k_0+1}^{k_0,\gamma}$ implies the estimate (2.120) for any $s \in \mathbb{N}$. The estimate $\|\partial_\eta K_G[\widehat{\eta}]\|_{s_1} \le_{s_1} \|\widehat{\eta}\|_{s_1+6}$ is used in section 6 (in particular in section 6.2) with a Sobolev index s_1 which has to be considered fixed, see (6.11). A sharper tame version of this estimate could be proved, but it is not needed. Note also that it does not involve the $\| \ \|_{s_1}^{k_0,\gamma}$ norm. □

The rest of this section is devoted to the proof of Proposition 2.28.
In order to analyze the Dirichlet-Neumann operator it is convenient to transform the boundary value problem (1.2) defined in the free domain $\{(x,y) : y < \eta(x)\}$ into an elliptic problem in the lower half-plane $\Sigma_0 := \{(X,Y) : Y < 0\}$ via a conformal diffeomorphism

$$x = U(X,Y), \quad y = V(X,Y)\,. \tag{2.121}$$

The following conformal transformation (2.122), the formulation of the problem as the fixed point equation (2.125), Lemma 2.31 and (2.129) is due to Baldi [5].

The conformal transformation. Let $p : \mathbb{R} \to \mathbb{R}$ be a smooth 2π-periodic function with zero average and $\|\partial_X^2 p\|_{L^2(\mathbb{T})} \le c_0 := 1/(2\sqrt{2\pi})$. We define the functions

$$U(X,Y) := X + \sum_{k\neq 0} p_k\, e^{|k|Y}\, e^{\mathrm{i}kX}, \qquad V(X,Y) := Y + \sum_{k\neq 0} \mathrm{i}\,\mathrm{sign}(k)\, p_k\, e^{|k|Y}\, e^{\mathrm{i}kX} + c \tag{2.122}$$

with $c \in \mathbb{R}$. The functions U and V are both harmonic on Σ_0 and satisfy the Cauchy-Riemann equations $U_X = V_Y$, $U_Y = -V_X$ so that $U + \mathrm{i}V$ is holomorphic on Σ_0. The gradient $(U_X, U_Y) \to (1,0)$ as $Y \to -\infty$.

Since, $\forall Y \le 0$, $\|U_{XX}(X,Y)\|_{L^2(\mathbb{T})} \le \|p_{XX}\|_{L^2(\mathbb{T})} \le c_0$, it results $U_X \ge 1/2$ on Σ_0, and, by $V_Y = U_X \ge 1/2$, we also get $V(X,Y) < V(X,0)$ for $Y < 0$. The Jacobian

$$\det\begin{pmatrix} U_X & U_Y \\ V_X & V_Y \end{pmatrix} = \det\begin{pmatrix} U_X & U_Y \\ -U_Y & U_X \end{pmatrix} = U_X^2 + U_Y^2 \ge \frac{1}{4}\,, \quad \forall (X,Y) \in \Sigma_0\,,$$

so that $U + \mathrm{i}V$ is a global diffeomorphism from Σ_0 onto its image. Since $U(X,Y) - X$ is 2π-periodic in X (see (2.122)) the map $U + \mathrm{i}V$ is the lift of a diffeomorphism from $\mathbb{T} \times (-\infty, 0]$ onto its image. The image of the map $U + \mathrm{i}V$ is the subset of $\mathbb{C} \simeq \mathbb{R}^2$ that is below the profile described parametrically by

$$(U(X,0), V(X,0)) = (X + p(X), -\mathcal{H}p(X) + c) \tag{2.123}$$

where $\mathcal{H}$ is the Hilbert transform in (2.110). The profile (2.123) coincides with the graph $Y = \eta(X)$ if

$$-\mathcal{H}p(X) + c = \eta(X + p(X))\,, \quad \forall X \in \mathbb{R}\,. \tag{2.124}$$

Since, by (2.110), the range of the Hilbert transform $\mathcal{H}$ is the space of functions with zero average and $\mathcal{H}^2 = -\Pi$ where $\Pi[f] := f - f_0$, the equation (2.124) is equivalent to

$$c = \frac{1}{2\pi}\int_0^{2\pi} \eta(X + p(X))\, dX$$

and

$$p(X) = \mathcal{H}[\eta(X + p(X))]\,. \tag{2.125}$$

Lemma 2.30. *Let η satisfy $\partial_\lambda^k \eta(\lambda,\cdot) \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$, for all $|k| \le k_0$. There exists $\delta := \delta(s_0,k_0) > 0$, such that, if $\|\eta\|_{2s_0+k_0+1}^{k_0,\gamma} \le \delta$, then there exists a unique solution $p = p(\lambda,\cdot)$ of (2.125) satisfying the estimates*

$$\|p\|_s \le_s \|\eta\|_s \,, \qquad \|p\|_s^{k_0,\gamma} \le_s \|\eta\|_{s+k_0}^{k_0,\gamma} \,, \qquad \forall s \ge s_0 \,. \tag{2.126}$$

Let $s_1 \ge 2s_0+1$. There exists $\delta(s_1) > 0$ such that the map $\{\|\eta\|_{s_1+2} < \delta(s_1)\} \to H^{s_1}$, $\eta \mapsto p(\eta)$, is $\mathcal{C}^1$.

Proof. We find a solution of (2.125) as a fixed point of the map

$$p(\varphi,X) \mapsto \Phi(p)(\varphi,X) := \mathcal{H}[\eta(\varphi, X + p(\varphi,X))] \,.$$

For any $n \in \mathbb{N}$, we consider the finite dimensional subspace $E_n := \mathrm{span}\{e^{\mathrm{i}(\ell\cdot\varphi + jx)} : |(\ell,j)| \le n\}$ and the regularized map $\Phi_n := \Pi_n \Phi : E_n \to E_n$ where Π_n denotes the L^2-orthogonal projector on E_n. We show that there is $r > 0$ small, such that, for any $n \in \mathbb{N}$, the map

$$\Phi_n : \mathcal{B}_{2s_0+1}(r) \cap E_n \to \mathcal{B}_{2s_0+1}(r) \cap E_n \,, \quad \mathcal{B}_{2s_0+1}(r) := \big\{ p \in H^{2s_0+1} : \|p\|_{2s_0+1} \le r \big\} \,,$$

is a contraction. We fix $r > 0$ such that $\|p\|_{\mathcal{C}^{s_0+1}} \le C(s_0)\|p\|_{2s_0+1} \le 1/2$, for all $p \in \mathcal{B}_{2s_0+1}(r)$, i.e $r := 1/(2C(s_0))$, so that the hyphothesis (2.87) of Lemma 2.21 is fulfilled. Then, using that $\mathcal{H}$ is an isometry on the Sobolev spaces H^s (see (2.110)), that $\|\Pi_n h\|_s \le \|h\|_s$, and applying (2.88), we get

$$\|\Phi_n(p)\|_{2s_0+1} \le \|\eta(\cdot + p(\cdot))\|_{2s_0+1} \le C_1(s_0)\|\eta\|_{2s_0+1} \le r$$

taking $\|\eta\|_{2s_0+1} \le r/C_1(s_0)$. Moreover for any $p_1\,,\,p_2 \in \mathcal{B}_{2s_0+1}(r) \cap E_n$, we have

$$\|\Phi_n(p_1) - \Phi_n(p_2)\|_{2s_0+1} \le C(s_0)\|\eta\|_{2s_0+2}\|p_1 - p_2\|_{2s_0+1} \le \|p_1 - p_2\|_{2s_0+1}/2 \,,$$

by taking $C(s_0)\|\eta\|_{2s_0+2} \le 1/2$. Then, by the contraction mapping theorem there exists a unique fixed point solution $p_n \in \mathcal{B}_{2s_0+1}(r) \cap E_n$ solving $\Phi_n(p_n) = p_n$. Note that $p_n \in E_n \subset \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$. Using again that the Hilbert transform is a unitary operator, and the estimate (2.88), we get, for all $s \ge s_0$

$$\|p_n\|_s = \|\Phi_n(p_n)\|_s = \|\Pi_n \mathcal{H}\eta(\cdot + p_n(\cdot))\| \le C(s)\|\eta\|_s + C(s_0)\|p_n\|_s\|\eta\|_{s_0+1} \tag{2.127}$$

which implies $\|p_n\|_s \le 2C(s)\|\eta\|_s$ taking $C(s_0)\|\eta\|_{s_0+1} \le 1/2$. Since $H^s \hookrightarrow H^{s-1}$ compactly, for any $s \ge s_0$, the sequence p_n converges strongly in H^s (up to subsequence) to a function $p \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$ which satisfies $\|p\|_s \le 2C(s)\|\eta\|_s$ for any $s \ge s_0$. The function p solves the equation (2.125) because

$$\begin{aligned}
\|\Phi(p) - \Phi_n(p_n)\|_{s_0} &\le \|\Pi_n\mathcal{H}\eta(\cdot + p(\cdot)) - \Pi_n\mathcal{H}\eta(\cdot + p_n(\cdot))\|_{s_0} + \|(\mathrm{Id} - \Pi_n)\mathcal{H}\eta(\cdot + p(\cdot))\|_{s_0} \\
&\le_{s_0} \|\eta\|_{s_0+1}\|p - p_n\|_{s_0} + \frac{1}{n}\|\eta\|_{s_0+1}(1 + \|p\|_{s_0+1}) \le_{s_0} \|p - p_n\|_{s_0} + \frac{1}{n} \to 0
\end{aligned}$$

as $n \to +\infty$. This implies that $\Phi(p) = p$. Arguing as in Lemma 2.21 one can prove that if $\partial_\lambda^k \eta(\lambda,\cdot) \in \mathcal{C}^\infty$ for all $|k| \le k_0$, then also $\partial_\lambda^k p(\lambda,\cdot) \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$, for all $|k| \le k_0$. The second estimate in (2.126) can be proved as the estimate (2.91) in Lemma 2.21, using the condition $\|\eta\|_{s_0+k_0+1}^{k_0,\gamma} \le \delta(s_0,k_0)$ for some $\delta(s_0,k_0) > 0$ small enough.

The differentiability of $\eta \mapsto p(\eta)$ follows by the implicit function theorem using the $\mathcal{C}^1$ map

$$F : H^{s_1+2} \times H^{s_1} \to H^{s_1} \,, \quad F(\eta,p)(\varphi,X) := p(\varphi,X) - \mathcal{H}\big[\eta(\varphi, X + p(\varphi,X))\big] \,.$$

Since $F(0,0) = 0$ and $\partial_p F(0,0) = \mathrm{Id}$, by the implicit function theorem there exists $\delta(s_1) > 0$ and a $\mathcal{C}^1$ map $\{\|\eta\|_{s_1+2} \le \delta(s_1)\} \ni \eta \mapsto p(\eta) \in H^{s_1}$, such that $F(\eta, p(\eta)) = 0$. $\square$

We transform (1.2) via the conformal diffeomorphism (2.122). Denote

$$Pu(X) := u(X + p(X)) \,.$$

The potential $\phi(X,Y) := \Phi(U(X,Y), V(X,Y))$ satisfies, using also (2.123)-(2.124),

$$\Delta\phi = 0 \text{ in } \{Y < 0\} \,, \quad \phi(X,0) = (P\psi)(X) \,, \quad \nabla\phi \to (0,0) \text{ as } Y \to -\infty \,. \tag{2.128}$$

Recall that the Dirichlet-Neumann operator at the flat surface $Y = 0$ is $\partial_X \mathcal{H}$.

Lemma 2.31. $G(\eta) = \partial_x P^{-1}\mathcal{H}P$.

Proof. Since $\eta(U(X,0)) = V(X,0)$ (see (2.124)) we derive $-U_Y = V_X = \eta_x U_X$ on $Y=0$. Moreover, by

$$\Phi_x = \frac{\phi_X U_X + \phi_Y U_Y}{U_X^2 + U_X^2}\,, \quad \Phi_y = \frac{\phi_Y U_X - \phi_X U_Y}{U_X^2 + U_X^2}\,,$$

and the definition (1.4) of the Dirichlet-Neumann operator we get

$$\begin{aligned} G(\eta)\psi(x) &= \frac{1}{U_X^2 + U_Y^2}\Big(\phi_X(-U_Y - \eta_x U_X) + \phi_Y(U_X - \eta_x U_Y)\Big) = \frac{1}{U_X(X,0)}\,\phi_Y(X,0) \\ &\overset{(2.122),(2.128)}{=} \frac{1}{1+p_X(X)}\,\partial_X \mathcal{H}(P\psi)(X) = \Big\{\frac{1}{1+p_X}\,\partial_X\mathcal{H}P\psi\Big\}(x+\tilde p(x)) \end{aligned}$$

where $X = x + \tilde p(x)$ is the inverse diffeomorphism of $x = X + p(X)$. In operatorial notation we have

$$\begin{aligned} G(\eta) &= P^{-1}\frac{1}{1+p_X}\,\partial_X\mathcal{H}P = \frac{1}{1+P^{-1}p_X}\,P^{-1}\partial_X P\,P^{-1}\mathcal{H}P \\ &= \frac{1}{1+P^{-1}p_X}\,(1+P^{-1}p_X)\,\partial_x P^{-1}\mathcal{H}P = \partial_x P^{-1}\mathcal{H}P \end{aligned}$$

by the rule $P^{-1}\partial_X P = (1+P^{-1}p_X)\,\partial_x$ for the changes of coordinates. ◻

Lemma 2.31 provides the representation (2.118) of the Dirichlet-Neumann operator with

$$\mathcal{R}_G(\eta) := \partial_x\big(P^{-1}\mathcal{H}P - \mathcal{H}\big)\,. \tag{2.129}$$

By Lemma 2.27, in particular by formula (2.115), the operator $\mathcal{R}_G(\eta)$ is an integral operator with kernel

$$K_G := K_G(\eta) := -\frac{1}{\pi}\partial_{xz}\log\big(1 + g(\varphi,x,z)\big) \tag{2.130}$$

where

$$g(\varphi,x,z) := \cos\Big(\frac{q(\lambda,\varphi,x) - q(\lambda,\varphi,z)}{2}\Big) - 1 + \cos\Big(\frac{x-z}{2}\Big)\frac{\sin(\frac12[q(\lambda,\varphi,x) - q(\lambda,\varphi,z)])}{\sin(\frac12[x-z])} \tag{2.131}$$

and $x \mapsto x + q(\varphi,x)$ is the inverse diffeomorphism of $X \mapsto X + p(\varphi,X)$ (the functions p,q depend on η).

PROOF OF PROPOSITION 2.28 CONCLUDED. By (2.119) we apply Lemma 2.30 and then (2.126) implies $\|p\|^{k_0,\gamma}_{2s_0+k_0+1} \leq_{s_0} \|\eta\|^{k_0,\gamma}_{2s_0+2k_0+1}$. Hence, by (2.119), the smallness assumption of Lemma 2.27 is verified. Hence the estimate (2.120) follows by (2.101), (2.114), (2.126).

We now prove that the function $\{\|\eta\|_{s_1+6} \leq \delta(s_1)\} \mapsto H^{s_1}(\mathbb{T}^\nu\times\mathbb{T}\times\mathbb{T})$, $\eta \mapsto K_G(\eta)$ is $\mathcal{C}^1$. Indeed, by applying Lemma 2.30 (with s_1+4 instead of s_1), the map $\{\|\eta\|_{s_1+6} < \delta(s_1)\} \mapsto H^{s_1+4}$, $\eta \mapsto p(\eta)$ is $\mathcal{C}^1$. Then, since $q(\varphi,x) = -p(\varphi, x+q(\varphi,x))$, by the implicit function theorem, for p small in $\|\cdot\|_{s_1+4}$-norm, also the map $p \mapsto q(p) \in H^{s_1+2}$ is $\mathcal{C}^1$. By composition, the claim follows by recalling (2.130), (2.131). ◻

To conclude we provide the following tame estimates for the Dirichlet Neumann operator:

Lemma 2.32. *There is $\delta(s_0,k_0) > 0$ such that, if $\|\eta\|^{k_0,\gamma}_{2s_0+2k_0+5} \leq \delta(s_0,k_0)$, then, for all $s \geq s_0$*

$$\|\big(G(\eta) - |D|\big)\psi\|^{k_0,\gamma}_s \leq_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+s_0+2k_0+3}\|\psi\|^{k_0,\gamma}_{s_0} + \|\eta\|^{k_0,\gamma}_{2s_0+2k_0+3}\|\psi\|^{k_0,\gamma}_s\,, \tag{2.132}$$

$$\|G'(\eta)[\widehat\eta]\psi\|^{k_0,\gamma}_s \leq_{s,k_0} \|\psi\|^{k_0,\gamma}_{s+2}\|\widehat\eta\|^{k_0,\gamma}_{s_0+1} + \|\psi\|^{k_0,\gamma}_{s_0+2}\|\widehat\eta\|^{k_0,\gamma}_{s+1} + \|\eta\|^{k_0,\gamma}_{s+s_0+2k_0+4}\|\widehat\eta\|^{k_0,\gamma}_{s_0+1}\|\psi\|^{k_0,\gamma}_{s_0+2}\,, \tag{2.133}$$

$$\begin{aligned} \|G''(\eta)[\widehat\eta,\widehat\eta]\psi\|^{k_0,\gamma}_s \leq_{s,k_0}\ & \|\psi\|^{k_0,\gamma}_{s+3}\big(\|\widehat\eta\|^{k_0,\gamma}_{s_0+2}\big)^2 + \|\psi\|^{k_0,\gamma}_{s_0+3}\|\widehat\eta\|^{k_0,\gamma}_{s+2}\|\widehat\eta\|^{k_0,\gamma}_{s_0+2} \\ & + \|\eta\|^{k_0,\gamma}_{s+s_0+2k_0+5}\|\psi\|^{k_0,\gamma}_{s_0+3}\big(\|\widehat\eta\|^{k_0,\gamma}_{s_0+2}\big)^2\,. \end{aligned} \tag{2.134}$$

Proof. The estimate (2.132) follows by the formula (2.118), the bound (2.120) (for $m = \alpha = 0$) and Lemmata 2.13, 2.14. The estimate (2.133) follows by the shape derivative formula (2.116), applying (2.132), (2.72) and the fact that the functions B,V defined in (2.117) satisfy

$$\|B\|^{k_0,\gamma}_s,\ \|V\|^{k_0,\gamma}_s \leq_s \|\psi\|^{k_0,\gamma}_{s+1} + \|\eta\|^{k_0,\gamma}_{s+s_0+2k_0+3}\|\psi\|^{k_0,\gamma}_{s_0+1}\,.$$

The estimate (2.134) follows by differentiating the shape derivative formula (2.116) and by applying the same kind of arguments. ◻

3 Degenerate KAM theory

In this section we verify that it is possible to develop degenerate KAM theory as in [10].

Definition 3.1. *A function* $f := (f_1, \dots, f_N) : [\kappa_1, \kappa_2] \to \mathbb{R}^N$ *is called non-degenerate if, for any vector* $c := (c_1, \dots, c_N) \in \mathbb{R}^N \setminus \{0\}$ *the function* $f \cdot c = f_1 c_1 + \dots + f_N c_N$ *is not identically zero on the whole interval* $[\kappa_1, \kappa_2]$.

From a geometric point of view, f non-degenerate means that the image of the curve $f([\kappa_1, \kappa_2]) \subset \mathbb{R}^N$ is not contained in any hyperplane of $\mathbb{R}^N$. For such reason a curve f which satisfies the non-degeneracy property of Definition 3.1 is also referred as an *essentially non-planar* curve, or a curve with *full torsion*. For a smooth degenerate function f, differentiating $(N-1)$ times the identity $f(\kappa) \cdot c = 0$, we see that

$$f(\kappa) \text{ degenerate} \quad \Longrightarrow \quad f(\kappa), (\partial_\kappa f)(\kappa), \dots, (\partial_\kappa^{N-1} f)(\kappa) \text{ are linearly dependent } \forall \kappa \in [\kappa_1, \kappa_2]\,. \tag{3.1}$$

Given $\mathbb{S}^+ \subset \mathbb{N}^+$ we denote the unperturbed tangential and normal frequency vectors by

$$\vec{\omega}(\kappa) := (\omega_j(\kappa))_{j \in \mathbb{S}^+}\,, \quad \vec{\Omega}(\kappa) := (\Omega_j(\kappa))_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+} := (\omega_j(\kappa))_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+}\,. \tag{3.2}$$

Lemma 3.1. *The frequency vectors* $\vec{\omega}(\kappa) \in \mathbb{R}^\nu$, $(\sqrt{\kappa}, \vec{\omega}(\kappa)) \in \mathbb{R}^{\nu+1}$ *and*

$$(\vec{\omega}(\kappa), \Omega_j(\kappa)) \in \mathbb{R}^{\nu+1}, j \in \mathbb{N}^+ \setminus \mathbb{S}^+, \quad (\vec{\omega}(\kappa), \Omega_j(\kappa), \Omega_{j'}(\kappa)) \in \mathbb{R}^{\nu+2}\,,\ \forall j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+\,,\ j \neq j'\,,$$

are non-degenerate.

Proof. Set $\lambda_0(\kappa) := \sqrt{\kappa}$ and $\lambda_j(\kappa) := \sqrt{j(1+\kappa j^2)}$, $j \geq 1$. The lemma follows by proving that, for any N, for any $\lambda_{j_1}(\kappa), \dots, \lambda_{j_N}(\kappa)$, with $j_1, \dots, j_N \geq 0$, $j_i \neq j_k$ for all $i \neq k$, the function $[\kappa_1, \kappa_2] \ni \kappa \mapsto (\lambda_{j_1}(\kappa), \dots, \lambda_{j_N}(\kappa)) \in \mathbb{R}^N$ is non-degenerate according to Definition 3.1. By (3.1) it is sufficient to prove that the $N \times N$-matrix

$$\mathcal{A}(\kappa) := \begin{pmatrix} \lambda_{j_1}(\kappa) & \lambda_{j_2}(\kappa) & \dots & \lambda_{j_N}(\kappa) \\ \partial_\kappa \lambda_{j_1}(\kappa) & \partial_\kappa \lambda_{j_2}(\kappa) & \dots & \partial_\kappa \lambda_{j_N}(\kappa) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_\kappa^{N-1} \lambda_{j_1}(\kappa) & \partial_\kappa^{N-1} \lambda_{j_2}(\kappa) & \dots & \partial_\kappa^{N-1} \lambda_{j_N}(\kappa) \end{pmatrix}$$

is non-singular at some value of $\kappa \in [\kappa_1, \kappa_2]$. Actually, it turns out to be non-singular for all $\kappa \in [\kappa_1, \kappa_2]$.

Arguing by induction we get the following formula for the derivatives of $\lambda_j(\kappa)$: for all $r \geq 1$

$$\partial_\kappa^r \lambda_0(\kappa) = \frac{(-1)^{r+1}}{2^r} (2r-3)!! \kappa^{-\frac{2r-1}{2}} = (-1)^{r+1} (2r-3)!! \lambda_0(\kappa) x_0^r\,, \quad x_0 := \frac{1}{2\kappa}\,, \tag{3.3}$$

where $(-1)!! := 1$, $1!! := 1$ and if $n > 1$ is odd $n!! := \prod_{k=0}^{\frac{n-1}{2}} (n-2k)$. For all $j, r \geq 1$

$$\partial_\kappa^r \lambda_j(\kappa) = \frac{\sqrt{j} j^{2r}}{2^r} (-1)^{r+1} (2r-3)!! (1+\kappa j^2)^{-\frac{2r-1}{2}} = (-1)^{r+1} (2r-3)!! \lambda_j(\kappa) x_j^r\,, \quad x_j := \frac{j^2}{2(1+\kappa j^2)}\,. \tag{3.4}$$

Using the previous formulas (3.3)-(3.4) and the multi-linearity of the determinant we get

$$\det(\mathcal{A}(\kappa)) = \prod_{k=1}^{N} \lambda_{j_k}(\kappa) \prod_{r=1}^{N-1} (-1)^{r+1} (2r-3)!! \det(\mathcal{B}(\kappa))$$

where the $N \times N$ matrix

$$\mathcal{B}(\kappa) := \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_{j_1} & x_{j_2} & \dots & x_{j_N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{j_1}^{N-1} & x_{j_2}^{N-1} & \dots & x_{j_N}^{N-1} \end{pmatrix}$$

is the Vandermonde matrix. Its determinant is

$$\det(\mathcal{B}(\kappa)) = \prod_{1\le i<k\le N} (x_{j_i} - x_{j_k}) \,. \tag{3.5}$$

By the definition of x_j in (3.3)-(3.4), we have that, for all $\kappa \in [\kappa_1, \kappa_2]$,

$$x_j - x_{j'} = \frac{1}{2}\frac{j^2 - j'^2}{(1+\kappa j^2)(1+\kappa j'^2)} \neq 0 \,, \forall j \neq j', j, j' \geq 1 \,, \quad x_j - x_0 = -\frac{1}{2\kappa(1+\kappa j^2)} \neq 0 \,, \ \forall j \geq 1 \,.$$

Thus, by (3.5) the determinant $\det(\mathcal{B}(\kappa)) \neq 0$ and so $\det(\mathcal{A}(\kappa)) \neq 0$, $\forall \kappa \in [\kappa_1, \kappa_2]$, proving the lemma. □

In the next Proposition 3.2 we deduce, by the qualitative non-degeneracy condition proved in Lemma 3.1, the analyticity and the asymptotics of the linear frequencies $\kappa \mapsto \omega_j(\kappa) = \sqrt{j(1+\kappa j^2)}$, the quantitative bounds (3.6)-(3.9). The proof is similar to [10]. It does not follow immediately [10] because the linear frequencies $\omega_j(\kappa)$ depend on the parameter κ also at the highest order $O(\sqrt{\kappa} j^{3/2})$.

Proposition 3.2. *There exist $k_0 \in \mathbb{N}$, $\rho_0 > 0$ such that, for any $\kappa \in [\kappa_1, \kappa_2]$,*

$$\max_{k\le k_0} |\partial_\kappa^k \{\vec\omega(\kappa)\cdot \ell\}| \geq \rho_0 \langle \ell \rangle \,, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \tag{3.6}$$

$$\max_{k\le k_0} |\partial_\kappa^k \{\vec\omega(\kappa)\cdot \ell + \Omega_j(\kappa)\}| \geq \rho_0 \langle \ell \rangle \,, \quad \forall \ell \in \mathbb{Z}^\nu, \ j \in \mathbb{N}^+ \setminus \mathbb{S}^+, \tag{3.7}$$

$$\max_{k\le k_0} |\partial_\kappa^k \{\vec\omega(\kappa)\cdot \ell + \Omega_j(\kappa) - \Omega_{j'}(\kappa)\}| \geq \rho_0 \langle \ell \rangle \,, \quad \forall (\ell, j, j') \neq (0, j, j), \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+, \tag{3.8}$$

$$\max_{k\le k_0} |\partial_\kappa^k \{\vec\omega(\kappa)\cdot \ell + \Omega_j(\kappa) + \Omega_{j'}(\kappa)\}| \geq \rho_0 \langle \ell \rangle \,, \quad \forall \ell \in \mathbb{Z}^\nu, j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \,. \tag{3.9}$$

We call (following [41]) ρ_0 the "amount of non-degeneracy" and k_0 the "index of nondegeneracy".

Proof. All the inequalities (3.6)-(3.9) are proved by contradiction.

Proof of (3.6). Suppose that $\forall k_0 \in \mathbb{N}$, $\forall \rho_0 > 0$ there exist $\ell \in \mathbb{Z}^\nu \setminus \{0\}$, $\kappa \in [\kappa_1, \kappa_2]$ such that $\max_{k\le k_0} |\partial_\kappa^k \{\vec\omega(\kappa)\cdot\ell\}| < \rho_0 \langle \ell \rangle$. This implies that for all $m \in \mathbb{N}$, taking $\rho_0 = \frac{1}{1+m}$, there exist $\ell_m \in \mathbb{Z}^\nu \setminus \{0\}$, $\kappa_m \in [\kappa_1, \kappa_2]$ such that

$$\max_{k\le m} |\partial_\kappa^k \{\vec\omega(\kappa_m)\cdot \ell_m\}| < \frac{1}{1+m} \langle \ell_m \rangle$$

and therefore

$$\forall k \in \mathbb{N}, \quad m \geq k \,, \quad \Big|\partial_\kappa^k \vec\omega(\kappa_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle}\Big| < \frac{1}{1+m} \,. \tag{3.10}$$

The sequences $(\kappa_m)_{m\in\mathbb{N}} \subset [\kappa_1, \kappa_2]$ and $(\ell_m / \langle \ell_m \rangle)_{m\in\mathbb{N}} \subset \mathbb{R}^\nu \setminus \{0\}$ are bounded. By compactness there exists a sequence $m_h \to +\infty$ such that $\kappa_{m_h} \to \bar\kappa \in [\kappa_1, \kappa_2]$, $\ell_{m_h}/\langle \ell_{m_h} \rangle \to \bar c \neq 0$. Passing to the limit in (3.10) for $m_h \to +\infty$ we deduce that $\partial_\kappa^k \vec\omega(\bar\kappa) \cdot \bar c = 0$, $\forall k \in \mathbb{N}$. We conclude that the analytic function $\kappa \mapsto \vec\omega(\kappa) \cdot \bar c$ is identically zero. Since $\bar c \neq 0$, this is in contradiction with Lemma 3.1.

Proof of (3.7). Recalling that $\Omega_j(\kappa) = \sqrt{j(1+\kappa j^2)}$, we have the expansion

$$\Omega_j(\kappa) = \sqrt{\kappa} j^{\frac{3}{2}} + \frac{c_j(\kappa)}{\sqrt{\kappa j}} \,, \qquad c_j(\kappa) := \frac{1}{2}\int_0^1 \big(1 + \frac{t}{\kappa j^2}\big)^{-1/2} dt \tag{3.11}$$

where

$$\forall k \in \mathbb{N} \,, \quad \big|\partial_\kappa^k \frac{c_j(\kappa)}{\sqrt{\kappa}}\big| \leq C(k) \tag{3.12}$$

uniformly in $j \in \mathbb{S}^c$, $\kappa \in [\kappa_1, \kappa_2]$.

First of all note that $\forall \kappa \in [\kappa_1, \kappa_2]$, we have $|\vec\omega(\kappa)\cdot\ell + \Omega_j(\kappa)| \geq \Omega_j(\kappa) - |\vec\omega(\kappa)\cdot\ell| \geq \sqrt{\kappa_1} j^{3/2} - C|\ell| \geq |\ell|$ if $j^{3/2} \geq C_0|\ell|$ for some $C_0 > 0$. Therefore in (3.7) we can restrict to the indices $(\ell, j) \in \mathbb{Z}^\nu \times (\mathbb{N}^+ \setminus \mathbb{S}^+)$ satisfying

$$j^{\frac{3}{2}} < C_0 |\ell| \,. \tag{3.13}$$

Arguing by contradiction (as for proving (3.6)), we suppose that for all $m \in \mathbb{N}$ there exist $\ell_m \in \mathbb{Z}^\nu$, $j_m \in \mathbb{N}^+ \setminus \mathbb{S}^+$ and $\kappa_m \in [\kappa_1, \kappa_2]$, such that

$$\max_{k \le m} \Big| \partial_\kappa^k \Big\{ \vec{\omega}(\kappa_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle} + \frac{\Omega_{j_m}(\kappa_m)}{\langle \ell_m \rangle} \Big\} \Big| < \frac{1}{1+m}$$

and therefore

$$\forall k \in \mathbb{N}, \quad m \ge k \,, \quad \Big| \partial_\kappa^k \Big\{ \vec{\omega}(\kappa_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle} + \frac{\Omega_{j_m}(\kappa_m)}{\langle \ell_m \rangle} \Big\} \Big| < \frac{1}{1+m} \,. \tag{3.14}$$

Since the sequences $(\kappa_m)_{m \in \mathbb{N}} \subset [\kappa_1, \kappa_2]$ and $(\ell_m / \langle \ell_m \rangle)_{m \in \mathbb{N}} \in \mathbb{R}^\nu$ are bounded, there exist $m_h \to +\infty$ such that

$$\kappa_{m_h} \to \bar{\kappa} \in [\kappa_1, \kappa_2] \,, \quad \frac{\ell_{m_h}}{\langle \ell_{m_h} \rangle} \to \bar{c} \in \mathbb{R}^\nu \,. \tag{3.15}$$

We now distinguish two cases:

Case 1: $(\ell_{m_h}) \subset \mathbb{Z}^\nu$ is bounded. In this case, up to subsequence, $\ell_{m_h} \to \bar{\ell} \in \mathbb{Z}^\nu$, and since $|j_m| \le C |\ell_m|^{\frac{2}{3}}$ for all m (see (3.13)), we have $j_{m_h} \to \bar{\jmath}$. Passing to the limit for $m_h \to +\infty$ in (3.14) we deduce, by (3.15), that

$$\partial_\kappa^k \big\{ \vec{\omega}(\bar{\kappa}) \cdot \bar{c} + \Omega_{\bar{\jmath}}(\bar{\kappa}) \langle \bar{\ell} \rangle^{-1} \big\} = 0 \,, \quad \forall k \in \mathbb{N} \,.$$

Therefore the analytic function $\kappa \mapsto \vec{\omega}(\kappa) \cdot \bar{c} + \langle \bar{\ell} \rangle^{-1} \Omega_{\bar{\jmath}}(\kappa)$ is identically zero. Since $(\bar{c}, \langle \ell \rangle^{-1}) \neq 0$ this is in contradiction with Lemma 3.1.

Case 2: (ℓ_{m_h}) is unbounded. Up to subsequence $|\ell_{m_h}| \to +\infty$. In this case the constant $\bar{c} \neq 0$ in (3.15). Moreover, by (3.13), we also have that, up to subsequences,

$$j_{m_h}^{\frac{3}{2}} \langle \ell_{m_h} \rangle^{-1} \to \bar{d} \in \mathbb{R} \,. \tag{3.16}$$

By (3.11), (3.12), (3.15), (3.16), we get

$$\frac{\Omega_{j_{m_h}}(\kappa_{m_h})}{\langle \ell_{m_h} \rangle} = \sqrt{\kappa_{m_h}} \frac{j_{m_h}^{\frac{3}{2}}}{\langle \ell_{m_h} \rangle} + \frac{c_{j_{m_h}}(\kappa_{m_h})}{\sqrt{\kappa_{m_h} j_{m_h}} \langle \ell_{m_h} \rangle} \to \bar{d} \sqrt{\bar{\kappa}} \,, \qquad \partial_\kappa^k \frac{\Omega_{j_{m_h}}(\kappa_{m_h})}{\langle \ell_{m_h} \rangle} \to \bar{d} \, \partial_\kappa^k \sqrt{\bar{\kappa}} \tag{3.17}$$

as $m_h \to +\infty$. Passing to the limit in (3.14), by (3.17), (3.15) we deduce that $\partial_\kappa^k \big\{ \vec{\omega}(\bar{\kappa}) \cdot \bar{c} + \bar{d} \sqrt{\bar{\kappa}} \big\} = 0$, $\forall k \in \mathbb{N}$. Therefore the analytic function $\kappa \mapsto \vec{\omega}(\kappa) \cdot \bar{c} + \bar{d} \sqrt{\kappa} = 0$ is identically zero. Since $(\bar{c}, \bar{d}) \neq 0$ this is in contradiction with Lemma 3.1.

Proof of (3.8). Notice that, for all $\kappa \in [\kappa_1, \kappa_2]$,

$$\begin{aligned} |\vec{\omega}(\kappa) \cdot \ell + \Omega_j(\kappa) - \Omega_{j'}(\kappa)| &\ge |\Omega_j(\kappa) - \Omega_{j'}(\kappa)| - |\vec{\omega}(\kappa)| |\ell| \\ &\overset{(3.11),(3.12)}{\ge} \sqrt{\kappa_1} |j^{\frac{3}{2}} - j'^{\frac{3}{2}}| - C - C|\ell| \ge \langle \ell \rangle \end{aligned}$$

provided $|j^{\frac{3}{2}} - j'^{\frac{3}{2}}| \ge C_1 \langle \ell \rangle$, for some $C_1 > 0$. Therefore in (3.8) we can restrict to the indices such that

$$|j^{\frac{3}{2}} - j'^{\frac{3}{2}}| < C_1 \langle \ell \rangle \,. \tag{3.18}$$

Moreover in (3.8) we can also assume that $j \neq j'$ otherwise (3.8) reduces to (3.6), which is already proved.

Now if, by contradiction, (3.8) is false, we deduce, arguing as in the previous cases, that for all $m \in \mathbb{N}$, there exist $\ell_m \in \mathbb{Z}^\nu$, $j_m, j'_m \in \mathbb{N}^+ \setminus \mathbb{S}^+$, $j_m \neq j'_m$, $\kappa_m \in [\kappa_1, \kappa_2]$, such that for all

$$k \in \mathbb{N} \,, \quad \forall m \ge k \,, \quad \Big| \partial_\kappa^k \Big\{ \vec{\omega}(\kappa_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle} + \frac{\Omega_{j_m}(\kappa_m)}{\langle \ell_m \rangle} - \frac{\Omega_{j'_m}(\kappa_m)}{\langle \ell_m \rangle} \Big\} \Big| < \frac{1}{1+m} \,. \tag{3.19}$$

As in the previous cases, since the sequences $(\kappa_m)_{m \in \mathbb{N}}$, $(\ell_m / \langle \ell_m \rangle)_{m \in \mathbb{N}}$ are bounded, there exists $m_h \to +\infty$ such that

$$\kappa_{m_h} \to \bar{\kappa} \in [\kappa_1, \kappa_2] \,, \quad \ell_{m_h} / \langle \ell_{m_h} \rangle \to \bar{c} \in \mathbb{R}^\nu \,. \tag{3.20}$$

We distinguish again two cases:

CASE 1 : (ℓ_{m_h}) IS BOUNDED. In this case, up to subsequence, $\ell_{m_h} \to \bar{\ell} \in \mathbb{Z}^\nu$. Using that

$$|j^{\frac{3}{2}} - j'^{\frac{3}{2}}| \geq |j - j'|(\sqrt{j} + \sqrt{j'}) \geq \sqrt{j} + \sqrt{j'}\,, \quad \forall j \neq j'\,,$$

by (3.18) we deduce that also j_{m_h}, j'_{m_h} are bounded sequences and therefore, up to subsequence,

$$j_{m_h} \to \bar{\jmath}\,, \qquad j'_{m_h} \to \bar{\jmath}'\,, \quad \bar{\jmath} \neq \bar{\jmath}'\,. \tag{3.21}$$

Hence passing to the limit in (3.19) for $m_h \to +\infty$, we deduce by (3.20), (3.21) that

$$\partial_\kappa^k \big\{ \vec{\omega}(\bar{\kappa}) \cdot \bar{c} + \Omega_{\bar{\jmath}}(\bar{\kappa})\langle \bar{\ell} \rangle^{-1} - \Omega_{\bar{\jmath}'}(\bar{\kappa})\langle \bar{\ell} \rangle^{-1} \big\} = 0\,, \quad \forall k \in \mathbb{N}\,.$$

Therefore the analytic function $\kappa \mapsto \vec{\omega}(\kappa) \cdot \bar{c} + \Omega_{\bar{\jmath}}(\kappa)\langle \bar{\ell} \rangle^{-1} - \Omega_{\bar{\jmath}'}(\kappa)\langle \bar{\ell} \rangle^{-1}$ is identically zero. This in contradiction with Lemma 3.1.

CASE 2 : (ℓ_{m_h}) IS UNBOUNDED. Up to subsequence $|\ell_{m_h}| \to +\infty$. In this case the constant $\bar{c} \neq 0$ in (3.20). Using (3.11)-(3.12), for all $k \in \mathbb{N}$,

$$\partial_\kappa^k \frac{\Omega_{j_{m_h}}(\kappa_{m_h}) - \Omega_{j'_{m_h}}(\kappa_{m_h})}{\langle \ell_{m_h} \rangle} = \partial_\kappa^k \sqrt{\kappa_{m_h}} \frac{j^{\frac{3}{2}} - j'^{\frac{3}{2}}}{\langle \ell_{m_h} \rangle} + \frac{1}{\sqrt{j_{m_h}}\langle \ell_{m_h} \rangle} \partial_\kappa^k \frac{c_{j_{m_h}}(\kappa_{m_h})}{\sqrt{\kappa_{m_h}}} - \frac{1}{\sqrt{j'_{m_h}}\langle \ell_{m_h} \rangle} \partial_\kappa^k \frac{c_{j'_{m_h}}(\kappa_{m_h})}{\sqrt{\kappa_{m_h}}}$$

and

$$\Big| \frac{1}{\sqrt{j_{m_h}}\langle \ell_{m_h} \rangle} \partial_\kappa^k \frac{c_{j_{m_h}}(\kappa_{m_h})}{\sqrt{\kappa_{m_h}}} - \frac{1}{\sqrt{j'_{m_h}}\langle \ell_{m_h} \rangle} \partial_\kappa^k \frac{c_{j'_{m_h}}(\kappa_{m_h})}{\sqrt{\kappa_{m_h}}} \Big| \leq \frac{C}{\langle \ell_{m_h} \rangle} \sup_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+, \kappa \in [\kappa_1, \kappa_2]} \Big| \partial_\kappa^k \frac{c_j(\kappa)}{\sqrt{\kappa}} \Big| \leq \frac{C'(k)}{\langle \ell_{m_h} \rangle} \to 0$$

as $m_h \to +\infty$. Moreover, by (3.18), up to subsequences, $|j_{m_h}^{\frac{3}{2}} - j'^{\frac{3}{2}}_{m_h}|\langle \ell_{m_h} \rangle^{-1} \to \bar{d} \in \mathbb{R}$. Therefore, for all $k \in \mathbb{N}$,

$$\partial_\kappa^k \frac{\Omega_{j_{m_h}}(\kappa_{m_h}) - \Omega_{j'_{m_h}}(\kappa_{m_h})}{\langle \ell_{m_h} \rangle} \to \bar{d}\, \partial_\kappa^k \sqrt{\bar{\kappa}}\,.$$

Passing to the limit in (3.19) for $m_h \to +\infty$ we deduce that $\partial_\kappa^k \big\{ \vec{\omega}(\bar{\kappa}) \cdot \bar{c} + \bar{d}\sqrt{\bar{\kappa}} \big\} = 0$, $\forall k \in \mathbb{N}$. In conclusion the analytic function $\kappa \mapsto \vec{\omega}(\kappa) \cdot \bar{c} + \bar{d}\sqrt{\kappa}$ is identically zero. Since $(\bar{c}, d) \neq 0$, this is a contradiction with Lemma 3.1.

PROOF OF (3.9). The proof is similar to the previous ones and we omit it. □

4 Nash-Moser theorem and measure estimates

Instead of working in a shrinking neighborhood of the origin, it is a convenient devise to rescale the variable $u \mapsto \varepsilon u$ with $u = O(1)$, writing (1.3)-(1.5) as

$$\partial_t u = J\Omega u + \varepsilon X_{P_\varepsilon}(u) \tag{4.1}$$

where $J\Omega$ is the linearized Hamiltonian vector field in (1.14) and

$$X_{P_\varepsilon}(u) := X_{P_\varepsilon}(\kappa, u) := \begin{pmatrix} \varepsilon^{-1}(G(\varepsilon\eta) - G(0))\psi \\ -\frac{1}{2}\psi_x^2 + \frac{1}{2}\frac{\big(G(\varepsilon\eta)\psi + \varepsilon\eta_x\psi_x\big)^2}{1 + (\varepsilon\eta_x)^2} + \varepsilon^{-1}\kappa\, \eta_{xx}\Big(\big(1 + (\varepsilon\eta_x)^2\big)^{-3/2} - 1\Big) \end{pmatrix}. \tag{4.2}$$

Note that the dependence of the vector field X_{P_ε} with respect to κ is linear. System (4.1) is the Hamiltonian system generated by the Hamiltonian

$$\mathcal{H}_\varepsilon(u) := \varepsilon^{-2}H(\varepsilon u) = H_L(u) + \varepsilon P_\varepsilon(u) \tag{4.3}$$

where H is the water-waves Hamiltonian (1.6), H_L is defined in (1.15) and

$$P_\varepsilon(u) := P_\varepsilon(\kappa, u) := \frac{\varepsilon^{-1}}{2}(\psi, \big(G(\varepsilon\eta) - G(0)\big)\psi)_{L^2(\mathbb{T}_x)} + \varepsilon^{-3}\kappa \int_{\mathbb{T}} \Big(\sqrt{1+(\varepsilon\eta_x)^2} - 1 - \frac{(\varepsilon\eta_x)^2}{2}\Big)\, dx\,. \quad (4.4)$$

We decompose the phase space

$$H^1_{0,\mathrm{even}} := \Big\{u := (\eta, \psi) \in H^1_0(\mathbb{T}_x) \times H^1_0(\mathbb{T}_x)\,, \quad u(x) = u(-x)\Big\} \quad (4.5)$$

as the direct sum of the symplectic subspaces

$$H^1_{0,\mathrm{even}} = H_{\mathbb{S}^+} \oplus H^\perp_{\mathbb{S}^+} \qquad \text{where} \qquad H_{\mathbb{S}^+} := \Big\{v := \sum_{j\in\mathbb{S}^+} \begin{pmatrix}\eta_j\\ \psi_j\end{pmatrix}\cos(jx)\Big\} \quad (4.6)$$

and $H^\perp_{\mathbb{S}^+}$ denotes the L^2-orthogonal.

We now introduce action-angle variables on the tangential sites by setting

$$\eta_j := \sqrt{\frac{2}{\pi}}\Lambda_j^{1/2}\sqrt{\xi_j + I_j}\cos(\theta_j), \quad \psi_j := \sqrt{\frac{2}{\pi}}\Lambda_j^{-1/2}\sqrt{\xi_j + I_j}\,\sin(\theta_j)\,, \quad \Lambda_j := \sqrt{j(1+\kappa j^2)^{-1}}\,, \ j \in \mathbb{S}^+\,, \quad (4.7)$$

where $\xi_j > 0$, $j = 1, \dots, \nu$, are positive constants, the variables $|I_j| \le \xi_j$, and we leave unchanged the normal component z. The symplectic 2-form in (1.7) then reads (for simplicity of notation we denote it in the same way)

$$\mathcal{W} := \Big(\sum\nolimits_{j\in\mathbb{S}^+} d\theta_j \wedge dI_j\Big) \oplus \mathcal{W}_{|H^\perp_{\mathbb{S}^+}} = d\Lambda \quad (4.8)$$

where Λ is the Liouville 1-form

$$\Lambda_{(\theta,I,z)}[\widehat\theta, \widehat I, \widehat z] := -\sum_{j\in\mathbb{S}^+} I_j\widehat\theta_j - \frac{1}{2}\big(Jz\,,\,\widehat z\big)_{L^2_x}\,. \quad (4.9)$$

Hence the Hamiltonian system (4.1) transforms into the new Hamiltonian system

$$\dot\theta = \partial_I H_\varepsilon(\theta, I, z)\,, \ \dot I = -\partial_\theta H_\varepsilon(\theta, I, z)\,, \quad z_t = J\nabla_z H_\varepsilon(\theta, I, z) \quad (4.10)$$

generated by the Hamiltonian

$$H_\varepsilon := \mathcal{H}_\varepsilon \circ A = \varepsilon^{-2} H \circ \varepsilon A \quad (4.11)$$

where

$$A(\theta, I, z) := v(\theta, I) + z := \sum_{j\in\mathbb{S}^+}\sqrt{\frac{2}{\pi}}\begin{pmatrix}\Lambda_j^{1/2}\sqrt{\xi_j + I_j}\,\cos(\theta_j)\\ \Lambda_j^{-1/2}\sqrt{\xi_j+I_j}\,\sin(\theta_j)\end{pmatrix}\cos(jx) + z\,. \quad (4.12)$$

We denote by

$$X_{H_\varepsilon} := (\partial_I H_\varepsilon, -\partial_\theta H_\varepsilon, J\nabla_z H_\varepsilon)$$

the Hamiltonian vector field in the variables $(\theta, I, z) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H^\perp_{\mathbb{S}^+}$. The involution ρ in (1.11) becomes

$$\tilde\rho : (\theta, I, z) \mapsto (-\theta, I, \rho z)\,. \quad (4.13)$$

By (1.6) and (4.11) the Hamiltonian H_ε reads (up to a constant)

$$H_\varepsilon = \mathcal{N} + \varepsilon P\,, \quad \mathcal{N} := H_L \circ A = \vec\omega(\kappa)\cdot I + \frac{1}{2}(z, \Omega z)_{L^2_x}\,, \quad P := P_\varepsilon \circ A\,, \quad (4.14)$$

where $\vec\omega(\kappa)$ is defined in (3.2) and Ω in (1.14). We look for an embedded invariant torus

$$i : \mathbb{T}^\nu \to \mathbb{T}^\nu \times \mathbb{R}^\nu \times H^\perp_{\mathbb{S}^+}, \quad \varphi \mapsto i(\varphi) := (\theta(\varphi), I(\varphi), z(\varphi)) \quad (4.15)$$

of the Hamiltonian vector field X_{H_ε} filled by quasi-periodic solutions with diophantine frequency $\omega \in \mathbb{R}^\nu$ (and which satisfies also first and second order Melnikov-non-resonance conditions as in (4.25)).

4.1 Nash-Moser Théoréme de conjugaison hypothétique

The Hamiltonian H_ε in (4.14) is a perturbation of the isochronous Hamiltonian $\mathcal{N}$. The expected quasi-periodic solutions of the Hamiltonian system (4.10) will have a shifted frequency which depends on the nonlinear term P. In view of that we introduce the family of Hamiltonians

$$H_\alpha := \mathcal{N}_\alpha + \varepsilon P\,, \quad \mathcal{N}_\alpha := \alpha\cdot I + \frac{1}{2}(z,\Omega z)_{L^2_x}\,, \quad \alpha\in\mathbb{R}^\nu\,, \tag{4.16}$$

which depend on the constant vector $\alpha\in\mathbb{R}^\nu$. For the value $\alpha=\vec{\omega}(\kappa)$ we have $H_\alpha = H_\varepsilon$. Then we look for a zero (i,α) of the nonlinear operator

$$\begin{aligned}\mathcal{F}(i,\alpha) &:= \mathcal{F}(i,\alpha,\omega,\kappa,\varepsilon) := \omega\cdot\partial_\varphi i(\varphi) - X_{H_\alpha} = \omega\cdot\partial_\varphi i(\varphi) - (X_{\mathcal{N}_\alpha}+\varepsilon X_P)(i(\varphi)) \qquad (4.17)\\ &:= \begin{pmatrix} \omega\cdot\partial_\varphi\theta(\varphi) - \alpha - \varepsilon\partial_I P(i(\varphi)) \\ \omega\cdot\partial_\varphi I(\varphi) + \varepsilon\partial_\theta P(i(\varphi)) \\ \omega\cdot\partial_\varphi z(\varphi) - J(\Omega z + \varepsilon\nabla_z P(i(\varphi)))\end{pmatrix}\end{aligned}$$

for some diophantine vector $\omega\in\mathbb{R}^\nu$. Thus $\varphi\mapsto i(\varphi)$ is an embedded torus, invariant for the Hamiltonian vector field X_{H_α}, filled by quasi-periodic solutions with frequency ω.

Each Hamiltonian H_α in (4.16) is reversible, i.e. $H_\alpha\circ\tilde{\rho} = H_\alpha$ where the involution $\tilde{\rho}$ is defined in (4.13). We look for reversible solutions of $\mathcal{F}(i,\alpha)=0$, namely satisfying $\tilde{\varrho} i(\varphi) = i(-\varphi)$ (see (4.13)), i.e.

$$\theta(-\varphi) = -\theta(\varphi)\,, \quad I(-\varphi) = I(\varphi)\,, \quad z(-\varphi) = (\varrho z)(\varphi)\,. \tag{4.18}$$

The Sobolev norm of the periodic component of the embedded torus

$$\mathfrak{I}(\varphi) := i(\varphi) - (\varphi,0,0) := (\Theta(\varphi), I(\varphi), z(\varphi))\,, \quad \Theta(\varphi) := \theta(\varphi)-\varphi\,, \tag{4.19}$$

is

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

where $\|z\|_s := \|z\|_{H^s_{\varphi,x}} = \|\eta\|_s + \|\psi\|_s$, see (1.19).

For the next theorem, we recall that k_0 is the index of non-degeneracy provided by Proposition 3.2 and it depends only on the linear unperturbed frequencies. Therefore it is considered as an absolute constant and we will often omit to write explicitly the dependence of the constants with respect to k_0. We look for quasi periodic solutions with frequency ω belonging to a δ-neighborhood (independent of ε)

$$\Omega := \Big\{\omega\in\mathbb{R}^\nu \,:\, \mathrm{dist}\big(\omega,\vec{\omega}[\kappa_1,\kappa_2]\big) < \delta\,,\ \delta>0\Big\}\,. \tag{4.20}$$

of the unperturbed linear frequencies $\vec{\omega}[\kappa_1,\kappa_2]$ defined in (3.2).

Theorem 4.1. (Nash-Moser) *Fix finitely many tangential sites $\mathbb{S}^+\subset\mathbb{N}^+$ and let $\nu := |\mathbb{S}^+|$. Let $\tau\geq 1$. There exist constants $\varepsilon_0>0$, $a_0 := a_0(\nu,\tau,k_0)>0$ and $k_1 := k_1(\nu,k_0,\tau)>0$ such that, for all $\gamma=\varepsilon^a$, $0<a<a_0$, $\varepsilon\in(0,\varepsilon_0)$, there exist a k_0-times differentiable function*

$$\alpha_\infty : \Omega\times[\kappa_1,\kappa_2]\mapsto\mathbb{R}^\nu\,, \quad \alpha_\infty(\omega,\kappa) = \omega + r_\varepsilon(\omega,\kappa)\,, \quad \textit{with} \quad |r_\varepsilon|^{k_0,\gamma} \leq C\varepsilon\gamma^{-(1+k_1)}\,, \tag{4.21}$$

a family of embedded tori i_∞ defined for all $\omega\in\Omega$ and $\kappa\in[\kappa_1,\kappa_2]$ satisfying the reversibility property (4.18) *and*

$$\|i_\infty(\varphi) - (\varphi,0,0)\|^{k_0,\gamma}_{s_0} \leq C\varepsilon\gamma^{-(1+k_1)}\,, \tag{4.22}$$

a sequence of k_0-times differentiable functions $\mu_j^\infty : \Omega\times[\kappa_1,\kappa_2]\to\mathbb{R}$, $j\in\mathbb{N}^+\setminus\mathbb{S}^+$, of the form

$$\mu_j^\infty(\omega,\kappa) = \mathtt{m}_3^\infty(\omega,\kappa) j^{\frac12}(1+\kappa j^2)^{\frac12} + \mathtt{m}_1^\infty(\omega,\kappa) j^{\frac12} + r_j^\infty(\omega,\kappa) \tag{4.23}$$

(defined in (8.40)*) satisfying*

$$|\mathtt{m}_3^\infty - 1|^{k_0,\gamma} + |\mathtt{m}_1^\infty|^{k_0,\gamma} \leq C\varepsilon\,, \qquad \sup_{j\in\mathbb{S}^c}|r_j^\infty|^{k_0,\gamma} \leq C\varepsilon\gamma^{-k_1}\,, \tag{4.24}$$

such that for all (ω,κ) in the Cantor like set

$$
\begin{aligned}
\mathcal{C}^\gamma_\infty := \Big\{ & (\omega,\kappa) \in \Omega \times [\kappa_1,\kappa_2] \,:\, |\omega\cdot\ell| \geq \gamma\langle\ell\rangle^{-\tau},\ \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \\
& |\omega\cdot\ell + \mu_j^\infty(\omega,\kappa)| \geq 4\gamma j^{\frac{3}{2}}\langle\ell\rangle^{-\tau},\ \forall\ell\in\mathbb{Z}^\nu,\ j\in\mathbb{N}^+\setminus\mathbb{S}^+ \ (1\text{-}Melnikov\ conditions) \\
& |\omega\cdot\ell + \mu_j^\infty(\omega,\kappa) - \varsigma\mu_{j'}^\infty(\omega,\kappa)| \geq \frac{4\gamma|j^{\frac{3}{2}} - \varsigma j'^{\frac{3}{2}}|}{\langle\ell\rangle^\tau}, \forall\ell\in\mathbb{Z}^\nu, j,j'\in\mathbb{N}^+\setminus\mathbb{S}^+, \varsigma=\pm1\,, (2\text{-}Melnikov)\Big\}
\end{aligned}
\tag{4.25}
$$

the function $i_\infty(\varphi) := i_\infty(\omega,\kappa,\varepsilon)(\varphi)$ is a solution of $\mathcal{F}(i_\infty,\alpha_\infty(\omega,\kappa),\omega,\kappa,\varepsilon) = 0$. As a consequence the embedded torus $\varphi \mapsto i_\infty(\varphi)$ is invariant for the Hamiltonian vector field $X_{H_{\alpha_\infty(\omega,\kappa)}}$ and it is filled by quasi-periodic solutions with frequency ω.

Note that the Cantor-like set $\mathcal{C}^\gamma_\infty$ in (4.25) for which a solution exists is defined only in terms of the "final" solution i_∞ and the "final" normal perturbed frequencies μ_j^∞, $j\in\mathbb{N}^+\setminus\mathbb{S}^+$. In Theorem 4.1 we are not concerned about the measure of $\mathcal{C}^\gamma_\infty$, in particular in investigating if it is not empty (note that α_∞, i_∞ and each μ_j^∞ are anyway defined for all $(\omega,\kappa)\in\Omega\times[\kappa_1,\kappa_2]$).

4.2 Measure estimates

The aim is now to deduce Theorem 1.1 from Theorem 4.1.

By (4.21), for any $\kappa\in[\kappa_1,\kappa_2]$, the function $\alpha_\infty(\cdot,\kappa)$ from Ω into the image $\alpha_\infty(\Omega\times\{\kappa\})$ is invertible:

$$
\beta = \alpha_\infty(\omega,\kappa) = \omega + r_\varepsilon(\omega,\kappa) \quad \Longleftrightarrow \quad \omega = \alpha_\infty^{-1}(\beta,\kappa) = \beta + \tilde r_\varepsilon(\beta,\kappa) \quad \text{with} \quad |\tilde r_\varepsilon|^{k_0,\gamma} \leq C\varepsilon\gamma^{-(1+k_1)}\,. \tag{4.26}
$$

Proof of (4.26).The inverse map $\beta\mapsto\alpha_\infty^{-1}(\beta,\kappa) = \beta+\tilde r_\varepsilon(\beta,\kappa)$ satisfies the identities $\tilde r_\varepsilon(\beta,\kappa) + r_\varepsilon(\beta + \tilde r_\varepsilon(\beta,\kappa),\kappa) = 0$. By the implicit function theorem $\tilde r_\varepsilon$ is $\mathcal{C}^1$ with respect to (β,κ) and it satisfies the identities

$$
\begin{aligned}
D_\beta\tilde r_\varepsilon(\beta,\kappa) &= -\big(\mathrm{Id} + D_\omega r_\varepsilon(\beta+\tilde r_\varepsilon(\beta,\kappa),\kappa)\big)^{-1} D_\omega r_\varepsilon(\beta+\tilde r_\varepsilon(\beta,\kappa),\kappa)\,, \\
\partial_\kappa\tilde r_\varepsilon(\beta,\kappa) &= -\big(\mathrm{Id} + D_\omega r_\varepsilon(\beta+\tilde r_\varepsilon(\beta,\kappa),\kappa)\big)^{-1} \partial_\kappa r_\varepsilon(\beta+\tilde r_\varepsilon(\beta,\kappa),\kappa)
\end{aligned}
$$

where D_ω, D_β denote the Fréchet derivatives with respect to the variables ω and β. Arguing by induction on $|k|\leq k_0$, $\tilde r_\varepsilon$ is k_0-times differentiable and the estimate (4.26) follows as the estimate (2.97).
Then, for any $\beta\in\alpha_\infty(\mathcal{C}^\gamma_\infty)$, Theorem 4.1 proves the existence of an embedded invariant torus filled by quasi-periodic solutions with diophantine frequency $\omega = \alpha_\infty^{-1}(\beta,\kappa)$ for the Hamiltonian

$$
H_\beta = \beta\cdot I + \frac{1}{2}(z,\Omega z)_{L^2_x} + \varepsilon P\,.
$$

Consider the curve of the unperturbed linear frequencies

$$
[\kappa_1,\kappa_2] \ni \kappa \mapsto \vec\omega(\kappa) := \big(\sqrt{j(1+\kappa j^2)}\big)_{j\in\mathbb{S}^+} \in \mathbb{R}^\nu\,.
$$

We now prove that for "most" $\kappa\in[\kappa_1,\kappa_2]$ the vector $\beta = \vec\omega(\kappa)\in\alpha_\infty(\mathcal{C}^\gamma_\infty)$, see Theorem 4.2. Hence, for such values of κ we have found an embedded invariant torus for the Hamiltonian H_ε in (4.14), filled by quasi-periodic solutions with diophantine frequency $\omega = \alpha_\infty^{-1}(\vec\omega(\kappa),\kappa)$. This implies Theorem 1.1.

Theorem 4.2. (Measure estimates) *Let*

$$
\gamma = \varepsilon^a\,, \quad 0 < a < \min\{a_0, 1/(1+k_0+k_1)\}\,, \quad \tau > k_0(\nu+4)\,. \tag{4.27}
$$

Then the measure of the set

$$
\mathcal{G}_\varepsilon := \big\{\kappa\in[\kappa_1,\kappa_2]\ :\vec\omega(\kappa)\in\alpha_\infty(\mathcal{C}^\gamma_\infty)\big\} = \big\{\kappa\in[\kappa_1,\kappa_2]\ : \alpha_\infty^{-1}(\vec\omega(\kappa),\kappa)\in\mathcal{C}^\gamma_\infty\big\} \tag{4.28}
$$

satisfies $|\mathcal{G}_\varepsilon| \geq (\kappa_2-\kappa_1) - C\varepsilon^{a/k_0}$ as $\varepsilon\to 0$.

Theorems 4.1-4.2 prove Theorem 1.1 with the Cantor-like set $\mathcal{G} := \mathcal{G}_\varepsilon$ defined in (4.28) and frequency vector $\tilde{\omega} = \omega_\varepsilon(\kappa)$ defined in (4.29) below.

The rest of this section is devoted to the proof of Theorem 4.2. By (4.26) the vector

$$\omega_\varepsilon(\kappa) := \alpha_\infty^{-1}(\vec{\omega}(\kappa), \kappa) = \vec{\omega}(\kappa) + \mathtt{r}_\varepsilon(\kappa), \quad \mathtt{r}_\varepsilon(\kappa) := \tilde{r}_\varepsilon(\vec{\omega}(\kappa), \kappa), \tag{4.29}$$

satisfies

$$|\partial_\kappa^k \mathtt{r}_\varepsilon(\kappa)| \le C\varepsilon\gamma^{-(1+k_1+k)}, \ \forall 0 \le k \le k_0. \tag{4.30}$$

We also denote, with a small abuse of notations,

$$\mu_j^\infty(\kappa) := \mu_j^\infty(\omega_\varepsilon(\kappa), \kappa) := \mathtt{m}_3^\infty(\kappa) j^{\frac12}(1+\kappa j^2)^{\frac12} + \mathtt{m}_1^\infty(\kappa) j^{\frac12} + r_j^\infty(\kappa), \quad \forall j \in \mathbb{N}^+ \setminus \mathbb{S}^+, \tag{4.31}$$

where

$$\mathtt{m}_3^\infty(\kappa) := \mathtt{m}_3^\infty(\omega_\varepsilon(\kappa), \kappa), \quad \mathtt{m}_1^\infty(\kappa) := \mathtt{m}_1^\infty(\omega_\varepsilon(\kappa), \kappa), \quad r_j^\infty(\kappa) := r_j^\infty(\omega_\varepsilon(\kappa), \kappa). \tag{4.32}$$

By (4.24), (4.32) and (4.29), using that $\varepsilon\gamma^{-(1+k_1+k_0)} \le 1$ (that by (4.27) is satisfied for ε small), we get

$$|\partial_\kappa^k[\mathtt{m}_3^\infty(\kappa) - 1]|, |\partial_\kappa^k \mathtt{m}_1^\infty(\kappa)| \le C\varepsilon\gamma^{-k}, \quad \sup_{j\in\mathbb{S}^c}|\partial_\kappa^k r_j^\infty(\kappa)| \le C\varepsilon\gamma^{-(k+k_1)}, \quad \forall 0 \le k \le k_0. \tag{4.33}$$

By (4.25), (4.29), (4.31) the Cantor set $\mathcal{G}_\varepsilon$ in (4.28) writes

$$\begin{aligned}\mathcal{G}_\varepsilon = \Big\{ &\kappa \in [\kappa_1, \kappa_2] : |\omega_\varepsilon(\kappa)\cdot\ell| \ge \gamma\langle\ell\rangle^{-\tau}, \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \\ &|\omega_\varepsilon(\kappa)\cdot\ell + \mu_j^\infty(\kappa)| \ge 4\gamma j^{\frac32}\langle\ell\rangle^{-\tau}, \forall \ell \in \mathbb{Z}^\nu, j \in \mathbb{N}^+\setminus\mathbb{S}^+, \\ &|\omega_\varepsilon(\kappa)\cdot\ell + \mu_j^\infty(\kappa) - \varsigma\mu_{j'}^\infty(\kappa)| \ge 4\gamma|j^{\frac32} - \varsigma j'^{\frac32}|\langle\ell\rangle^{-\tau}, \forall \ell \in \mathbb{Z}^\nu, j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+, \varsigma \in \{+,-\}\Big\}.\end{aligned}$$

We estimate the measure of the complementary set

$$\mathcal{G}_\varepsilon^c := [\kappa_1, \kappa_2] \setminus \mathcal{G}_\varepsilon := \Big(\bigcup_\ell R_\ell^{(0)}\Big) \bigcup \Big(\bigcup_{\ell,j} R_{\ell,j}^{(I)}\Big) \bigcup \Big(\bigcup_{\ell,j,j'} R_{\ell jj'}^{(II)}\Big) \bigcup \Big(\bigcup_{\ell,j,j'} Q_{\ell jj'}^{(II)}\Big) \tag{4.34}$$

where the "resonant sets" are

$$R_\ell^{(0)} := \big\{\kappa \in [\kappa_1, \kappa_2] : |\omega_\varepsilon(\kappa)\cdot\ell| < 4\gamma\langle\ell\rangle^{-\tau}\big\} \tag{4.35}$$

$$R_{\ell j}^{(I)} := \big\{\kappa \in [\kappa_1, \kappa_2] : |\omega_\varepsilon(\kappa)\cdot\ell + \mu_j^\infty(\kappa)| < 4\gamma j^{\frac32}\langle\ell\rangle^{-\tau}\big\} \tag{4.36}$$

$$R_{\ell jj'}^{(II)} := \big\{\kappa \in [\kappa_1, \kappa_2] : |\omega_\varepsilon(\kappa)\cdot\ell + \mu_j^\infty(\kappa) - \mu_{j'}^\infty(\kappa)| < 4\gamma|j^{\frac32} - j'^{\frac32}|\langle\ell\rangle^{-\tau}\big\} \tag{4.37}$$

$$Q_{\ell jj'}^{(II)} := \big\{\kappa \in [\kappa_1, \kappa_2] : |\omega_\varepsilon(\kappa)\cdot\ell + \mu_j^\infty(\kappa) + \mu_{j'}^\infty(\kappa)| < 4\gamma|j^{\frac32} + j'^{\frac32}|\langle\ell\rangle^{-\tau}\big\}. \tag{4.38}$$

Lemma 4.3. *If $R_{\ell j}^{(I)} \ne \emptyset$ then $j^{\frac32} \le C\langle\ell\rangle$. If $R_{\ell jj'}^{(II)} \ne \emptyset$ then $|j^{\frac32} - j'^{\frac32}| \le C\langle\ell\rangle$. If $Q_{\ell jj'}^{(II)} \ne \emptyset$ then $j^{\frac32} + j'^{\frac32} \le C\langle\ell\rangle$.*

Proof. We prove the lemma for $R_{\ell jj'}^{(II)}$. The other cases follow similarly. If $\kappa \in R_{\ell jj'}^{(II)}$ then

$$|\mu_j^\infty(\kappa) - \mu_{j'}^\infty(\kappa)| < 4\gamma|j^{\frac32} - j'^{\frac32}|\langle\ell\rangle^{-\tau} + |\omega_\varepsilon(\kappa)||\ell| \le 4\gamma|j^{\frac32} - j'^{\frac32}| + C|\ell|. \tag{4.39}$$

Moreover (4.31) and (4.33) imply

$$\begin{aligned}|\mu_j^\infty - \mu_{j'}^\infty| &\ge |\mathtt{m}_3^\infty(\kappa)||j^{\frac12}(1+\kappa j^2)^{\frac12} - j'^{\frac12}(1+\kappa j'^2)^{\frac12}| - |\mathtt{m}_1^\infty(\kappa)||j^{\frac12} - j'^{\frac12}| - 2\sup_{j\in\mathbb{S}^c}|r_j^\infty(\kappa)| \\ &\ge C_1|j^{\frac32} - j'^{\frac32}| - C\varepsilon|j^{\frac12} - j'^{\frac12}| - C\varepsilon\gamma^{-k_1} \ge C_1|j^{\frac32} - j'^{\frac32}|/2\end{aligned} \tag{4.40}$$

for $2C\varepsilon\gamma^{-k_1} \le C_1/2$, which is fulfilled taking ε small enough by (4.27). The lemma follows by (4.39), (4.40), for $C_1/4 \ge 4\gamma$. □

The perturbed frequencies satisfy estimates similar to (3.6)-(3.9) in Proposition 3.2.

Lemma 4.4. *For ε small enough, for all $\kappa \in [\kappa_1, \kappa_2]$,*

$$\max_{k\le k_0}|\partial_\kappa^k\{\omega_\varepsilon(\kappa)\cdot\ell\}| \ge \rho_0\langle\ell\rangle/2\,, \quad \forall \ell\in\mathbb{Z}^\nu\setminus\{0\}, \tag{4.41}$$

$$\max_{k\le k_0}|\partial_\kappa^k\{\omega_\varepsilon(\kappa)\cdot\ell+\mu_j^\infty(\kappa)\}| \ge \rho_0\langle\ell\rangle/2\,, \quad \forall \ell\in\mathbb{Z}^\nu,\, j\in\mathbb{N}^+\setminus\mathbb{S}^+, \tag{4.42}$$

$$\max_{k\le k_0}|\partial_\kappa^k\{\omega_\varepsilon(\kappa)\cdot\ell+\mu_j^\infty(\kappa)-\mu_{j'}^\infty(\kappa)\}| \ge \rho_0\langle\ell\rangle/2\,, \quad \forall (\ell,j,j')\neq(0,j,j),\, j,j'\in\mathbb{N}^+\setminus\mathbb{S}^+, \tag{4.43}$$

$$\max_{k\le k_0}|\partial_\kappa^k\{\omega_\varepsilon(\kappa)\cdot\ell+\mu_j^\infty(\kappa)+\mu_{j'}^\infty(\kappa)\}| \ge \rho_0\langle\ell\rangle/2\,, \quad \forall \ell\in\mathbb{Z}^\nu, j,j'\in\mathbb{N}^+\setminus\mathbb{S}^+\,. \tag{4.44}$$

Proof. We prove (4.43). The other estimates follow analogously. First of all, by Lemma 4.3 we may restrict to the set of indices satisfying

$$|j^{\frac32}-j'^{\frac32}| \le C\langle\ell\rangle\,. \tag{4.45}$$

Split $\mu_j^\infty(\kappa)=\Omega_j(\kappa)+(\mu_j^\infty-\Omega_j)(\kappa)$ where $\Omega_j(\kappa):=j^{\frac12}(1+\kappa j^2)^{\frac12}$. A direct calculation shows that

$$|\partial_\kappa^k\{\Omega_j(\kappa)-\Omega_{j'}(\kappa)\}| \le C_k|j^{\frac32}-j'^{\frac32}|\,, \quad \forall\, k\ge 0\,. \tag{4.46}$$

Then all $0\le k\le k_0$ one has

$$\begin{aligned}|\partial_\kappa^k\big\{(\mu_j^\infty-\mu_{j'}^\infty)(\kappa)-(\Omega_j-\Omega_{j'})(\kappa)\big\}| &\le |\partial_\kappa^k\big\{(\mathtt{m}_3^\infty(\kappa)-1)(\Omega_j(\kappa)-\Omega_{j'}(\kappa)\big)| \\ &\quad + |\partial_\kappa^k\mathtt{m}_1^\infty(\kappa)||j^{\frac12}-j'^{\frac12}| + 2\sup_{j\in\mathbb{N}^+\setminus\mathbb{S}^+}|\partial_\kappa^k r_j^\infty(\kappa)| \\ &\overset{(4.46),(4.33)}{\le} C\varepsilon\gamma^{-(k+k_1)}|j^{\frac32}-j'^{\frac32}|\,.\end{aligned} \tag{4.47}$$

By (4.29), (4.30) and (4.47) we get

$$\begin{aligned}\max_{k\le k_0}|\partial_\kappa^k\{\omega_\varepsilon(\kappa)\cdot\ell+\mu_j^\infty(\kappa)-\mu_{j'}^\infty(\kappa)\}| &\ge \max_{k\le k_0}|\partial_\kappa^k\{\vec\omega(\kappa)\cdot\ell+\Omega_j(\kappa)-\Omega_{j'}(\kappa)\}| - C\varepsilon\gamma^{-(1+k_0+k_1)}|\ell| \\ &\quad - C\varepsilon\gamma^{-(k_0+k_1)}|j^{\frac32}-j'^{\frac32}| \\ &\overset{(4.45)}{\ge} \max_{k\le k_0}|\partial_\kappa^k\{\vec\omega(\kappa)\cdot\ell+\Omega_j(\kappa)-\Omega_{j'}(\kappa)\}| - C\varepsilon\gamma^{-(1+k_0+k_1)}\langle\ell\rangle \\ &\overset{(3.8)}{\ge} \rho_0\langle\ell\rangle - C\varepsilon\gamma^{-(1+k_0+k_1)}\langle\ell\rangle \ge \rho_0\langle\ell\rangle/2\end{aligned}$$

provided $\varepsilon\gamma^{-(1+k_0+k_1)}\le\rho_0/(2C)$, that, by (4.27), is satisfied for ε small. ☐

Lemma 4.5 (Estimates of the resonant sets). *The measures of the sets in* (4.35)-(4.38) *satisfy*

$$\begin{aligned}&|R_\ell^{(0)}| \lessdot \big(\gamma\langle\ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0}}\,, |R_{\ell j}^{(I)}| \lessdot \big(\gamma j^{\frac32}\langle\ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0}}\,, \\ &|R_{\ell jj'}^{(II)}| \lessdot \big(\gamma|j^{\frac32}-j'^{\frac32}|\langle\ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0}},\ |Q_{\ell jj'}^{(II)}| \lessdot \big(\gamma|j^{\frac32}+j'^{\frac32}|\langle\ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0}}\,.\end{aligned}$$

Proof. We prove the estimate of $R_{\ell jj'}^{(II)}$. The other cases are simpler. We write

$$R_{\ell jj'}^{(II)} = \big\{\kappa\in[\kappa_1,\kappa_2] : |g_{\ell jj'}(\kappa)| < 4\gamma|j^{\frac32}-j'^{\frac32}|\langle\ell\rangle^{-(\tau+1)}\big\}$$

where $g_{\ell jj'}(\kappa) := (\omega_\varepsilon(\kappa)\cdot\ell+\mu_j^\infty(\kappa)-\mu_{j'}^\infty(\kappa))\langle\ell\rangle^{-1}$. We apply Theorem 17.1 in [41]. We estimate the measure of $R_{\ell jj'}^{(II)}$ only if $4\gamma|j^{\frac32}-j'^{\frac32}|\langle\ell\rangle^{-(\tau+1)}\le\frac{\rho_0}{4(1+k_0)}$. Otherwise, for γ small enough, the set $R_{\ell jj'}^{(II)}=\emptyset$ is empty. By (4.43) we derive that

$$\max_{k\le k_0}|\partial_\kappa^k g_{\ell jj'}(\kappa)| \ge \rho_0/2\,, \quad \forall\kappa\in[\kappa_1,\kappa_2]\,.$$

In addition, (4.29)-(4.32) and Lemma 4.3 imply that $\max_{k\le k_0}|\partial_\kappa^k g_{\ell jj'}(\kappa)|\le C_1$, $\forall\kappa\in[\kappa_1,\kappa_2]$, provided $\varepsilon\gamma^{-(1+k_0+k_1)}$ is small enough. By Theorem 17.1 in [41] the Lemma follows. ☐

Proof of Theorem 4.2 completed. The measure of the set $\mathcal{G}_\varepsilon^c$ in (4.34) is estimated by

$$
\begin{aligned}
|\mathcal{G}_\varepsilon^c| & \leq \sum\nolimits_\ell |R_\ell^{(0)}| + \sum\nolimits_{\ell,j} |R_{\ell j}^{(I)}| + \sum\nolimits_{\ell,j,j'} |R_{\ell jj'}^{(II)}| + \sum\nolimits_{\ell,j,j'} |Q_{\ell jj'}^{(II)}| \\
& \overset{Lemma\, 4.3}{\leq} \sum\nolimits_\ell |R_\ell^{(0)}| + \sum\nolimits_{j \leq C\langle \ell\rangle^{2/3}} |R_{\ell j}^{(I)}| + \sum\nolimits_{j,j' \leq C\langle \ell\rangle^{2}} |R_{\ell jj'}^{(II)}| + \sum\nolimits_{j,j' \leq C\langle \ell\rangle^{2/3}} |Q_{\ell jj'}^{(II)}| \\
& \overset{Lemma\, 4.5}{\lessdot} \sum_\ell \big(\gamma \langle \ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0}} + \sum_{j \leq C\langle \ell\rangle^{2/3}} \big(\gamma j^{\frac{3}{2}} \langle \ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0}} + \sum_{j,j' \leq C\langle \ell\rangle^{2}} \big(\gamma |j^{\frac{3}{2}} - j'^{\frac{3}{2}}| \langle \ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0}} \\
& \quad + \sum_{j,j' \leq C\langle \ell\rangle^{2/3}} \big(\gamma |j^{\frac{3}{2}} + j'^{\frac{3}{2}}| \langle \ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0}} \overset{Lemma\, 4.3}{\leq} C\gamma^{\frac{1}{k_0}} \sum_\ell \langle \ell\rangle^{4-\frac{\tau}{k_0}} \overset{(4.27)}{\leq} C' \varepsilon^{\frac{a}{k_0}} .
\end{aligned}
$$

Hence $|\mathcal{G}_\varepsilon| \geq \kappa_2 - \kappa_1 - C'\varepsilon^{a/k_0}$ and the proof of Theorem 4.2 is concluded.

5 Approximate inverse

5.1 Estimates on the perturbation P

We prove tame estimates for the composition operator induced by the Hamiltonian vector field $X_P = (\partial_I P, -\partial_\theta P, J\nabla_z P)$ in (4.17).

We first estimate the composition operator induced by $v(\theta, y)$ defined in (4.12). Since the functions $I_j \mapsto \sqrt{\xi_j + I_j}$, $\theta \mapsto \cos(\theta)$, $\theta \mapsto \sin(\theta)$ are analytic for $|I| \leq r$ small, the composition Lemma 2.22 implies that, for all $\Theta, y \in H^s(\mathbb{T}^\nu, \mathbb{R}^\nu)$, $\|\Theta\|_{s_0}, \|y\|_{s_0} \leq r$, setting $\theta(\varphi) := \varphi + \Theta(\varphi)$,

$$
\|\partial_\theta^\alpha \partial_I^\beta v(\theta(\cdot), I(\cdot))\|_s^{k_0,\gamma} \leq_s 1 + \|\mathfrak{I}\|_s^{k_0,\gamma} , \quad \forall \alpha, \beta \in \mathbb{N}^\nu , \ \ |\alpha| + |\beta| \leq 3 . \tag{5.1}
$$

Lemma 5.1. *Let $\mathfrak{I}(\varphi)$ in (4.19) satisfy $\|\mathfrak{I}\|_{2s_0+2k_0+5}^{k_0,\gamma} \leq 1$. Then the following estimates hold:*

$$
\|X_P(i)\|_s^{k_0,\gamma} \leq_s 1 + \|\mathfrak{I}\|_{s+s_0+2k_0+3}^{k_0,\gamma} , \tag{5.2}
$$

and for all $\widehat{\imath} := (\widehat{\theta}, \widehat{I}, \widehat{z})$

$$
\|d_i X_P(i)[\widehat{\imath}]\|_s^{k_0,\gamma} \leq_s \|\widehat{\imath}\|_{s+2}^{k_0,\gamma} + \|\mathfrak{I}\|_{s+s_0+2k_0+4}^{k_0,\gamma} \|\widehat{\imath}\|_{s_0+2}^{k_0,\gamma} , \tag{5.3}
$$

$$
\|d_i^2 X_P(i)[\widehat{\imath}, \widehat{\imath}]\|_s^{k_0,\gamma} \leq_s \|\widehat{\imath}\|_{s+2}^{k_0,\gamma} \|\widehat{\imath}\|_{s_0+2}^{k_0,\gamma} + \|\mathfrak{I}\|_{s+s_0+2k_0+5}^{k_0,\gamma} (\|\widehat{\imath}\|_{s_0+2}^{k_0,\gamma})^2 . \tag{5.4}
$$

Proof. By the definition (4.14), $P = P_\varepsilon \circ A$, where A is defined in (4.12) and P_ε is defined in (4.4). Hence

$$
X_P = \Big([\partial_I v(\theta, I)]^T \nabla P_\varepsilon(A(\theta, I, z)) \, , \, -[\partial_\theta v(\theta, I)]^T \nabla P_\varepsilon(A(\theta, I, z)) \, , \, \Pi_{\mathbb{S}_+}^\perp J \nabla P_\varepsilon(A(\theta, I, z)) \Big) \tag{5.5}
$$

where $\Pi_{\mathbb{S}_+^\perp}$ is the L^2-projector on the space $H_{\mathbb{S}_+}^\perp$ defined in (4.6). Now $\nabla P_\varepsilon = -J X_{P_\varepsilon}$ (see (4.1)) where X_{P_ε} is the explicit Hamiltonian vector field in (4.2). The smallness condition of Lemma 2.32 is fulfilled because $\|\eta\|_{2s_0+2k_0+5}^{k_0,\gamma} \leq \varepsilon \|A(\theta(\cdot), I(\cdot), z(\cdot, \cdot))\|_{2s_0+2k_0+5}^{k_0,\gamma} \leq C(s_0)\varepsilon(1 + \|\mathfrak{I}\|_{2s_0+2k_0+5}^{k_0,\gamma}) \leq C_1(s_0)\varepsilon \leq \delta(s_0, k_0)$ for ε small. Thus by the tame estimate (2.132) for the Dirichlet Neumann operator, the interpolation inequality (2.72), and (5.1), we get

$$
\|\nabla P_\varepsilon(A(\theta(\cdot), I(\cdot), z(\cdot, \cdot)))\|_s^{k_0,\gamma} \leq_s \|A(\theta(\cdot), I(\cdot), z(\cdot, \cdot))\|_{s+s_0+2k_0+3}^{k_0,\gamma} \leq_s 1 + \|\mathfrak{I}\|_{s+s_0+2k_0+3}^{k_0,\gamma} .
$$

Hence (5.2) follows by (5.5), interpolation and (5.1).

The estimates (5.3), (5.4) for $d_i X_P$ and $d_i^2 X_P$ follow by differentiating the expression of X_P in (5.5) and applying the estimates (2.133), (2.134) on the Dirichlet Neumann operator, the estimate (5.1) on $v(\theta, y)$ and using the interpolation inequality (2.72). $\square$

5.2 Almost approximate inverse

In order to implement a convergent Nash-Moser scheme that leads to a solution of $\mathcal{F}(i, \alpha) = 0$ (the operator $\mathcal{F}(i, \alpha)$ is defined in (4.17)) we construct an *almost-approximate right inverse* of the linearized operator

$$d_{i,\alpha}\mathcal{F}(i_0, \alpha_0)[\widehat{\imath}, \widehat{\alpha}] = \omega \cdot \partial_\varphi \widehat{\imath} - d_i X_{H_\alpha}(i_0(\varphi))[\widehat{\imath}] - (\widehat{\alpha}, 0, 0)$$

at an approximate torus $i_0(\varphi) = (\theta_0(\varphi), I_0(\varphi), z_0(\varphi))$ (recall (4.15)), at a given value of α_0, see Theorem 5.10. Note that $d_{i,\alpha}\mathcal{F}(i_0, \alpha_0) = d_{i,\alpha}\mathcal{F}(i_0)$ is independent of α_0, see (4.17) and recall that the perturbation P in (4.14) does not depend on α (it depends on κ).

We implement the general strategy proposed in [15] and [8]. An invariant torus i_0 for the Hamiltonian vector field X_{H_α} with diophantine flow (i.e. ω satisfies (1.32)) is isotropic (see e.g. Lemma 1 in [15]), namely the pull-back 1-form $i_0^*\Lambda$ is closed, where Λ is the Liouville 1-form defined in (4.9). This is tantamount to say that the 2-form $i_0^*\mathcal{W} = i_0^* d\Lambda = d i_0^*\Lambda = 0$ where $\mathcal{W} = d\Lambda$ is defined in (4.8). For an "approximately invariant" torus i_0, which supports a linear flow which is only approximately diophantine, i.e. $\omega \in \mathtt{DC}^\gamma_{K_n}$ defined in (1.40), the 1-form $i_0^*\Lambda$ is only "approximately closed". In order to make this statement quantitative we consider

$$i_0^*\Lambda = \sum\nolimits_{k=1}^{\nu} a_k(\varphi) d\varphi_k\,, \quad a_k(\varphi) := -\big([\partial_\varphi \theta_0(\varphi)]^T I_0(\varphi)\big)_k - \frac{1}{2}(\partial_{\varphi_k} z_0(\varphi), J z_0(\varphi))_{L^2(\mathbb{T}_x)} \tag{5.6}$$

and we quantify how small is

$$i_0^*\mathcal{W} = d\, i_0^*\Lambda = \sum\nolimits_{1 \le k < j \le \nu} A_{kj}(\varphi) d\varphi_k \wedge d\varphi_j\,, \quad A_{kj}(\varphi) := \partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi)\,, \tag{5.7}$$

in terms of the "error function"

$$Z(\varphi) := (Z_1, Z_2, Z_3)(\varphi) := \mathcal{F}(i_0, \alpha_0)(\varphi) = \omega \cdot \partial_\varphi i_0(\varphi) - X_{H_\alpha}(i_0(\varphi), \alpha_0)\,, \tag{5.8}$$

and the "ultra-violet" cut-off $K_n = K_0^{\chi^n}$, $\chi = 3/2$, in (1.39), used in the definition (1.40) of $\mathtt{DC}^\gamma_{K_n}$. The main difference with respect to [15] and [8] is that we do not assume ω to be diophantine (i.e. (1.32)) but only $\omega \in \mathtt{DC}^\gamma_{K_n}$.

Along this section we will always assume the following hypothesis, which will be verified at each step of the Nash-Moser iteration of section 8:

- Ansatz. The map $(\omega, \kappa) \mapsto \mathfrak{I}_0(\omega, \kappa) := i_0(\varphi; \omega, \kappa) - (\varphi, 0, 0)$ is k_0-times differentiable with respect to the parameters $(\omega, \kappa) \in \mathbb{R}^\nu \times [\kappa_1, \kappa_2]$, and for some $\mu := \mu(\tau, \nu) > 0$, $\gamma \in (0, 1)$,

$$\|\mathfrak{I}_0\|^{k_0,\gamma}_{s_0+\mu} + |\alpha_0 - \omega|^{k_0,\gamma} \le C\varepsilon\gamma^{-(1+k_1)}\,, \tag{5.9}$$

  where the constant $k_1 = k_1(\nu, k_0) > 0$ is given in Theorem 4.1. We shall always assume $\varepsilon\gamma^{-(1+k_1)}$ small enough (in section 4.2 we have even required the stronger condition $\varepsilon\gamma^{-(1+k_0+k_1)} \ll 1$).

We suppose that the torus $i_0(\omega, \kappa)$ is defined for all the values of $(\omega, \kappa) \in \mathbb{R}^\nu \times [\kappa_1, \kappa_2]$ because, in the Nash-Moser iteration of section 8, we construct a k_0-times differtentiable extension of each approximate solution on the whole $\mathbb{R}^\nu \times [\kappa_1, \kappa_2]$, see Lemma 8.5.

Lemma 5.2. $\|Z\|^{k_0,\gamma}_s \le_s \varepsilon\gamma^{-(1+k_1)} + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+2}$.

Proof. By (4.17), (5.2), (5.9). □

In the following, we will assume that $\omega \in \mathtt{DC}^\gamma_{K_n}$ (defined in (1.40)) and we split the coefficients $A_{kj} = A_{kj}(\varphi)$ in (5.7) as

$$A_{kj} = A^{(n)}_{kj} + A^{(n),\perp}_{kj}\,, \qquad A_{kj} := \Pi_{K_n} A_{kj}\,, \qquad A^{(n),\perp}_{kj} := \Pi^\perp_{K_n} A_{kj} \tag{5.10}$$

where $K_n := K_0^{\chi^n}$, $\chi := 3/2$, is defined in (1.39), the operator Π_{K_n} is the orthogonal projection on the Fourier modes $|(\ell, j)| \le K_n$ and $\Pi^\perp_{K_n} := \mathrm{Id} - \Pi_{K_n}$, see (2.7). The "ultra-violet" cut-off functions K_n are introduced in view of the nonlinear Nash-Moser iteration of section 8.

Lemma 5.3. *Assume that $\omega \in \mathtt{DC}^{\gamma}_{K_n}$ defined in* (1.40)*. Then the coefficients $A^{(n)}_{kj}$ and $A^{(n),\perp}_{kj}$ in* (5.10) *satisfy the following tame estimates*

$$\|A^{(n)}_{kj}\|^{k_0,\gamma}_s \le_s \gamma^{-1}\big(\|Z\|^{k_0,\gamma}_{s+\tau(k_0+1)+k_0+1} + \|Z\|^{k_0,\gamma}_{s_0+1}\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\tau(k_0+1)+k_0+1}\big)\,, \tag{5.11}$$

$$\|A^{(n),\perp}_{kj}\|^{k_0,\gamma}_s \le_s \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+2}\,, \qquad \|A^{(n),\perp}_{kj}\|^{k_0,\gamma}_{s_0+\mathtt{c}} \le_{s_0,b} K_n^{-b}\|\mathfrak{I}_0\|_{s_0+b+\mathtt{c}}\,, \quad \forall b>0\,, \tag{5.12}$$

and for any $\mathtt{c}>0$ such that (5.9) *holds with $\mu \ge \tau(k_0+1)+k_0+1+\mathtt{c}$.*

Proof. Proof of (5.11). The coefficients A_{kj} satisfy the identity (see [15], Lemma 5)

$$\omega\cdot\partial_\varphi A_{kj} = \mathcal{W}\big(\partial_\varphi Z(\varphi)\underline{e}_k, \partial_\varphi i_0(\varphi)\underline{e}_j\big) + \mathcal{W}\big(\partial_\varphi i_0(\varphi)\underline{e}_k, \partial_\varphi Z(\varphi)\underline{e}_j\big)$$

where $\underline{e}_k$ denote the k-th versor of $\mathbb{R}^\nu$. Therefore applying the projector Π_{K_n} we have

$$\omega\cdot\partial_\varphi A^{(n)}_{kj} = \Pi_{K_n}\big[\mathcal{W}\big(\partial_\varphi Z(\varphi)\underline{e}_k, \partial_\varphi i_0(\varphi)\underline{e}_j\big) + \mathcal{W}\big(\partial_\varphi i_0(\varphi)\underline{e}_k, \partial_\varphi Z(\varphi)\underline{e}_j\big)\big]\,.$$

Then by (2.72) and (5.9) we get

$$\|\omega\cdot\partial_\varphi A^{(n)}_{kj}\|^{k_0,\gamma}_s \le_s \|Z\|^{k_0,\gamma}_{s+1} + \|Z\|^{k_0,\gamma}_{s_0+1}\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+1} \tag{5.13}$$

and (5.11) follows applying $(\omega\cdot\partial_\varphi)^{-1}$, and using that, for all $\omega \in \mathtt{DC}^{\gamma}_{K_n}$ defined in (1.40), it results $\|(\omega\cdot\partial_\varphi)^{-1}\Pi_{K_n} g\|^{k_0,\gamma}_s \le_s \gamma^{-1}\|g\|^{k_0,\gamma}_{s+\tau(k_0+1)+k_0}$.

Proof of (5.12). Recalling (5.7) and (5.10), the function $A^{(n),\perp}_{kj}(\varphi) = \Pi^{\perp}_{K_n}\big(\partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi)\big)$ where $a_k(\varphi)$, $k=1,\ldots,\nu$, are defined in (5.6). Then (5.12) follows by the smoothing properties (2.8) and by (2.72), (5.9). ☐

Remark 5.4. If the frequency ω is diophantine, i.e. ω satisfies (1.32), then (5.11) holds with A_{kj} instead of $A^{(n)}_{kj}$ (i.e. $A^{(n),\perp}_{kj}=0$). Furthermore if $Z = \mathcal{F}(i_0,\alpha_0)=0$, then $A_{kj}=0$. ☐

As in [15], [8] we first modify the approximate torus i_0 to obtain an isotropic torus i_δ which is still approximately invariant. We denote the Laplacian $\Delta_\varphi := \sum_{k=1}^{\nu}\partial^2_{\varphi_k}$.

Lemma 5.5. (Isotropic torus) *The torus $i_\delta(\varphi) := (\theta_0(\varphi), I_\delta(\varphi), z_0(\varphi))$ defined by*

$$I_\delta := I_0 + [\partial_\varphi\theta_0(\varphi)]^{-T}\rho(\varphi)\,, \quad \rho_j(\varphi) := \Delta_\varphi^{-1}\sum_{k=1}^{\nu}\partial_{\varphi_j}A_{kj}(\varphi)\,, \quad j=1,\ldots,\nu\,, \tag{5.14}$$

is isotropic. Moreover I_δ admits the splitting $I_\delta = I^{(n)}_\delta + I^{(n),\perp}_\delta$ where

$$I^{(n)}_\delta := I_0 + [\partial_\varphi\theta_0(\varphi)]^{-T}\rho^{(n)}(\varphi)\,, \quad \rho^{(n)}_j(\varphi) := \Delta_\varphi^{-1}\sum_{k=1}^{\nu}\partial_{\varphi_j}A^{(n)}_{kj}(\varphi)\,, \tag{5.15}$$

$$I^{(n),\perp}_\delta := [\partial_\varphi\theta_0(\varphi)]^{-T}\rho^{(n),\perp}(\varphi)\,, \quad \rho^{(n),\perp}_j(\varphi) := \Delta_\varphi^{-1}\sum_{k=1}^{\nu}\partial_{\varphi_j}A^{(n),\perp}_{kj}(\varphi)\,. \tag{5.16}$$

There is $\sigma := \sigma(\nu,\tau,k_0)$ and $\mathtt{c}>0$ such that if (5.9) *holds with $\sigma+\mathtt{c}\le\mu$, then*

$$\|I_\delta - I_0\|^{k_0,\gamma}_s \le \|I^{(n)}_\delta - I_0\|^{k_0,\gamma}_s + \|I^{(n),\perp}_\delta\|^{k_0,\gamma}_s \le_s \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+1} \tag{5.17}$$

$$\|I^{(n)}_\delta - I_0\|^{k_0,\gamma}_s \le_s \gamma^{-1}\big(\|Z\|^{k_0,\gamma}_{s+\sigma} + \|Z\|^{k_0,\gamma}_{s_0+\sigma}\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma}\big)\,, \tag{5.18}$$

$$\|I^{(n),\perp}_\delta\|^{k_0,\gamma}_{s_0+\mathtt{c}} \le_{s_0,b} K_n^{-b}\|\mathfrak{I}_0\|_{s_0+\mathtt{c}+b}\,, \qquad \forall b>0\,, \tag{5.19}$$

$$\|\partial_i[i_\delta][\widehat{\imath}]\|^{k_0,\gamma}_s \le_s \|\widehat{\imath}\|^{k_0,\gamma}_s + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma}\|\widehat{\imath}\|^{k_0,\gamma}_{s_0}\,. \tag{5.20}$$

Moreover the "error" function $Z_\delta := \mathcal{F}(i_\delta,\alpha_0)$ of the isotropic torus i_δ (defined analogously to (5.8)*) may be splitted as $Z_\delta = Z^{(n)}_\delta + Z^{(n),\perp}_\delta$ with*

$$\|Z^{(n)}_\delta\|^{k_0,\gamma}_s \le_s \|Z\|^{k_0,\gamma}_{s+\sigma} + \|Z\|^{k_0,\gamma}_{s_0+\sigma}\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma} \tag{5.21}$$

$$\|Z^{(n),\perp}_\delta\|^{k_0,\gamma}_s \le_s \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma}\,, \qquad \|Z^{(n),\perp}_\delta\|^{k_0,\gamma}_{s_0+\mathtt{c}} \le_{s_0,b} K_n^{-b}\|\mathfrak{I}_0\|^{k_0,\gamma}_{s_0+\sigma+\mathtt{c}+b}\,, \quad \forall b>0\,. \tag{5.22}$$

In the paper we denote equivalently the differential by ∂_i or d_i. Moreover we denote by $\sigma := \sigma(\nu, \tau, k_0)$ possibly different (larger) "loss of derivatives" constants.

Proof. The isotropy of the torus i_δ, defined by (5.14), is proved in Lemma 6 of [15]. The estimate (5.17) follows by (5.14), (5.6), (5.7), (2.72) and (5.9). The estimate (5.18) follows by (5.15) and (5.11). The estimate (5.19) follows by (5.16) and (5.12). The bound (5.20) follows by (5.14), (5.7), (5.6), (5.9). We now prove (5.21), (5.22). One has

$$
\begin{aligned}
\mathcal{F}(i_\delta, \alpha_0) &= \mathcal{F}(i_0, \alpha_0) + \begin{pmatrix} 0 \\ \omega\cdot\partial_\varphi (I_\delta - I_0) \\ 0 \end{pmatrix} + \varepsilon\big(X_P(i_\delta) - X_P(i_0)\big) \\
&= \mathcal{F}(i_0, \alpha_0) + \begin{pmatrix} 0 \\ \omega\cdot\partial_\varphi (I_\delta - I_0) \\ 0 \end{pmatrix} + \varepsilon \int_0^1 \partial_I X_P(t\, i_\delta + (1-t) i_0) \cdot (I_\delta - I_0)\, dt = Z_\delta^{(n)} + Z_\delta^{(n),\perp}
\end{aligned}
$$

where

$$
Z_\delta^{(n)} := \mathcal{F}(i_0, \alpha_0) + \begin{pmatrix} 0 \\ \omega\cdot\partial_\varphi (I_\delta^{(n)} - I_0) \\ 0 \end{pmatrix} + \varepsilon \int_0^1 \partial_I X_P(t\, i_\delta + (1-t) i_0) \cdot (I_\delta^{(n)} - I_0)\, dt\,, \tag{5.23}
$$

$$
Z_\delta^{(n),\perp} := \begin{pmatrix} 0 \\ \omega\cdot\partial_\varphi I_\delta^{(n),\perp} \\ 0 \end{pmatrix} + \varepsilon \int_0^1 \partial_I X_P(t\, i_\delta + (1-t) i_0) \cdot I_\delta^{(n),\perp}\, dt\,. \tag{5.24}
$$

By differentiating (5.15) and, arguing as in [15], [8], we get

$$
\omega\cdot\partial_\varphi (I_\delta^{(n)} - I_0) = [\partial_\varphi \theta_0(\varphi)]^{-T} \omega\cdot\partial_\varphi \rho^{(n)}(\varphi) - \big([\partial_\varphi \theta_0(\varphi)]^{-T} \big(\omega\cdot\partial_\varphi [\partial_\varphi \theta_0(\varphi)]^T\big) [\partial_\varphi \theta_0(\varphi)]^{-T}\big) \rho^{(n)}(\varphi) \tag{5.25}
$$

$$
\omega\cdot\partial_\varphi [\partial_\varphi \theta_0(\varphi)] = \varepsilon \partial_\varphi (\partial_I P)(i_0(\varphi)) + \partial_\varphi Z_1(\varphi)\,. \tag{5.26}
$$

Then (5.21) follows by (5.23), (5.25)-(5.26), (5.3), (2.72), (5.18), (5.9), Lemma 5.2, (5.15), (5.13), (5.11). The estimates (5.22) follow by (5.24), (5.16), (2.72), (5.12), (5.3), (5.17), (5.9) and (5.19). □

In order to find an approximate inverse of the linearized operator $d_{i,\alpha}\mathcal{F}(i_\delta)$ we introduce the symplectic diffeomorpshim $G_\delta : (\phi, y, w) \to (\theta, I, z)$ of the phase space $\mathbb{T}^\nu \times \mathbb{R}^\nu \times H_{\mathbb{S}^+}^\perp$ defined by

$$
\begin{pmatrix} \theta \\ I \\ z \end{pmatrix} := G_\delta \begin{pmatrix} \phi \\ y \\ w \end{pmatrix} := \begin{pmatrix} \theta_0(\phi) \\ I_\delta(\phi) + [\partial_\phi \theta_0(\phi)]^{-T} y - \big[(\partial_\theta \tilde z_0)(\theta_0(\phi))\big]^T J w \\ z_0(\phi) + w \end{pmatrix} \tag{5.27}
$$

where $\tilde z_0(\theta) := z_0(\theta_0^{-1}(\theta))$. It is proved in [15] that G_δ is symplectic, because the torus i_δ is isotropic (Lemma 5.5). In the new coordinates, i_δ is the trivial embedded torus $(\phi, y, w) = (\phi, 0, 0)$. Under the symplectic change of variables G_δ the Hamiltonian vector field X_{H_α} (the Hamiltonian H_α is defined in (4.16)) changes into

$$
X_{K_\alpha} = (DG_\delta)^{-1} X_{H_\alpha} \circ G_\delta \qquad \text{where} \qquad K_\alpha := H_\alpha \circ G_\delta\,. \tag{5.28}
$$

By (4.18) the transformation G_δ is also reversibility preserving and so K_α is reversible, $K_\alpha \circ \tilde\rho = K_\alpha$.

The Taylor expansion of K_α at the trivial torus $(\phi, 0, 0)$ is

$$
\begin{aligned}
K_\alpha(\phi, y, w) = {} & K_{00}(\phi, \alpha) + K_{10}(\phi, \alpha) \cdot y + (K_{01}(\phi, \alpha), w)_{L^2(\mathbb{T}_x)} + \frac{1}{2} K_{20}(\phi) y \cdot y \\
& + \big(K_{11}(\phi) y, w\big)_{L^2(\mathbb{T}_x)} + \frac{1}{2}\big(K_{02}(\phi) w, w\big)_{L^2(\mathbb{T}_x)} + K_{\geq 3}(\phi, y, w)
\end{aligned} \tag{5.29}
$$

where $K_{\geq 3}$ collects the terms at least cubic in the variables (y, w). The Taylor coefficient $K_{00}(\phi, \alpha) \in \mathbb{R}$, $K_{10}(\phi, \alpha) \in \mathbb{R}^\nu$, $K_{01}(\phi, \alpha) \in H_{\mathbb{S}^+}^\perp$, $K_{20}(\phi)$ is a $\nu \times \nu$ real matrix, $K_{02}(\phi)$ is a linear self-adjoint operator of $H_{\mathbb{S}^+}^\perp$ and $K_{11}(\phi) \in \mathcal{L}(\mathbb{R}^\nu, H_{\mathbb{S}^+}^\perp)$.

Note that, by (4.16) and (5.27), the only Taylor coefficients which depend on α are K_{00}, K_{10}, K_{01}. The Hamilton equations associated to (5.29) are

$$\begin{cases} \dot\phi = K_{10}(\phi,\alpha) + K_{20}(\phi)y + K_{11}^T(\phi)w + \partial_y K_{\geq 3}(\phi,y,w) \\ \dot y = \partial_\phi K_{00}(\phi,\alpha) - [\partial_\phi K_{10}(\phi,\alpha)]^T y - [\partial_\phi K_{01}(\phi,\alpha)]^T w \\ \qquad -\partial_\phi\big(\frac12 K_{20}(\phi)y\cdot y + (K_{11}(\phi)y,w)_{L^2(\mathbb{T}_x)} + \frac12 (K_{02}(\phi)w,w)_{L^2(\mathbb{T}_x)} + K_{\geq 3}(\phi,y,w)\big) \\ \dot w = J\big(K_{01}(\phi,\alpha) + K_{11}(\phi)y + K_{02}(\phi)w + \nabla_w K_{\geq 3}(\phi,y,w)\big) \end{cases} \tag{5.30}$$

where $\partial_\phi K_{10}^T$ is the $\nu\times\nu$ transposed matrix and $\partial_\phi K_{01}^T$, $K_{11}^T : H_{\mathbb{S}^+}^\perp \to \mathbb{R}^\nu$ are defined by the duality relation $(\partial_\phi K_{01}[\hat\phi], w)_{L^2_x} = \hat\phi\cdot[\partial_\phi K_{01}]^T w$, $\forall \hat\phi\in\mathbb{R}^\nu, w\in H_{\mathbb{S}^+}^\perp$, and similarly for K_{11}. Explicitly, for all $w\in H_{\mathbb{S}^+}^\perp$, and denoting $\underline{e}_k$ the k-th versor of $\mathbb{R}^\nu$,

$$K_{11}^T(\phi)w = \sum\nolimits_{k=1}^\nu \big(K_{11}^T(\phi)w\cdot\underline{e}_k\big)\underline{e}_k = \sum\nolimits_{k=1}^\nu \big(w, K_{11}(\phi)\underline{e}_k\big)_{L^2(\mathbb{T}_x)}\underline{e}_k \in \mathbb{R}^\nu\,. \tag{5.31}$$

In the next lemma we provide estimates of the coefficients K_{00}, K_{10}, K_{01} in the Taylor expansion (5.29).

Lemma 5.6. *There is $\sigma := \sigma(\tau,\nu,k_0) > 0$ and a decomposition*

$$\partial_\phi K_{00} = \partial_\phi K_{00}^{(n)} + \partial_\phi K_{00}^{(n),\perp}\,,\quad K_{10} = K_{10}^{(n)} + K_{10}^{(n),\perp}\,,\quad K_{01} = K_{01}^{(n)} + K_{01}^{(n),\perp}\,, \tag{5.32}$$

such that, if (5.9) holds with $\mu\geq\sigma+\mathtt{c}$, $\mathtt{c}>0$, then

$$\|\partial_\phi K_{00}^{(n)}(\cdot,\alpha_0)\|_s^{k_0,\gamma} + \|K_{10}^{(n)}(\cdot,\alpha_0)-\omega\|_s^{k_0,\gamma} + \|K_{01}^{(n)}(\cdot,\alpha_0)\|_s^{k_0,\gamma} \leq_s \|Z\|_{s+\sigma}^{k_0,\gamma} + \|Z\|_{s_0+\sigma}^{k_0,\gamma}\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\,, \tag{5.33}$$

$$\|\partial_\phi K_{00}^{(n),\perp}(\cdot,\alpha_0)\|_s^{k_0,\gamma} + \|K_{10}^{(n),\perp}(\cdot,\alpha_0)\|_s^{k_0,\gamma} + \|K_{01}^{(n),\perp}(\cdot,\alpha_0)\|_s^{k_0,\gamma} \leq_s \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\,, \tag{5.34}$$

$$\|\partial_\phi K_{00}^{(n),\perp}(\cdot,\alpha_0)\|_{s_0+\mathtt{c}}^{k_0,\gamma} + \|K_{10}^{(n),\perp}(\cdot,\alpha_0)\|_{s_0+\mathtt{c}}^{k_0,\gamma} + \|K_{01}^{(n),\perp}(\cdot,\alpha_0)\|_{s_0+\mathtt{c}}^{k_0,\gamma} \leq_{s_0,b} K_n^{-b}\|\mathfrak{I}_0\|_{s_0+\sigma+\mathtt{c}+b}^{k_0,\gamma} \tag{5.35}$$

for all $b>0$.

Proof. In Lemma 8 of [15] or Lemma 6.4 of [8] the following identities are proved

$$\begin{aligned} \partial_\phi K_{00}(\phi,\alpha_0) &= -[\partial_\phi\theta_0(\phi)]^T\big(-Z_{2,\delta} - [\partial_\phi I_\delta][\partial_\phi\theta_0]^{-1}Z_{1,\delta} - [(\partial_\theta\tilde z_0)(\theta_0(\phi))]^T J Z_{3,\delta} \\ &\quad - [(\partial_\theta\tilde z_0)(\theta_0(\phi))]^T J\partial_\phi z_0(\phi)[\partial_\phi\theta_0(\phi)]^{-1}Z_{1,\delta}\big)\,, \\ K_{10}(\phi,\alpha_0) &= \omega - [\partial_\phi\theta_0(\phi)]^{-1}Z_{1,\delta}(\phi)\,, \\ K_{01}(\phi,\alpha_0) &= JZ_{3,\delta} - J\partial_\phi z_0(\phi)[\partial_\phi\theta_0(\phi)]^{-1}Z_{1,\delta}(\phi) \end{aligned}$$

where $Z_\delta = (Z_{1,\delta},Z_{2,\delta},Z_{3,\delta}) := \mathcal{F}(i_\delta,\alpha_0)$. According to the splitting $Z_\delta = Z_\delta^{(n)} + Z_\delta^{(n),\perp}$ given in Lemma 5.5, setting $Z_\delta^{(n)} = (Z_{1,\delta}^{(n)}, Z_{2,\delta}^{(n)}, Z_{3,\delta}^{(n)})$, $Z_\delta^{(n),\perp} = (Z_{1,\delta}^{(n),\perp}, Z_{2,\delta}^{(n),\perp}, Z_{3,\delta}^{(n),\perp})$, we get the decomposition (5.32) with

$$\begin{aligned} \partial_\phi K_{00}^{(n)}(\phi,\alpha_0) &= -[\partial_\phi\theta_0(\phi)]^T\big(-Z_{2,\delta}^{(n)} - [\partial_\phi I_\delta][\partial_\phi\theta_0]^{-1}Z_{1,\delta}^{(n)} - [(\partial_\theta\tilde z_0)(\theta_0(\phi))]^T J Z_{3,\delta}^{(n)} \\ &\quad - [(\partial_\theta\tilde z_0)(\theta_0(\phi))]^T J\partial_\phi z_0(\phi)[\partial_\phi\theta_0(\phi)]^{-1}Z_{1,\delta}^{(n)}\big)\,, \\ \partial_\phi K_{00}^{(n),\perp}(\phi,\alpha_0) &= -[\partial_\phi\theta_0(\phi)]^T\big(-Z_{2,\delta}^{(n),\perp} - [\partial_\phi I_\delta][\partial_\phi\theta_0]^{-1}Z_{1,\delta}^{(n),\perp} - [(\partial_\theta\tilde z_0)(\theta_0(\phi))]^T J Z_{3,\delta}^{(n),\perp} \\ &\quad - [(\partial_\theta\tilde z_0)(\theta_0(\phi))]^T J\partial_\phi z_0(\phi)[\partial_\phi\theta_0(\phi)]^{-1}Z_{1,\delta}^{(n),\perp}\big)\,, \\ K_{10}^{(n)}(\phi,\alpha_0) &= \omega - [\partial_\phi\theta_0(\phi)]^{-1}Z_{1,\delta}^{(n)}(\phi)\,, \\ K_{10}^{(n),\perp}(\phi,\alpha_0) &= -[\partial_\phi\theta_0(\phi)]^{-1}Z_{1,\delta}^{(n),\perp}(\phi)\,, \\ K_{01}^{(n)}(\phi,\alpha_0) &= JZ_{3,\delta}^{(n)} - J\partial_\phi z_0(\phi)[\partial_\phi\theta_0(\phi)]^{-1}Z_{1,\delta}^{(n)}(\phi) \\ K_{01}^{(n),\perp}(\phi,\alpha_0) &= JZ_{3,\delta}^{(n),\perp} - J\partial_\phi z_0(\phi)[\partial_\phi\theta_0(\phi)]^{-1}Z_{1,\delta}^{(n),\perp}(\phi)\,. \end{aligned}$$

Then the estimates (5.33)-(5.35) follow by (5.17), (5.21), (5.22), using (2.72) and (5.9). □

We now estimate the variation of the coefficients K_{00}, K_{10}, K_{01} with respect to α. Note, in particular, that $\partial_\alpha K_{10} \approx \mathrm{Id}$ says that the tangential frequencies vary with $\alpha \in \mathbb{R}^\nu$. We also estimate K_{20} and K_{11}.

Lemma 5.7. *We have*

$$\|\partial_\alpha K_{00}\|_s^{k_0,\gamma} + \|\partial_\alpha K_{10} - \mathrm{Id}\|_s^{k_0,\gamma} + \|\partial_\alpha K_{01}\|_s^{k_0,\gamma} \leq_s \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}, \quad \|K_{20}\|_s \leq_s \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big),$$
$$\|K_{11} y\|_s^{k_0,\gamma} \leq_s \varepsilon\big(\|y\|_{s+2}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|y\|_{s_0+2}^{k_0,\gamma}\big), \quad \|K_{11}^T w\|_s^{k_0,\gamma} \leq_s \varepsilon\big(\|w\|_{s+2}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|w\|_{s_0+2}^{k_0,\gamma}\big).$$

Proof. By [15], [8] we have

$$\partial_\alpha K_{00}(\phi) = I_\delta(\phi), \quad \partial_\alpha K_{10}(\phi) = [\partial_\phi \theta_0(\phi)]^{-1}, \quad \partial_\alpha K_{01}(\phi) = J\partial_\theta \tilde z_0(\theta_0(\phi)),$$
$$K_{20}(\varphi) = \varepsilon[\partial_\varphi \theta_0(\varphi)]^{-1}\partial_{II} P(i_\delta(\varphi))[\partial_\varphi\theta_0(\varphi)]^{-T},$$
$$K_{11}(\varphi) = \varepsilon\big(\partial_I \nabla_z P(i_\delta(\varphi))[\partial_\varphi\theta_0(\varphi)]^{-T} + J(\partial_\theta \tilde z_0)(\theta_0(\varphi))(\partial_{II}P)(i_\delta(\varphi))[\partial_\varphi\theta_0(\varphi)]^{-T}\big).$$

Then (5.2), (5.9), (5.17) imply the lemma (the bound for K_{11}^T follows by (5.31)). □

Under the linear change of variables

$$DG_\delta(\varphi,0,0)\begin{pmatrix}\widehat\phi\\ \widehat y\\ \widehat w\end{pmatrix} := \begin{pmatrix}\partial_\phi\theta_0(\varphi) & 0 & 0\\ \partial_\phi I_\delta(\varphi) & [\partial_\phi\theta_0(\varphi)]^{-T} & -[(\partial_\theta \tilde z_0)(\theta_0(\varphi))]^T J\\ \partial_\phi z_0(\varphi) & 0 & I\end{pmatrix}\begin{pmatrix}\widehat\phi\\ \widehat y\\ \widehat w\end{pmatrix} \tag{5.36}$$

the linearized operator $d_{i,\alpha}\mathcal{F}(i_\delta)$ is transformed (approximately, see (5.71) for the precise expression of the error) into the one obtained when we linearize the Hamiltonian system (5.30) at $(\phi, y, w) = (\varphi, 0, 0)$, differentiating also in α at α_0, and changing $\partial_t \rightsquigarrow \omega\cdot\partial_\varphi$, namely

$$\begin{pmatrix}\widehat\phi\\ \widehat y\\ \widehat w\\ \widehat\alpha\end{pmatrix} \mapsto \begin{pmatrix}\omega\cdot\partial_\varphi\widehat\phi - \partial_\phi K_{10}(\varphi)[\widehat\phi] - \partial_\alpha K_{10}(\varphi)[\widehat\alpha] - K_{20}(\varphi)\widehat y - K_{11}^T(\varphi)\widehat w\\ \omega\cdot\partial_\varphi \widehat y + \partial_{\phi\phi}K_{00}(\varphi)[\widehat\phi] + \partial_\phi\partial_\alpha K_{00}(\varphi)[\widehat\alpha] + [\partial_\phi K_{10}(\varphi)]^T\widehat y + [\partial_\phi K_{01}(\varphi)]^T\widehat w\\ \omega\cdot\partial_\varphi\widehat w - J\{\partial_\phi K_{01}(\varphi)[\widehat\phi] + \partial_\alpha K_{01}(\varphi)[\widehat\alpha] + K_{11}(\varphi)\widehat y + K_{02}(\varphi)\widehat w\}\end{pmatrix}. \tag{5.37}$$

As in [8], by (5.36), (5.9), (5.17), the induced composition operator satisfies: for all $\widehat\imath := (\widehat\phi, \widehat y, \widehat w)$

$$\|DG_\delta(\varphi,0,0)[\widehat\imath]\|_s^{k_0,\gamma} + \|DG_\delta(\varphi,0,0)^{-1}[\widehat\imath]\|_s^{k_0,\gamma} \leq_s \|\widehat\imath\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|\widehat\imath\|_{s_0}^{k_0,\gamma}, \tag{5.38}$$
$$\|D^2G_\delta(\varphi,0,0)[\widehat\imath_1,\widehat\imath_2]\|_s^{k_0,\gamma} \leq_s \|\widehat\imath_1\|_s^{k_0,\gamma}\|\widehat\imath_2\|_{s_0}^{k_0,\gamma} + \|\widehat\imath_1\|_{s_0}^{k_0,\gamma}\|\widehat\imath_2\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|\widehat\imath_1\|_{s_0}^{k_0,\gamma}\|\widehat\imath_2\|_{s_0}^{k_0,\gamma}. \tag{5.39}$$

In order to construct an "almost-approximate" inverse of (5.37) we need that

$$\mathcal{L}_\omega := \Pi_{\mathbb{S}^+}^\perp\big(\omega\cdot\partial_\varphi - JK_{02}(\varphi)\big)_{|H_{\mathbb{S}^+}^\perp} \tag{5.40}$$

is "almost invertible" up to the scales $K_n := K_0^{\chi^n}$, $\chi := 3/2$, defined in (1.39), and used for the nonlinear Nash-Moser iteration of section 8. Let $H_\perp^s(\mathbb{T}^{\nu+1}) := H^s(\mathbb{T}^{\nu+1}) \cap H_{\mathbb{S}^+}^\perp$.

- Almost-invertibility assumption. There exists a subset $\Lambda_o \subset \Omega \times [\kappa_1,\kappa_2]$ such that, for all $(\omega,\kappa) \in \Lambda_o$ the operator $\mathcal{L}_\omega$ in (5.40) may be decomposed as

$$\mathcal{L}_\omega = \mathbf{L}_\omega + \mathbf{R}_\omega + \mathbf{R}_\omega^\perp \tag{5.41}$$

where $\mathbf{L}_\omega$ is invertible and $\mathbf{R}_\omega$, $\mathbf{R}_\omega^\perp$ satisfy the estimates (7.93)-(7.95). More precisely for every function $g \in H_\perp^{s+\sigma}(\mathbb{T}^{\nu+1})$ and such that $g(-\varphi) = -\rho g(\varphi)$, there is a solution $h := \mathbf{L}_\omega^{-1} g \in H_\perp^s(\mathbb{T}^{\nu+1})$ such that $h(-\varphi) = \rho h(\varphi)$, of the linear equation $\mathbf{L}_\omega h = g$ which satisfies for all $s_0 \leq s \leq S$ the tame estimate

$$\|\mathbf{L}_\omega^{-1} g\|_s^{k_0,\gamma} \leq_S \gamma^{-1}\big(\|g\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|g\|_{s_0+\sigma}^{k_0,\gamma}\big) \tag{5.42}$$

for some $\sigma := \sigma(\tau,\nu,k_0) > 0$, and the constant $\mu(\mathtt{b}) > 0$ is defined in (7.10).

This assumption shall be verified by Theorem 7.12 at each n-th step of the Nash-Moser nonlinear iteration. It is obtained, in sections 6 and 7, by the process of almost-diagonalization of $\mathcal{L}_\omega$ up to remainders of size $O(\varepsilon N_{n-1}^{\mathtt{a}-1})$ where the larger scales N_n are

$$N_n := K_n^p\,, \quad i.e. \quad N_0 = K_0^p\,, \tag{5.43}$$

and the constant $p > 1$ is large enough, i.e. it satisfies (8.5). The set of "good" parameters Λ_o is contained in particular in the set $\mathtt{DC}^\gamma_{K_n} \times [\kappa_1, \kappa_2]$ defined in (1.40). Actually the parameters in $(\omega, \kappa) \in \Lambda_o$ have to satisfy also first and second order Melnikov non-resonance conditions, see (7.90).

In order to find an almost-approximate inverse of the linear operator in (5.37) (and so of $d_{i,\alpha}\mathcal{F}(i_\delta)$) it is sufficient to almost invert the operator

$$\mathbb{D}[\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\alpha}] := \begin{pmatrix} \omega\cdot\partial_\varphi\widehat{\phi} - \partial_\alpha K_{10}(\varphi)[\widehat{\alpha}] - K_{20}(\varphi)\widehat{y} - K_{11}^T(\varphi)\widehat{w} \\ \omega\cdot\partial_\varphi\widehat{y} + \partial_\phi\partial_\alpha K_{00}(\varphi)[\widehat{\alpha}] \\ \mathbf{L}_\omega\widehat{w} - J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}] - JK_{11}(\varphi)\widehat{y} \end{pmatrix} \tag{5.44}$$

which is obtained by neglecting in (5.37) the terms $\partial_\phi K_{10}$, $\partial_{\phi\phi}K_{00}$, $\partial_\phi K_{00}$, $\partial_\phi K_{01}$ (which vanish at an exact solution by Lemma 5.6), and the small remainders $\mathbf{R}_\omega$, $\mathbf{R}_\omega^\perp$ which appear in (5.41). In addition, since we require only the finitely many non-resonance conditions (1.40), we also decompose $\omega\cdot\partial_\varphi$ as

$$\omega\cdot\partial_\varphi = \mathcal{D}_\omega^{(n)} + \mathcal{D}_\omega^{(n),\perp}\,, \quad \mathcal{D}_\omega^{(n)} := \Pi_{K_n}\omega\cdot\partial_\varphi\Pi_{K_n} + \Pi_{K_n}^\perp\,, \quad \mathcal{D}_\omega^{(n),\perp} := \Pi_{K_n}^\perp\omega\cdot\partial_\varphi\Pi_{K_n}^\perp - \Pi_{K_n}^\perp \tag{5.45}$$

and we further split the operator $\mathbb{D}$ in (5.44) as

$$\mathbb{D} = \mathbb{D}_n + \mathbb{D}_n^\perp \qquad \text{where} \qquad \mathbb{D}_n^\perp[\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\alpha}] := \begin{pmatrix} \mathcal{D}_\omega^{(n),\perp}\widehat{\phi} \\ \mathcal{D}_\omega^{(n),\perp}\widehat{y} \\ 0 \end{pmatrix} \tag{5.46}$$

and

$$\mathbb{D}_n[\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\alpha}] := \begin{pmatrix} \mathcal{D}_\omega^{(n)}\widehat{\phi} - \partial_\alpha K_{10}(\varphi)[\widehat{\alpha}] - K_{20}(\varphi)\widehat{y} - K_{11}^T(\varphi)\widehat{w} \\ \mathcal{D}_\omega^{(n)}\widehat{y} + \partial_\alpha\partial_\phi K_{00}(\varphi)[\widehat{\alpha}] \\ \mathbf{L}_\omega\widehat{w} - J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}] - JK_{11}(\varphi)\widehat{y} \end{pmatrix}\,. \tag{5.47}$$

By the smoothing properties (2.8), the operator $\mathcal{D}_\omega^{(n),\perp}$ satisfies

$$\|\mathcal{D}_\omega^{(n),\perp}h\|_{s_0}^{k_0,\gamma} \leq K_n^{-b}\|h\|_{s_0+b+1}^{k_0,\gamma}\,, \quad \forall b > 0\,, \qquad \|\mathcal{D}_\omega^{(n),\perp}h\|_s^{k_0,\gamma} \leq \|h\|_{s+1}^{k_0,\gamma}\,. \tag{5.48}$$

Lemma 5.8. *Assume that $\omega \in \mathtt{DC}^\gamma_{K_n}$, see (1.40). Then, for all $g \in H^s$ with zero average, the linear equation $\mathcal{D}_\omega^{(n)}h = g$ has a unique solution $h := [\mathcal{D}_\omega^{(n)}]^{-1}g$ with zero average, which satisfies*

$$\|[\mathcal{D}_\omega^{(n)}]^{-1}g\|_s^{k_0,\gamma} \leq \gamma^{-1}\|g\|_{s+\tau_1}^{k_0,\gamma}\,, \qquad \tau_1 := \tau + k_0(\tau+1)\,. \tag{5.49}$$

We look for an exact inverse of $\mathbb{D}_n$ defined in (5.47) by solving the system

$$\mathbb{D}_n[\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\alpha}] := \begin{pmatrix} \mathcal{D}_\omega^{(n)}\widehat{\phi} - \partial_\alpha K_{10}(\varphi)[\widehat{\alpha}] - K_{20}(\varphi)\widehat{y} - K_{11}^T(\varphi)\widehat{w} \\ \mathcal{D}_\omega^{(n)}\widehat{y} + \partial_\alpha\partial_\phi K_{00}(\varphi)[\widehat{\alpha}] \\ \mathbf{L}_\omega\widehat{w} - J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}] - JK_{11}(\varphi)\widehat{y} \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \tag{5.50}$$

where (g_1, g_2, g_3) satisfy the reversibility property

$$g_1(\varphi) = g_1(-\varphi)\,, \quad g_2(\varphi) = -g_2(-\varphi)\,, \quad g_3(\varphi) = -(\rho g_3)(-\varphi)\,. \tag{5.51}$$

We first consider the second equation in (5.50), namely $\mathcal{D}_\omega^{(n)}\widehat{y} = g_2 - \partial_\alpha\partial_\phi K_{00}(\varphi)[\widehat{\alpha}]$. By reversibility, the φ-average of the right hand side of this equation is zero, and so, by Lemma 5.8, its solution is

$$\widehat{y} := [\mathcal{D}_\omega^{(n)}]^{-1}\big(g_2 - \partial_\alpha\partial_\phi K_{00}(\varphi)[\widehat{\alpha}]\big)\,. \tag{5.52}$$

Then we consider the third equation $\mathbf{L}_\omega \widehat{w} = g_3 + JK_{11}(\varphi)\widehat{y} + J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}]$ that, by the inversion assumption (5.42), has a solution

$$\widehat{w} := \mathbf{L}_\omega^{-1}\big(g_3 + JK_{11}(\varphi)\widehat{y} + J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}]\big)\,. \tag{5.53}$$

Finally, we solve the first equation in (5.50), which, substituting (5.52), (5.53), becomes

$$\mathcal{D}_\omega^{(n)}\widehat{\phi} = g_1 + M_1(\varphi)[\widehat{\alpha}] + M_2(\varphi)g_2 + M_3(\varphi)g_3\,, \tag{5.54}$$

where

$$M_1(\varphi) := \partial_\alpha K_{10}(\varphi) - M_2(\varphi)\partial_\alpha\partial_\phi K_{00}(\varphi) + M_3(\varphi)J\partial_\alpha K_{01}(\varphi)\,, \tag{5.55}$$

$$M_2(\varphi) := K_{20}(\varphi)[\mathcal{D}_\omega^{(n)}]^{-1} + K_{11}^T(\varphi)\mathbf{L}_\omega^{-1}JK_{11}(\varphi)[\mathcal{D}_\omega^{(n)}]^{-1}\,, \quad M_3(\varphi) := K_{11}^T(\varphi)\mathbf{L}_\omega^{-1}\,. \tag{5.56}$$

In order to solve the equation (5.54) we have to choose $\widehat{\alpha}$ such that the right hand side has zero average. By Lemma 5.7, (5.9), (5.49) the φ-averaged matrix $\langle M_1\rangle = \mathrm{Id} + O(\varepsilon\gamma^{-(1+k_1)})$. Therefore, for $\varepsilon\gamma^{-(1+k_1)}$ small enough, $\langle M_1\rangle$ is invertible and $\langle M_1\rangle^{-1} = \mathrm{Id} + O(\varepsilon\gamma^{-(1+k_1)})$. Thus we define

$$\widehat{\alpha} := -\langle M_1\rangle^{-1}(\langle g_1\rangle + \langle M_2 g_2\rangle + \langle M_3 g_3\rangle)\,. \tag{5.57}$$

With this choice of $\widehat{\alpha}$, by Lemma 5.8, the equation (5.54) has the solution

$$\widehat{\phi} := [\mathcal{D}_\omega^{(n)}]^{-1}\big(g_1 + M_1(\varphi)[\widehat{\alpha}] + M_2(\varphi)g_2 + M_3(\varphi)g_3\big)\,. \tag{5.58}$$

In conclusion, we have obtained a solution $(\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\alpha})$ of the linear system (5.50).

Proposition 5.9. *Assume* (5.9) *(with $\mu = \mu(\mathtt{b}) + \sigma$) and* (5.42)*. Then, $\forall(\omega,\kappa) \in \Lambda_o$, $\forall g := (g_1, g_2, g_3)$ satisfying* (5.51)*, the system* (5.50) *has a solution $\mathbb{D}_n^{-1}g := (\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\alpha})$ where $(\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\alpha})$ are defined in* (5.58)*,* (5.52)*,* (5.53)*,* (5.57)*, which satisfies* (4.18) *and for any $s_0 \le s \le S$*

$$\|\mathbb{D}_n^{-1}g\|_s^{k_0,\gamma} \le_S \gamma^{-1}\big(\|g\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|g\|_{s_0+\sigma}^{k_0,\gamma}\big). \tag{5.59}$$

Proof. To shorten notations we write $\| \|_s$ instead of $\| \|_s^{k_0,\gamma}$. Recalling (5.56), by Lemma 5.7, (5.42), (5.9), (5.49), we get $\|M_2 g\|_{s_0} + \|M_3 g\|_{s_0} \le C\|g\|_{s_0+\sigma}$. Then, by (5.57) and $\langle M_1\rangle^{-1} = 1 + O(\varepsilon\gamma^{-(1+k_1)}) = O(1)$, we deduce $|\widehat{\alpha}| \le C\|g\|_{s_0+\sigma}$ and (5.52), (5.49) imply $\|\widehat{y}\|_s \le_s \gamma^{-1}\big(\|g\|_{s+\sigma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}\|g\|_{s_0}\big)$. The bound (5.59) is sharp for $\widehat{w}$ because $\mathbf{L}_\omega^{-1}g_3$ in (5.53) is estimated using (5.42). Finally also $\widehat{\phi}$ satisfies (5.59) using (5.58), (5.56), (5.42), (5.49) and Lemma 5.7. □

Finally we prove that the operator

$$\mathbf{T}_0 := \mathbf{T}_0(i_0) := (D\widetilde{G}_\delta)(\varphi, 0, 0) \circ \mathbb{D}_n^{-1} \circ (DG_\delta)(\varphi, 0, 0)^{-1} \tag{5.60}$$

is an almost-approximate right inverse for $d_{i,\alpha}\mathcal{F}(i_0)$ where $\widetilde{G}_\delta(\phi, y, w, \alpha) := \big(G_\delta(\phi, y, w), \alpha\big)$ is the identity on the α-component. We denote the norm $\|(\phi, y, w, \alpha)\|_s^{k_0,\gamma} := \max\{\|(\phi, y, w)\|_s^{k_0,\gamma},\ |\alpha|^{k_0,\gamma}\}$.

Theorem 5.10. (Almost-approximate inverse) *Assume the inversion assumption* (5.41)-(5.42)*. Then, there exists $\bar{\sigma} := \bar{\sigma}(\tau, \nu, k_0) > 0$ such that, if* (5.9) *holds with $\mu = \mu(\mathtt{b}) + \bar{\sigma}$, then for all $(\omega,\kappa) \in \Lambda_o$, for all $g := (g_1, g_2, g_3)$ satisfying* (5.51)*, the operator $\mathbf{T}_0$ defined in* (5.60) *satisfies, for all $s_0 \le s \le S$,*

$$\|\mathbf{T}_0 g\|_s^{k_0,\gamma} \le_S \gamma^{-1}\big(\|g\|_{s+\bar{\sigma}}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar{\sigma}}^{k_0,\gamma}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\big)\,. \tag{5.61}$$

Moreover $\mathbf{T}_0$ is an almost-approximate inverse of $d_{i,\alpha}\mathcal{F}(i_0)$, namely

$$d_{i,\alpha}\mathcal{F}(i_0) \circ \mathbf{T}_0 - \mathrm{Id} = \mathcal{P}(i_0) + \mathcal{P}_\omega(i_0) + \mathcal{P}_\omega^\perp(i_0) \tag{5.62}$$

where, for all $s_0 \le s \le S$,

$$\|\mathcal{P}g\|_s^{k_0,\gamma} \le_S \gamma^{-1}\Big(\|\mathcal{F}(i_0,\alpha_0)\|_{s_0+\bar\sigma}^{k_0,\gamma}\|g\|_{s+\bar\sigma}^{k_0,\gamma} + \big\{\|\mathcal{F}(i_0,\alpha_0)\|_{s+\bar\sigma}^{k_0,\gamma} + \|\mathcal{F}(i_0,\alpha_0)\|_{s_0+\bar\sigma}^{k_0,\gamma}\|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar\sigma}^{k_0,\gamma}\big\}\|g\|_{s_0+\bar\sigma}^{k_0,\gamma}\Big), \tag{5.63}$$

$$\|\mathcal{P}_\omega g\|_s^{k_0,\gamma} \le_S \varepsilon\gamma^{-2}N_{n-1}^{-\mathtt{a}}\big(\|g\|_{s+\bar\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar\sigma}^{k_0,\gamma}\|g\|_{s_0+\bar\sigma}^{k_0,\gamma}\big), \tag{5.64}$$

$$\|\mathcal{P}_\omega^\perp g\|_{s_0}^{k_0,\gamma} \le_{S,b} \gamma^{-1}K_n^{-b}\big(\|g\|_{s_0+\bar\sigma+b}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s_0+\mu(\mathtt{b})+b+\bar\sigma}^{k_0,\gamma}\|g\|_{s_0+\bar\sigma}^{k_0,\gamma}\big), \quad \forall b>0, \tag{5.65}$$

$$\|\mathcal{P}_\omega^\perp g\|_s^{k_0,\gamma} \le_S \gamma^{-1}\big(\|g\|_{s+\bar\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar\sigma}^{k_0,\gamma}\|g\|_{s_0+\bar\sigma}^{k_0,\gamma}\big). \tag{5.66}$$

Proof. The bound (5.61) follows from (5.60), (5.59), (5.38). By (4.17), since $X_{\mathcal{N}}$ does not depend on I, and i_δ differs by i_0 only in the I component (see (5.14)), we have

$$d_{i,\alpha}\mathcal{F}(i_0) - d_{i,\alpha}\mathcal{F}(i_\delta) = \varepsilon\int_0^1 \partial_I d_i X_P(\theta_0, I_\delta + s(I_0-I_\delta), z_0)[I_0 - I_\delta, \Pi[\,\cdot\,]]ds =: \mathcal{E}_0 = \mathcal{E}_0^{(n)} + \mathcal{E}_0^{(n),\perp} \tag{5.67}$$

where Π is the projection $(\widehat\imath,\widehat\alpha)\mapsto\widehat\imath$ and, recalling (5.15), (5.16),

$$\mathcal{E}_0^{(n)} := \varepsilon\int_0^1 \partial_I d_i X_P(\theta_0, I_\delta + s(I_0-I_\delta), z_0)[I_0 - I_\delta^{(n)}, \Pi[\,\cdot\,]]ds, \tag{5.68}$$

$$\mathcal{E}_0^{(n),\perp} := -\varepsilon\int_0^1 \partial_I d_i X_P(\theta_0, I_\delta + s(I_0-I_\delta), z_0)[I_\delta^{(n),\perp}, \Pi[\,\cdot\,]]ds. \tag{5.69}$$

Denote by $\mathtt{u} := (\phi, y, w)$ the symplectic coordinates induced by G_δ in (5.27). Under the symplectic map G_δ, the nonlinear operator $\mathcal{F}$ in (4.17) is transformed into

$$\mathcal{F}(G_\delta(\mathtt{u}(\varphi)),\alpha) = DG_\delta(\mathtt{u}(\varphi))\big(\mathcal{D}_\omega\mathtt{u}(\varphi) - X_{K_\alpha}(\mathtt{u}(\varphi),\alpha)\big) \tag{5.70}$$

where $K_\alpha = H_\alpha\circ G_\delta$, see (5.28) and (5.30). Differentiating (5.70) at the trivial torus $\mathtt{u}_\delta(\varphi) = G_\delta^{-1}(i_\delta)(\varphi) = (\varphi,0,0)$, at $\alpha=\alpha_0$, we get

$$d_{i,\alpha}\mathcal{F}(i_\delta) = DG_\delta(\mathtt{u}_\delta)\big(\omega\cdot\partial_\varphi - d_{\mathtt{u},\alpha}X_{K_\alpha}(\mathtt{u}_\delta,\alpha_0)\big)D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1} + \mathcal{E}_1, \tag{5.71}$$

$$\mathcal{E}_1 := D^2G_\delta(\mathtt{u}_\delta)\big[DG_\delta(\mathtt{u}_\delta)^{-1}\mathcal{F}(i_\delta,\alpha_0),\ DG_\delta(\mathtt{u}_\delta)^{-1}\Pi[\,\cdot\,]\big] = \mathcal{E}_1^{(n)} + \mathcal{E}_1^{(n),\perp} \tag{5.72}$$

where, recalling the splitting $\mathcal{F}(i_\delta,\alpha_0) = Z_\delta = Z_\delta^{(n)} + Z_\delta^{(n),\perp}$ in Lemma 5.5, we have

$$\mathcal{E}_1^{(n)} := D^2G_\delta(\mathtt{u}_\delta)\big[DG_\delta(\mathtt{u}_\delta)^{-1}Z_\delta^{(n)},\ DG_\delta(\mathtt{u}_\delta)^{-1}\Pi[\,\cdot\,]\big] \tag{5.73}$$

$$\mathcal{E}_1^{(n),\perp} := D^2G_\delta(\mathtt{u}_\delta)\big[DG_\delta(\mathtt{u}_\delta)^{-1}Z_\delta^{(n),\perp},\ DG_\delta(\mathtt{u}_\delta)^{-1}\Pi[\,\cdot\,]\big]. \tag{5.74}$$

In expanded form $\omega\cdot\partial_\varphi - d_{\mathtt{u},\alpha}X_{K_\alpha}(\mathtt{u}_\delta,\alpha_0)$ is provided in (5.37). By (5.44), (5.46), (5.47), (5.40), (5.41) and Lemma 5.6 we split

$$\omega\cdot\partial_\varphi - d_{\mathtt{u},\alpha}X_K(\mathtt{u}_\delta,\alpha_0) = \mathbb{D}_n + \mathbb{D}_n^\perp + R_Z^{(n)} + R_Z^{(n),\perp} + \mathbb{R}_\omega + \mathbb{R}_\omega^\perp \tag{5.75}$$

where

$$R_Z^{(n)}[\widehat\phi,\widehat y,\widehat w,\widehat\alpha] := \begin{pmatrix} -\partial_\phi K_{10}^{(n)}(\varphi,\alpha_0)[\widehat\phi] \\ \partial_{\phi\phi}K_{00}^{(n)}(\varphi,\alpha_0)[\widehat\phi] + [\partial_\phi K_{10}^{(n)}(\varphi,\alpha_0)]^T\widehat y + [\partial_\phi K_{01}^{(n)}(\varphi,\alpha_0)]^T\widehat w \\ -J\{\partial_\phi K_{01}^{(n)}(\varphi,\alpha_0)[\widehat\phi]\} \end{pmatrix},$$

$$R_Z^{(n),\perp}[\widehat\phi,\widehat y,\widehat w,\widehat\alpha] := \begin{pmatrix} -\partial_\phi K_{10}^{(n),\perp}(\varphi,\alpha_0)[\widehat\phi] \\ \partial_{\phi\phi}K_{00}^{(n),\perp}(\varphi,\alpha_0)[\widehat\phi] + [\partial_\phi K_{10}^{(n),\perp}(\varphi,\alpha_0)]^T\widehat y + [\partial_\phi K_{01}^{(n),\perp}(\varphi,\alpha_0)]^T\widehat w \\ -J\{\partial_\phi K_{01}^{(n),\perp}(\varphi,\alpha_0)[\widehat\phi]\} \end{pmatrix}$$

and

$$\mathbb{R}_\omega[\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha}] := \begin{pmatrix} 0 \\ 0 \\ \mathbf{R}_\omega[\widehat{w}] \end{pmatrix}, \qquad \mathbb{R}_\omega^\perp[\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha}] := \begin{pmatrix} 0 \\ 0 \\ \mathbf{R}_\omega^\perp[\widehat{w}] \end{pmatrix}.$$

By (5.67), (5.71), (5.72), (5.75) we get the decomposition

$$d_{i,\alpha}\mathcal{F}(i_0) = DG_\delta(\mathtt{u}_\delta) \circ \mathbb{D}_n \circ D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1} + \mathcal{E}^{(n)} + \mathcal{E}_\omega + \mathcal{E}_\omega^\perp \tag{5.76}$$

where

$$\mathcal{E}^{(n)} := \mathcal{E}_0^{(n)} + \mathcal{E}_1^{(n)} + DG_\delta(\mathtt{u}_\delta)R_Z^{(n)}D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}, \quad \mathcal{E}_\omega := DG_\delta(\mathtt{u}_\delta)\mathbb{R}_\omega D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}, \tag{5.77}$$

$$\mathcal{E}_\omega^\perp := \mathcal{E}_0^{(n),\perp} + \mathcal{E}_1^{(n),\perp} + DG_\delta(\mathtt{u}_\delta)[\mathbb{R}_\omega^\perp + \mathbb{D}_n^\perp + R_Z^{(n),\perp}]D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}. \tag{5.78}$$

Applying $\mathbf{T}_0$ defined in (5.60) to the right in (5.76) (recall that $\mathtt{u}_\delta(\varphi) := (\varphi, 0, 0)$), since $\mathbb{D}_n \circ \mathbb{D}_n^{-1} = \mathrm{Id}$ (Proposition 5.9), we get

$$d_{i,\alpha}\mathcal{F}(i_0) \circ \mathbf{T}_0 - \mathrm{Id} = \mathcal{P} + \mathcal{P}_\omega + \mathcal{P}_\omega^\perp,$$
$$\mathcal{P} := \mathcal{E}^{(n)} \circ \mathbf{T}_0, \quad \mathcal{P}_\omega := \mathcal{E}_\omega \circ \mathbf{T}_0, \quad \mathcal{P}_\omega^\perp := \mathcal{E}_\omega^\perp \circ \mathbf{T}_0.$$

Lemma 5.1 and (5.9), (5.33), (5.17), (5.18), (5.21), (5.38)-(5.39), imply the estimate

$$\|\mathcal{E}^{(n)}[\widehat{\imath},\widehat{\alpha}]\|_s^{k_0,\gamma} \le_s \|Z\|_{s_0+\sigma}^{k_0,\gamma}\|\widehat{\imath}\|_{s+\sigma}^{k_0,\gamma} + \|Z\|_{s+\sigma}^{k_0,\gamma}\|\widehat{\imath}\|_{s_0+\sigma}^{k_0,\gamma} + \|Z\|_{s_0+\sigma}^{k_0,\gamma}\|\widehat{\imath}\|_{s_0+\sigma}^{k_0,\gamma}\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \tag{5.79}$$

where $Z := \mathcal{F}(i_0, \alpha_0)$, recall (5.8). Then (5.63) follows from (5.61), (5.79), (5.9). The estimates (5.64), (5.65), (5.66) follow by (7.93)-(7.95), (5.61), (5.38), (5.17), (5.19), (5.22), (5.34), (5.35), (5.9), (5.48). □

6 The linearized operator in the normal directions

In order to write an explicit expression of the linear operator $\mathcal{L}_\omega$ defined in (5.40) we compute the quadratic term $\frac{1}{2}(K_{02}(\phi)w, w)_{L^2(\mathbb{T}_x)}$ in the Taylor expansion of the Hamiltonian $K_\alpha(\phi, 0, w)$ in (5.29).

Lemma 6.1. *The operator $K_{02}(\phi)$ reads*

$$K_{02}(\phi) = \Pi_{\mathbb{S}^+}^\perp \partial_u \nabla_u H(T_\delta(\phi)) + \varepsilon R(\phi) \tag{6.1}$$

where H is the water-waves Hamiltonian defined in (1.6), *evaluated at the torus*

$$T_\delta(\phi) := \varepsilon A(i_\delta(\phi)) = \varepsilon A(\theta_0(\phi), I_\delta(\phi), z_0(\phi)) = \varepsilon v(\theta_0(\phi), I_\delta(\phi)) + \varepsilon z_0(\phi) \tag{6.2}$$

with $A(\theta, I, z)$, $v(\theta, I)$ defined in (4.12). *The operator $K_{02}(\phi)$ is even and reversibility preserving. The remainder $R(\phi)$ has the "finite dimensional" form*

$$R(\phi)[h] = \sum\nolimits_{j=1}^{\nu} \big(h\,,\,g_j\big)_{L^2_x}\chi_j\,, \quad \forall h \in H_{\mathbb{S}^+}^\perp, \tag{6.3}$$

for functions $g_j, \chi_j \in H_{\mathbb{S}^+}^\perp$ which satisfy the tame estimates: for some $\sigma := \sigma(\tau, \nu) > 0$, $\forall s \ge s_0$,

$$\|g_j\|_s^{k_0,\gamma} + \|\chi_j\|_s^{k_0,\gamma} \le 1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{k_0,\gamma}, \quad \|\partial_i g_j[\widehat{\imath}]\|_s + \|\partial_i \chi_j[\widehat{\imath}]\|_s \le \|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma}. \tag{6.4}$$

Proof. The operator $K_{02}(\phi)$ is

$$K_{02}(\phi) = \partial_w \nabla_w K_\alpha(\phi, 0, 0) = \partial_w \nabla_w (H_\alpha \circ G_\delta)(\phi, 0, 0) = \Omega_{|H_{\mathbb{S}^+}^\perp} + \varepsilon \partial_w \nabla_w (P \circ G_\delta)(\phi, 0, 0) \tag{6.5}$$

where $H_\alpha = \mathcal{N}_\alpha + \varepsilon P$ is defined in (4.16) and Ω in (1.14). Differentiating with respect to w the Hamiltonian

$$(P \circ G_\delta)(\phi, y, w) = P(\theta_0(\phi), I_\delta(\phi) + L_1(\phi)y + L_2(\phi)w, z_0(\phi) + w)$$

where (see (5.27)) $L_1(\phi) := [\partial_\phi \theta_0(\phi)]^{-T}$, $L_2(\phi) := -[\partial_\theta \tilde{z}_0(\theta_0(\phi))]^T J$, we get $\nabla_w (P \circ G_\delta)(\phi, y, w) = L_2(\phi)^T \partial_I P(G_\delta(\phi, y, w)) + \nabla_z P(G_\delta(\phi, y, w))$, and therefore

$$\partial_w \nabla_w (P \circ G_\delta)(\phi, 0, 0) = \partial_z \nabla_z P(i_\delta(\phi)) + R(\phi) \qquad \text{with} \qquad R(\phi) := R_1(\phi) + R_2(\phi) + R_3(\phi)\,, \quad (6.6)$$
$$R_1(\phi) := L_2(\phi)^T \partial_{II} P(i_\delta(\phi)) L_2(\phi)\,, \quad R_2(\phi) := L_2(\phi)^T \partial_z \partial_I P(i_\delta(\phi))\,, \quad R_3(\phi) := \partial_I \nabla_z P(i_\delta(\phi)) L_2(\phi)\,.$$

Each operator R_1, R_2, R_3 has the finite dimensional form (6.3) because it is the composition of at least one operator with finite rank $\mathbb{R}^\nu$. For example, writing the operator $L_2(\phi) : H_{\mathbb{S}^+}^\perp \to \mathbb{R}^\nu$ as $L_2(\phi)[h] = \sum_{i=1}^\nu \big(h\,,\, L_2(\phi)^T[e_i]\big)_{L^2_x} e_i$, $\forall h \in H_{\mathbb{S}^+}^\perp$, we get

$$R_1(\phi)[h] = \sum\nolimits_{i=1}^\nu \big(h\,,\, L_2(\phi)^T[e_i]\big)_{L^2_x} A_1[e_i]\,, \quad A_1 := L_2(\phi)^T \partial_{II} P(i_\delta(\phi))\,.$$

Similarly $R_3(\phi)[h] = \sum_{i=1}^\nu \big(h, L_2(\phi)^T[e_i]\big)_{L^2_x} A_3[e_i]$ with $A_3 := \partial_y \nabla_z P(i_\delta(\phi))$, and since $A_2 := \partial_z \partial_I P(i_\delta(\phi)) : H_{\mathbb{S}^+}^\perp \to \mathbb{R}^\nu$, we get $R_2(\phi)[h] = \sum_{i=1}^\nu \big(h, A_2^T[e_i]\big)_{L^2_x} L_2(\phi)^T[e_i]$. The estimate (6.4) follows by Lemma 5.1.

By (6.5), (6.6), and (4.14), (4.12), (4.3), (1.15), we get

$$\begin{aligned} K_{02}(\phi) &= \Omega_{|H_{\mathbb{S}^+}^\perp} + \varepsilon \partial_z \nabla_z P(i_\delta(\phi)) + \varepsilon R(\phi) = \Omega_{|H_{\mathbb{S}^+}^\perp} + \varepsilon \Pi_{\mathbb{S}^+}^\perp \partial_u \nabla_u P_\varepsilon(A(i_\delta(\phi))) + \varepsilon R(\phi) \\ &= \Pi_{\mathbb{S}^+}^\perp \partial_u \nabla_u \mathcal{H}_\varepsilon(A(i_\delta(\phi))) + \varepsilon R(\phi) \end{aligned}$$

which proves (6.1) because $A(i_\delta(\phi)) = T_\delta(\phi)$, see (6.2). □

By Lemma 6.1 the linear operator $\mathcal{L}_\omega$ defined in (5.40) has the form

$$\mathcal{L}_\omega = \Pi_{\mathbb{S}^+}^\perp (\mathcal{L} + \varepsilon R)_{|H_{\mathbb{S}^+}^\perp} \qquad \text{where} \qquad \mathcal{L} := \omega\cdot\partial_\varphi \mathbb{I}_2 - J \partial_u \nabla_u H(T_\delta(\varphi)) \quad (6.7)$$

is obtained linearizing the original water waves system (1.3), (1.5) at the torus $u = (\eta, \psi) = T_\delta(\varphi)$ defined in (6.2), changing $\partial_t \rightsquigarrow \omega\cdot\partial_\varphi$, and denoting the 2×2-identity matrix by

$$\mathbb{I}_2 := \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Id} \end{pmatrix}.$$

Using formula (2.116) the linearized operator $\mathcal{L}$ is

$$\mathcal{L} = \omega\cdot\partial_\varphi \mathbb{I}_2 + \begin{pmatrix} \partial_x V + G(\eta) B & -G(\eta) \\ (1 + B V_x) + B G(\eta) B - \kappa \partial_x c \partial_x & V \partial_x - B G(\eta) \end{pmatrix} \quad (6.8)$$

where the functions $B := B(\varphi, x)$, $V := V(\varphi, x)$ are defined by (2.117) with $(\eta, \psi) = (\eta(\varphi, x), \psi(\varphi, x)) = T_\delta(\varphi)$ defined in (6.2), and

$$c := c(\varphi, x) := (1 + \eta_x^2)^{-3/2}. \quad (6.9)$$

By (6.2), (4.12), (4.18) the function $u = (\eta, \psi) = T_\delta(\varphi)$ is in (even(φ)-even(x), odd(φ)-even(x)), and c is even(φ)-even(x), $B \in$ odd(φ)-even(x), $V =$ odd(φ), odd(x). The operators $\mathcal{L}_\omega$ and $\mathcal{L}$ are real, even and reversible.

Notation. In (6.8) and hereafter any function a is identified with the corresponding multiplication operators $h \mapsto ah$, and, where there is no parenthesis, composition of operators is understood. For example, $\partial_x c \partial_x$ means: $h \mapsto \partial_x(c \partial_x h)$.

In the next sections we focus on reducing the linear operator $\mathcal{L}$ in (6.8) to constant coefficients up to a pseudo-differential operator of order 0 (and up to a small remainder supported on the high modes). The finite dimensional remainder εR transforms under conjugation into an operator of the same form (Lemma 6.30) and therefore it will be dealt only once at the end of the section.

For the sequel we will always assume the following ansatz in "low norm" (that will be satisfied by the approximate solutions along the Nash-Moser iteration): for some $\mu := \mu(\tau, \nu) > 0$, $\gamma \in (0, 1)$,

$$\|\mathfrak{I}_0\|_{s_0+\mu}^{k_0,\gamma} \le 1\ , \qquad \text{and so, by (5.17),}\ \ \|\mathfrak{I}_\delta\|_{s_0+\mu}^{k_0,\gamma} \le 2\,. \quad (6.10)$$

Actually $\mu := \mu(\mathtt{b}) + \sigma_1$, where $\mu(\mathtt{b})$ is defined in (7.10) and σ_1 in (8.4), is fixed in the Nash Moser iteration of section 8 (see also (8.8)). In order to estimate the variation of the eigenvalues with respect to the approximate invariant torus, we need also to estimate the derivatives with respect to the torus $i(\varphi)$ in another low norm $\| \ \|_{s_1}$, for all the Sobolev indices s_1 such that

$$s_1 + \sigma \leq s_0 + \mu \,, \quad \text{for some} \quad \sigma := \sigma(\tau, \nu) > 0 \,. \tag{6.11}$$

Thus by (6.10) we have

$$\|\mathfrak{I}_0\|_{s_1+\sigma}^{k_0,\gamma} \leq 1 \qquad \text{and so, by (5.17),} \ \|\mathfrak{I}_\delta\|_{s_1+\sigma}^{k_0,\gamma} \leq 2 \,. \tag{6.12}$$

The constants μ and σ represent the *loss of derivatives* at any step of the reduction procedure of this section and it possibly increases along the (finitely many) steps of this reduction procedure. In Lemma 7.2 we fix the largest loss of derivatives $\sigma := \sigma(\mathtt{b})$.

Remark 6.2. Let us shortly motivate the role of the intermediate Sobolev index s_1. In the reducibility scheme in section 7 we require that the remainders $\mathbf{R}_0$, $\mathbf{Q}_0$ satisfy the estimates (7.8). In Lemma 7.2 we take $\mathbf{R}_0 := \mathbf{R}_M^{(3)}$, $\mathbf{Q}_0 := \mathbf{Q}_M^{(3)}$ defined in Proposition 6.31 and so we want that (6.250) holds with $s_1 = s_0$. For that we need to estimate, along section 6, the derivatives ∂_i of functions, operators, etc, in intermediate $\| \ \|_{s_1}$ norms, i.e. for s_1 which satisfies (6.11). □

As a consequence of Moser composition Lemma 2.22, the Sobolev norm of $u = T_\delta$ (see (6.2)) satisfies

$$\|u\|_s^{k_0,\gamma} = \|\eta\|_s^{k_0,\gamma} + \|\psi\|_s^{k_0,\gamma} \leq \varepsilon C(s)\big(1 + \|\mathfrak{I}_0\|_s^{k_0,\gamma}\big) \,, \quad \forall s \geq s_0 \tag{6.13}$$

(the funtion A defined in (4.12) is smooth). Similarly

$$\|\partial_i u[\hat\imath]\|_{s_1} \leq_{s_1} \varepsilon \|\hat\imath\|_{s_1} \,. \tag{6.14}$$

We remark that it would be sufficient to give Lipschitz estimates of u (and of operators, transformations, eigenvalues) with respect to the variable i, namely to estimate the finite difference $\Delta_{12} u := u(i_1) - u(i_2)$ in terms of the difference $\|i_1 - i_2\|_{s_1+\sigma}$, but for convenience we compute the derivatives ∂_i. We repeat that it is sufficient to estimate the derivatives (or the finite difference) with respect to i only in low norm s_1 is because this information is only needed to control the variation of the eigenvalues with respect to i, see remark 7.4.

Finally we recall that $\mathfrak{I}_0 := \mathfrak{I}_0(\omega, \kappa)$ is defined for all $\omega \in \mathbb{R}^\nu$ by the extension procedure of section 8. Moreover all the functions appearing in $\mathcal{L}$ in (6.8) are $\mathcal{C}^\infty$ in (φ, x) as the approximate torus $u = (\eta, \psi) = T_\delta(\varphi)$. This enables to use directly pseudo-differential operator theory as reminded in section 2.

6.1 Linearized good unknown of Alinhac

We first conjugate the linearized operator $\mathcal{L}$ in (6.8) by the change of variable

$$\mathcal{Z} := \begin{pmatrix} 1 & 0 \\ B & 1 \end{pmatrix} , \qquad \mathcal{Z}^{-1} = \begin{pmatrix} 1 & 0 \\ -B & 1 \end{pmatrix}$$

obtaining

$$\mathcal{L}_0 := \mathcal{Z}^{-1} \mathcal{L} \mathcal{Z} = \omega \cdot \partial_\varphi \mathbb{I}_2 + \begin{pmatrix} \partial_x V & -G(\eta) \\ a - \kappa \partial_x c \partial_x & V \partial_x \end{pmatrix} \tag{6.15}$$

where a is the function

$$a := a(\varphi, x) = 1 + \omega \cdot \partial_\varphi B + V B_x \,. \tag{6.16}$$

The matrix $\mathcal{Z}$ amounts to introduce (a linearized version of) the "good unknown of Alinhac".

Lemma 6.3. *The maps $\mathcal{Z}^{\pm 1} - \mathrm{Id}$ are even, reversibility preserving and $\mathcal{D}^{k_0}$-tame with tame constant satisfying, for all $s_0 \le s \le S$,*

$$\mathfrak{M}_{\mathcal{Z}^{\pm1}-\mathrm{Id}}(s)\,,\,\mathfrak{M}_{(\mathcal{Z}^{\pm1}-\mathrm{Id})^*}(s) \le_s \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{6.17}$$

The operator $\mathcal{L}_0$ is even and reversible. There is $\sigma := \sigma(\tau,\nu) > 0$ such that the functions

$$\|a-1\|_s^{k_0,\gamma} + \|V\|_s^{k_0,\gamma} + \|B\|_s^{k_0,\gamma} \le_s \varepsilon\big(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \|c-1\|_s^{k_0,\gamma} \le_s \varepsilon^2\big(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{6.18}$$

Moreover

$$\|\partial_i a[\hat\imath]\|_{s_1} + \|\partial_i V[\hat\imath]\|_{s_1} + \|\partial_i B[\hat\imath]\|_{s_1} \le_{s_1} \varepsilon\|\hat\imath\|_{s_1+\sigma}\,, \quad \|\partial_i c[\hat\imath]\|_{s_1} \le_{s_1} \varepsilon^2\|\hat\imath\|_{s_1+\sigma} \tag{6.19}$$

$$\|\partial_i(\mathcal{Z}^{\pm1}[\hat\imath])h\|_{s_1}\,,\,\|\partial_i((\mathcal{Z}^{\pm1})^*[\hat\imath])h\|_{s_1} \le_{s_1} \varepsilon\|\hat\imath\|_{s_1+\sigma}\|h\|_{s_1}\,. \tag{6.20}$$

Proof. The estimate (6.18), follows by the explicit expressions of a, V, B, c in (6.16), (2.117), (6.9), by applying Lemma 2.22 and the estimates (2.72), (2.120), (2.68) and Lemma 2.14. The operators $\mathcal{Z}^{\pm1}$ are reversibility preserving because B is oddφ. The estimate (6.17) holds by (2.39), (2.68), (6.18) and since the adjoint $\mathcal{Z}^* = \begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix}$. The estimates involving $\mathcal{Z}^{-1}$ follow similarly. The estimate (6.19) follows by differentiating the explicit expressions of a, B, V, c in (6.16), (2.117), (6.9), by applying Lemma 2.22, (2.116), (2.120), (2.72) and (6.14). The estimates (6.20) follow by the estimate of $\partial_i B$ in (6.19) and (2.72). □

6.2 Symmetrization and space reduction of the highest order

The aim of this section is to conjugate the linear operator $\mathcal{L}_0$ in (6.15) to the operator $\mathcal{L}_3$ in (6.58) whose coefficient $m_3(\varphi)$ of the highest order is independent of the space variable. By (2.118) we first rewrite

$$\mathcal{L}_0 = \omega\cdot\partial_\varphi \mathbb{I}_2 + \begin{pmatrix} V\partial_x + V_x & -|D_x| - \mathcal{R}_G \\ a - \kappa c\partial_{xx} - \kappa c_x\partial_x & V\partial_x \end{pmatrix}\,. \tag{6.21}$$

Step 1. We first conjugate $\mathcal{L}_0$ with a change of variable

$$(\mathcal{B}h)(\varphi,x) := h(\varphi, x+\beta(\varphi,x)) \tag{6.22}$$

induced by a φ-dependent family of diffeomorphisms of the torus

$$y = x + \beta(\varphi,x) \quad \Leftrightarrow \quad x = y + \tilde\beta(\varphi,y) \tag{6.23}$$

where $\beta(\varphi,x)$ is a small periodic function to be determined. Under the change of variable (6.22) the differential operators ∂_x, ∂_{xx}, $\omega\cdot\partial_\varphi$, and the multiplication operator by a, transform into

$$\mathcal{B}^{-1}\partial_x\mathcal{B} = \{\mathcal{B}^{-1}(1+\beta_x)\}\partial_y, \qquad \mathcal{B}^{-1}\partial_{xx}\mathcal{B} = \{\mathcal{B}^{-1}(1+\beta_x)\}^2\partial_{yy} + (\mathcal{B}^{-1}\beta_{xx})\partial_y, \tag{6.24}$$

$$\mathcal{B}^{-1}\omega\cdot\partial_\varphi\mathcal{B} = \omega\cdot\partial_\varphi + (\mathcal{B}^{-1}\omega\cdot\partial_\varphi\beta)\partial_y\,, \qquad \mathcal{B}^{-1}a\mathcal{B} = (\mathcal{B}^{-1}a). \tag{6.25}$$

Moreover, using (6.24),

$$\begin{aligned}\mathcal{B}^{-1}|D_x|\mathcal{B} = \mathcal{B}^{-1}\partial_x\mathcal{H}\mathcal{B} = (\mathcal{B}^{-1}\partial_x\mathcal{B})(\mathcal{B}^{-1}\mathcal{H}\mathcal{B}) &= \{\mathcal{B}^{-1}(1+\beta_x)\}\partial_y[\mathcal{H} + (\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H})] \\ &= \{\mathcal{B}^{-1}(1+\beta_x)\}|D_y| + \mathcal{R}_{\mathcal{B}}\end{aligned} \tag{6.26}$$

where, by Lemma 2.27,

$$\mathcal{R}_{\mathcal{B}} := \{\mathcal{B}^{-1}(1+\beta_x)\}\partial_y(\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H}) \in OPS^{-\infty}\,. \tag{6.27}$$

Thus, by (6.24)-(6.26), the operator $\mathcal{L}_0$ in (6.21) transforms into

$$\mathcal{L}_1 := \mathcal{B}^{-1}\mathcal{L}_0\mathcal{B} = \omega\cdot\partial_\varphi\mathbb{I}_2 + \begin{pmatrix} a_1\partial_y + a_2 & -a_3|D_y| + \mathcal{R}_1 \\ -\kappa a_4\partial_{yy} - \kappa a_5\partial_y + a_6 & a_1\partial_y \end{pmatrix} \tag{6.28}$$

where $a_i = a_i(\varphi, y)$ are

$$a_1 := \mathcal{B}^{-1}[\omega\cdot\partial_\varphi\beta + V(1+\beta_x)], \qquad a_2 := \mathcal{B}^{-1}(V_x), \qquad a_3 := \mathcal{B}^{-1}(1+\beta_x), \tag{6.29}$$

$$a_4 := \mathcal{B}^{-1}[c(1+\beta_x)^2], \qquad a_5 := \mathcal{B}^{-1}[c\beta_{xx} + c_x(1+\beta_x)], \qquad a_6 := \mathcal{B}^{-1}a, \tag{6.30}$$

and

$$\mathcal{R}_1 := -\mathcal{R}_{\mathcal{B}} - \mathcal{B}^{-1}\mathcal{R}_G\mathcal{B} \in OPS^{-\infty}\,. \tag{6.31}$$

We look for $\beta(\varphi, x)$ such that

$$(a_3a_4)(\varphi, y) = m(\varphi) \tag{6.32}$$

for some function $m(\varphi)$, independent of the space variable y. By (6.29)-(6.30), the equation (6.32) is

$$c(\varphi, x)(1+\beta_x(\varphi, x))^3 = m(\varphi)$$

which is solved by

$$m(\varphi) := \Big(\frac{1}{2\pi}\int_{\mathbb{T}} c(\varphi, x)^{-\frac{1}{3}}\,dx\Big)^{-3}, \qquad \beta(\varphi, x) := \partial_x^{-1}\big(m(\varphi)^{\frac{1}{3}}c(\varphi, x)^{-\frac{1}{3}} - 1\big)\,, \tag{6.33}$$

where ∂_x^{-1} is the Fourier multiplier

$$\partial_x^{-1}e^{\mathrm{i}jx} := \frac{e^{\mathrm{i}jx}}{\mathrm{i}j}\,,\ \forall j \neq 0\,, \qquad \partial_x^{-1}1 := 0\,.$$

Remark 6.4. Since c is even(φ)-even(x), it follows that $\beta = \mathrm{even}(\varphi), \mathrm{odd}(x)$. As a consequence, $\mathcal{B}, \mathcal{B}^{-1}$ are even and reversibility preserving. Therefore $a_1 = \mathrm{odd}(\varphi), \mathrm{odd}(x)$, $a_2 = \mathrm{odd}(\varphi), \mathrm{even}(x)$, $a_3, a_4, a_6 = \mathrm{even}(\varphi), \mathrm{even}(x)$, $a_5 = \mathrm{even}(\varphi), \mathrm{odd}(x)$. $\square$

Step 2. We conjugate $\mathcal{L}_1$ in (6.28) by the linear map

$$\mathcal{Q} := \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix}, \quad \mathcal{Q}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & q^{-1} \end{pmatrix}\,,$$

where $q(\varphi, x)$ is a real valued function close to 1 to be determined. We compute

$$\mathcal{L}_2 := \mathcal{Q}^{-1}\mathcal{L}_1\mathcal{Q} = \omega\cdot\partial_\varphi\mathbb{I}_2 + \begin{pmatrix} a_1\partial_y + a_2 & -a_3q|D_y| - a_3q_y\mathcal{H} + \mathcal{R}_2 \\ -\kappa q^{-1}a_4\partial_{yy} - \kappa q^{-1}a_5\partial_y + q^{-1}a_6 & a_1\partial_y + q^{-1}(\omega\cdot\partial_\varphi q) + q^{-1}a_1q_y \end{pmatrix} \tag{6.34}$$

where, by Lemma 2.26 and (6.31), the remainder

$$\mathcal{R}_2 := \mathcal{R}_1q - a_3[\mathcal{H}, q]\partial_y - a_3[\mathcal{H}, q_y] \in OPS^{-\infty}\,. \tag{6.35}$$

We choose the function q so that the coefficients of the off diagonal highest order terms satisfy

$$a_3q = q^{-1}a_4\,, \quad \textit{i.e.} \quad q := \sqrt{a_4/a_3} \tag{6.36}$$

(note that a_3, a_4 are close to 1). Thus by (6.36), (6.32), (6.33), (6.9) we get

$$a_3q = q^{-1}a_4 = m_3(\varphi)\,, \quad m_3(\varphi) := \sqrt{m(\varphi)} = \Big(\frac{1}{2\pi}\int_{\mathbb{T}}\sqrt{1+\eta_x^2}\,dx\Big)^{-3/2}\,, \tag{6.37}$$

and, by (6.34),

$$\mathcal{L}_2 = \omega\cdot\partial_\varphi\mathbb{I}_2 + \begin{pmatrix} a_1\partial_y + a_2 & -m_3(\varphi)|D_y| + a_7\mathcal{H} + \mathcal{R}_2 \\ m_3(\varphi)(1-\kappa\partial_{yy}) + a_8\partial_y + b_9 & a_1\partial_y + b_{10} \end{pmatrix} \tag{6.38}$$

where

$$a_7 := -a_3q_y\,, \quad a_8 := -\kappa q^{-1}a_5\,, \quad b_9 := q^{-1}a_6 - m_3(\varphi)\,, \quad b_{10} := q^{-1}\big(\omega\cdot\partial_\varphi q + a_1q_y\big)\,. \tag{6.39}$$

Remark 6.5. Since $a_4, a_3 \in \text{even}(\varphi), \text{even}(x)$, the function $q \in \text{even}(\varphi), \text{even}(x)$, hence the operator $\mathcal{Q}$ is even and reversibility preserving. Moreover $a_7, a_8 = \text{even}(\varphi)\text{odd}(x)$, $b_9 \in \text{even}(\varphi), \text{even}(x)$, $b_{10} = \text{odd}(\varphi)\text{even}(x)$. □

Lemma 6.6. *The operators $\mathcal{B}^{\pm 1}$ are $\mathcal{D}^{k_0}$-(k_0+1)-tame, $\mathcal{Q}^{\pm 1}$ are $\mathcal{D}^{k_0}$-tame with tame constants satisfying*

$$\mathfrak{M}_{\mathcal{B}}(s)\,,\, \mathfrak{M}_{\mathcal{Q}}(s) \leq_S 1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\,, \quad \forall s_0 \leq s \leq S\,. \tag{6.40}$$

The operators $\mathcal{B}^{\pm 1} - \mathrm{Id}$, $(\mathcal{B}^{\pm 1} - \mathrm{Id})^*$ is $\mathcal{D}^{k_0}$-(k_0+2)-tame and $\mathcal{Q}^{\pm 1} - \mathrm{Id}$, $(\mathcal{Q}^{\pm 1} - \mathrm{Id})^*$ are $\mathcal{D}^{k_0}$-tame and, for all $s_0 \leq s \leq S$,

$$\mathfrak{M}_{\mathcal{B}^{\pm 1} - \mathrm{Id}}(s)\,,\, \mathfrak{M}_{(\mathcal{B}^{\pm 1} - \mathrm{Id})^*}(s)\,,\, \mathfrak{M}_{\mathcal{Q}^{\pm 1} - \mathrm{Id}}(s)\,,\, \mathfrak{M}_{(\mathcal{Q}^{\pm 1} - \mathrm{Id})^*}(s) \leq_S \varepsilon(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})\,. \tag{6.41}$$

The functions m_3 satisfies

$$\|m_3 - 1\|_s^{k_0,\gamma} \leq_s \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \|\partial_i m_3[\widehat{\imath}]\|_{s_1} \leq_{s_1} \varepsilon \|\widehat{\imath}\|_{s_1+\sigma} \tag{6.42}$$

and the functions a_i satisfy

$$\max\{\|a_1\|_s^{k_0,\gamma}, \|a_2\|_s^{k_0,\gamma}, \|a_7\|_s^{k_0,\gamma}, \|a_8\|_s^{k_0,\gamma}, \|b_9\|_s^{k_0,\gamma}, \|b_{10}\|_s^{k_0,\gamma}\} \leq_S \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{6.43}$$

The remainder $\mathcal{R}_2$ in (6.35) *is in $OPS^{-\infty}$ and, for some $\sigma := \sigma(\tau, \nu) > 0$, for all $m \geq 0$, $s \geq 0$, $\alpha \in \mathbb{N}$,*

$$|\mathcal{R}_2|_{-m,s,\alpha}^{k_0,\gamma} \leq_{m,S,\alpha} \varepsilon(1 + \|\mathfrak{I}_0\|_{s+\sigma+m+\alpha}^{k_0,\gamma})\,. \tag{6.44}$$

Moreover

$$\|(\partial_i A[\hat{\imath}])h\|_{s_1} \leq_S \varepsilon \|\hat{\imath}\|_{s_1+\sigma} \|h\|_{s_1+\sigma}\,, \qquad A \in \{\mathcal{B}^{\pm 1}, \mathcal{Q}^{\pm 1}, (\mathcal{B}^{\pm 1})^*, (\mathcal{Q}^{\pm 1})^*\}\,, \tag{6.45}$$

$$\|\partial_i a_1[\hat{\imath}]\|_{s_1}, \|\partial_i a_2[\hat{\imath}]\|_{s_1}, \|\partial_i a_7[\hat{\imath}]\|_{s_1}, \|\partial_i a_8[\hat{\imath}]\|_{s_1}, \|\partial_i b_9[\hat{\imath}]\|_{s_1}, \|\partial_i b_{10}[\hat{\imath}]\|_{s_1} \leq_S \varepsilon \|\hat{\imath}\|_{s_1+\sigma} \tag{6.46}$$

and for all $m \geq 0$, $\alpha \in \mathbb{N}$

$$|\partial_i \mathcal{R}_2[\hat{\imath}]|_{-m,s_1,\alpha} \leq_{m,S,\alpha} \varepsilon \|\hat{\imath}\|_{s_1+\sigma+m+\alpha}\,. \tag{6.47}$$

Proof. The estimates (6.40), (6.43) follows by (6.37), (6.29), (6.30), (6.39), using (2.72) and Lemmata 6.3, 2.22, 2.21, 2.13. The estimate (6.44) follows by Lemmas 2.21, 2.25, 2.26, 2.27, Proposition 2.28, (6.13), and (2.72). The estimate (6.41) for $\mathcal{Q} = \mathcal{Q}^*$ follows since the function $q(\varphi, x)$ is close to 1, and it satisfies $\|q - 1\|_s^{k_0,\gamma} \leq_s \varepsilon(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})$, for some $\sigma := \sigma(k_0, \tau, \nu) > 0$. The estimate for $\mathcal{B} - \mathrm{Id}$ follows by

$$(\mathcal{B} - \mathrm{Id})h = \beta\, \mathcal{B}_\tau[h_x]\,, \qquad \mathcal{B}_\tau[h](\varphi, x) := \int_0^1 h_x(\varphi, x + \tau\beta(\varphi, x))\, d\tau$$

and the estimate for the adjoint $(\mathcal{B} - \mathrm{Id})^*$ follows by the representation

$$\mathcal{B}^* h(\varphi, y) = (1 + \tilde{\beta}(\varphi, y)) h(\varphi, y + \tilde{\beta}(\varphi, y)) \tag{6.48}$$

where $y \mapsto y + \tilde{\beta}(\varphi, y)$ is the inverse diffeomorphism of $x \mapsto x + \beta(\varphi, x)$. The expressions of $\mathcal{B}^{-1} - \mathrm{Id}$ and $(\mathcal{B}^{-1})^*$ are similar.

Let us prove the estimate (6.45) for $\mathcal{B}$ and $\mathcal{B}^{-1}$. The other estimates follow analogously. By (6.33) and using the estimates (6.18), (6.19) on c we get

$$\|\partial_i \beta[\hat{\imath}]\|_{s_1} \leq_{s_1} \varepsilon \|\hat{\imath}\|_{s_1+\sigma} \tag{6.49}$$

then the estimate (6.45) for $\mathcal{B}$ follows since $(\partial_i \mathcal{B}[\hat{\imath}])h = \partial_i \beta[\hat{\imath}] \mathcal{B}[h_x]$. Since $y = x + \beta(x)$ if and only if $x = y + \tilde{\beta}(y)$, differentiating with respect to i we get $\partial_i \tilde{\beta}[\hat{\imath}] = (1 + \beta_x)^{-1} \mathcal{B}^{-1}[\partial_i \beta[\hat{\imath}]]$, hence $\partial_i \tilde{\beta}$ satisfies (6.49) (for a possibly larger $\sigma := \sigma(\tau, \nu) > 0$), and hence $\mathcal{B}^{-1}$ satisfies (6.45). The estimates (6.46) follows by differentiating the explicit expressions of the coefficients and applying (2.72), the estimates of Lemma 6.3, (6.45) for $\mathcal{B}^{\pm 1}$ and Lemma 2.22. By (6.36), $\partial_i q$ satisfies (6.46), therefore $\mathcal{Q}$ and $\mathcal{Q}^{-1}$ satisfy (6.45). For proving (6.47) for $\partial_i \mathcal{R}_2[\hat{\imath}]$ we show that the derivative ∂_i of each term in (6.35) satisfies the estimate (6.47). For instance the term $\partial_i[\mathcal{H}, q][\widehat{\imath}] = [\mathcal{H}, \partial_i q[\widehat{\imath}]]$ can be estimated by applying Lemma 2.26 and using that $\partial_i q[\widehat{\imath}]$ (the function q is defined in (6.36)) satisfies the same bound (6.46). For estimating $\partial_i \mathcal{R}_1[\widehat{\imath}]$ we estimate separately the derivatives of the two terms $\mathcal{B}^{-1} \mathcal{R}_G \mathcal{B}$ and $\mathcal{R}_{\mathcal{B}}$ in (6.31). The operator $\partial_i(\mathcal{B}^{-1} \mathcal{R}_G \mathcal{B})[\widehat{\imath}]$ satisfies the estimate (6.47) by (2.129)-(2.130) by Lemmata 2.23, 2.25, 2.27, Proposition 2.28 and (6.40), (6.41), (6.45), (6.14). The estimate of the operator $\partial_i \mathcal{R}_{\mathcal{B}}[\widehat{\imath}]$ in (6.27), follows similarly. □

Step 3. We "symmetrize" the order of derivatives in the off-diagonal terms of the operator $\mathcal{L}_2$ in (6.38). We conjugate $\mathcal{L}_2$ by the vector valued Fourier multiplier

$$\mathcal{S} = \begin{pmatrix} 1 & 0 \\ 0 & G \end{pmatrix}, \quad \mathcal{S}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & G^{-1} \end{pmatrix}, \quad G := \mathrm{Op}(g(\xi)) \in OPS^{1/2} \tag{6.50}$$

where g is a $\mathcal{C}^\infty$ even function satisfying

$$g(0) = 1\,, \quad g > 0\,, \quad g(\xi) = |\xi|^{-\frac{1}{2}}(1+\kappa\xi^2)^{\frac{1}{2}}\,, \quad \forall |\xi| \geq 1/3\,. \tag{6.51}$$

Note that $\mathcal{S}$ is a real and even operator, see Lemma 2.4. Recalling the definition of the cut off function χ in (2.26), the symbols $g \in S^{\frac{1}{2}}$ and $1/g \in S^{-\frac{1}{2}}$ admit the expansions

$$g(\xi) = \chi(\xi)g(\xi) + (1-\chi(\xi))g(\xi) = \chi(\xi)\frac{(1+\kappa\xi^2)^{\frac{1}{2}}}{|\xi|^{1/2}} + (1-\chi(\xi))g(\xi) = \sqrt{\kappa}\chi(\xi)|\xi|^{\frac{1}{2}} + g_{-\frac{3}{2}}(\xi) \tag{6.52}$$

where $g_{-\frac{3}{2}} \in S^{-\frac{3}{2}}$ and

$$\frac{1}{g(\xi)} = \frac{\chi(\xi)}{g(\xi)} + \frac{1-\chi(\xi)}{g(\xi)} = \chi(\xi)\frac{|\xi|^{\frac{1}{2}}}{(1+\kappa\xi^2)^{\frac{1}{2}}} + \frac{1-\chi(\xi)}{g(\xi)} = \frac{\chi(\xi)}{\sqrt{\kappa}|\xi|^{\frac{1}{2}}} + g_{-\frac{5}{2}}(\xi)\,, \quad g_{-\frac{5}{2}} \in S^{-\frac{5}{2}}\,. \tag{6.53}$$

Since $\frac{1-\chi(\xi)}{g(\xi)} = 0$, for $|\xi| \geq 1$, and $\frac{1-\chi(0)}{g(0)} = 1$, the operator $\mathrm{Op}\Big(\frac{1-\chi(\xi)}{g(\xi)}\Big) = \pi_0$ on the periodic functions, where π_0 is the projector

$$\pi_0(f) := \frac{1}{2\pi}\int_{\mathbb{T}} f(x)\,dx\,. \tag{6.54}$$

By (6.52)-(6.53) we get the expansions

$$G = \sqrt{\kappa}|D|^{\frac{1}{2}} + G_{-3/2}\,, \qquad G^{-1} = |D|^{\frac{1}{2}}(1-\kappa\partial_{xx})^{-\frac{1}{2}} + \pi_0 = \frac{1}{\sqrt{\kappa}}|D|^{-\frac{1}{2}} + G_{-5/2}\,, \tag{6.55}$$

where $G_{-3/2} = \mathrm{Op}(g_{-\frac{3}{2}}) \in OPS^{-3/2}$ and $G_{-5/2} = \mathrm{Op}(g_{-\frac{5}{2}}) \in OPS^{-5/2}$. Using (6.50), (6.51), (2.25), (6.55) we get

$$|D|G = \mathrm{Op}(\chi(\xi)|\xi|g(\xi)) = T(D)\,, \quad G^{-1}(1-\kappa\partial_{xx}) = T(D) + \pi_0 \tag{6.56}$$

where $T(D)$ is the Fourier multiplier

$$T := T(D) := |D|^{1/2}(1-\kappa\partial_{xx})^{1/2} = \mathrm{Op}\big(\chi(\xi)|\xi|^{\frac{1}{2}}(1+\kappa\xi^2)^{\frac{1}{2}}\big) \in OPS^{3/2}\,. \tag{6.57}$$

Hence using (6.55)-(6.56) (and renaming ∂_y as ∂_x) we get

$$\begin{aligned} \mathcal{L}_3 &\overset{(6.38)}{:=} \mathcal{S}^{-1}\mathcal{L}_2\mathcal{S} \overset{(6.34),(6.28)}{=} \mathcal{S}^{-1}\mathcal{Q}^{-1}\mathcal{B}^{-1}\mathcal{L}_0\mathcal{B}\mathcal{Q}\mathcal{S} = \\ &= \omega\cdot\partial_\varphi \mathbb{I}_2 + \begin{pmatrix} a_1\partial_x + a_2 & -m_3(\varphi)T(D) + \sqrt{\kappa}\,a_7\mathcal{H}|D|^{\frac{1}{2}} + \mathcal{R}_{3,B} \\ m_3(\varphi)T(D) - \frac{a_8}{\sqrt{\kappa}}|D|^{\frac{1}{2}}\mathcal{H} + m_3(\varphi)\pi_0 + \mathcal{R}_{3,C} & a_1\partial_x + \mathcal{R}_{3,D} \end{pmatrix} \end{aligned} \tag{6.58}$$

where the remainders are the pseudo-differential operators in OPS^0

$$\mathcal{R}_{3,B} := a_7\mathcal{H}G_{-3/2} + \mathcal{R}_2\Lambda\,, \quad \mathcal{R}_{3,D} := [G^{-1}, a_1]\partial_x G + G^{-1}b_{10}G\,, \tag{6.59}$$

$$\mathcal{R}_{3,C} := a_8 G_{-5/2}\partial_x + [G^{-1}, a_8]\partial_x + G^{-1}b_9\,. \tag{6.60}$$

Lemma 6.7. *Each $\mathcal{R} = \mathcal{R}_{3,B}, \mathcal{R}_{3,C}, \mathcal{R}_{3,D}$ is in OPS^0 and satisfy, for all $s_0 \leq s \leq S$,*

$$|\mathcal{R}|_{0,s,\alpha}^{k_0,\gamma} \lesssim_{S,\alpha} \varepsilon(1 + \|\mathfrak{I}_0\|_{s+\sigma+\alpha}^{k_0,\gamma})\,, \quad |\partial_i\mathcal{R}[\hat{\imath}]|_{0,s_1,\alpha} \lesssim_{S,\alpha} \varepsilon\|\hat{\imath}\|_{s_1+\sigma+\alpha} \tag{6.61}$$

for some $\sigma := \sigma(\tau,\nu) > 0$. The real operator $\mathcal{L}_3$ is even and reversible.

Proof. Use Lemma 2.8 to estimate the commutators in (6.59)-(6.60). □

6.3 Complex variables

We now write the real operator $\mathcal{L}_3$ in (6.58), which acts on the real variables $(\eta,\psi)\in\mathbb{R}^2$, as an operator acting on the complex variables (see (2.16))

$$h := \eta + \mathrm{i}\psi\,, \quad \bar h := \eta - \mathrm{i}\psi\,, \qquad i.e.\ \ \eta = (h+\bar h)/2\,, \quad \psi = (h-\bar h)/(2\mathrm{i})\,.$$

By (2.17) we get the real, even and reversible operator (for simplicity of notation we still denote it by $\mathcal{L}_3$)

$$\mathcal{L}_3 = \omega\cdot\partial_\varphi \mathbb{I}_2 + \mathrm{i}m_3(\varphi)\mathbf{T}(D) + \mathbf{A}_1(\varphi,x)\partial_x + \mathrm{i}(\mathbf{A}_0^{(I)}(\varphi,x) + \mathbf{A}_0^{(II)}(\varphi,x))\mathcal{H}|D|^{\frac12} + \mathrm{i}m_3(\varphi)\Pi_0 + \mathbf{R}_3^{(I)} + \mathbf{R}_3^{(II)} \tag{6.62}$$

where

$$\mathbf{T} := \mathbf{T}(D) := \begin{pmatrix} T(D) & 0 \\ 0 & -T(D)\end{pmatrix}, \quad \mathbf{A}_1(\varphi,x) := \begin{pmatrix} a_1(\varphi,x) & 0 \\ 0 & a_1(\varphi,x)\end{pmatrix}, \tag{6.63}$$

$$\mathbf{A}_0^{(I)}(\varphi,x) := \begin{pmatrix} a_9 & 0 \\ 0 & -a_9\end{pmatrix}, \quad a_9 := -\frac12\Big(\sqrt{\kappa}a_7 + \frac{a_8}{\sqrt{\kappa}}\Big), \tag{6.64}$$

$$\mathbf{A}_0^{(II)}(\varphi,x) := \begin{pmatrix} 0 & a_{10} \\ -a_{10} & 0\end{pmatrix}, \quad a_{10} := \frac12\Big(\sqrt{\kappa}a_7 - \frac{a_8}{\sqrt{\kappa}}\Big), \tag{6.65}$$

$$\Pi_0 := \frac12\begin{pmatrix} 1 & 1 \\ -1 & -1\end{pmatrix}\pi_0\,, \tag{6.66}$$

$$\mathbf{R}_3^{(I)} := \begin{pmatrix} r_3^{(I)}(x,D) & 0 \\ 0 & \overline{r_3^{(I)}(x,D)}\end{pmatrix} \in OPS^0\,, \quad r_3^{(I)}(x,D) := \frac12\big(a_2 + \mathcal{R}_{3,D} - \mathrm{i}\mathcal{R}_{3,B} + \mathrm{i}\mathcal{R}_{3,C}\big)\,,$$

$$\mathbf{R}_3^{(II)} := \begin{pmatrix} 0 & r_3^{(II)}(x,D) \\ \overline{r_3^{(II)}(x,D)} & 0\end{pmatrix} \in OPS^0\,, \quad r_3^{(II)}(x,D) := \frac12\big(a_2 - \mathcal{R}_{3,D} + \mathrm{i}\mathcal{R}_{3,B} + \mathrm{i}\mathcal{R}_{3,C}\big)\,.$$

Lemma 6.6 and (6.61) imply for all $s_0 \le s \le S$, the estimates

$$|r_3^{(I)}(x,D)|_{0,s,\alpha}^{k_0,\gamma},\ |r_3^{(II)}(x,D)|_{0,s,\alpha}^{k_0,\gamma} \lesssim_{S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\alpha+\sigma}^{k_0,\gamma}\big)\,, \tag{6.67}$$

$$|\partial_i r_3^{(I)}(x,D)[\hat\imath]|_{0,s_1,\alpha},\ |\partial_i r_3^{(II)}(x,D)[\hat\imath]|_{0,s_1,\alpha} \lesssim_{S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\alpha+\sigma}\,. \tag{6.68}$$

Note that $\mathcal{L}_3$ in (6.62) is block-diagonal (in $(u,\bar u)$) up to order $|D|^{1/2}$. The introduction of the complex formulation is convenient in section 6.5 where we eliminate iteratively the off-diagonal terms of $\mathcal{L}_3$ up to very smoothing remainders, see Proposition 6.11.

In the next sections we reduce the real, even and reversible operator $\mathcal{L}_3$ neglecting the term $\mathrm{i}m_3(\varphi)\Pi_0$ in (6.66). For simplicity of notation we denote it as $\mathcal{L}_3$ as well. The projector $m_3(\varphi)\mathrm{i}\Pi_0$ transforms under conjugation into a finite dimensional operator and we will conjugate it only once in section 6.8.

6.4 Time-reduction of the highest order

The purpose of this section is to remove the dependence on φ from the highest order term $\mathrm{i}m_3(\varphi)\mathbf{T}(D)$ in the operator $\mathcal{L}_3$ defined in (6.62) (without Π_0). Actually, since we only assume that the frequency ω belongs to $\mathtt{DC}^\gamma_{K_n}$ defined in (1.40), we shall only transform $\mathrm{i}\Pi_{K_n} m_3(\varphi)\mathbf{T}(D)$ (where K_n is defined in (1.39)) into a constant coefficient operator, and we keep the term (6.80) which is Fourier supported on the high harmonics, and thus contributes to (7.94)-(7.95).

To this aim we perform a quasi periodic reparametrization of time

$$\vartheta := \varphi + \omega p(\varphi) \quad \Leftrightarrow \quad \varphi = \vartheta + \omega\tilde p(\vartheta) \tag{6.69}$$

where $p(\varphi)$ is a small periodic function to be determined. We conjugate $\mathcal{L}_3$ by the real operator

$$\mathcal{P}\,\mathbb{I}_2 = \begin{pmatrix} \mathcal{P} & 0 \\ 0 & \mathcal{P}\end{pmatrix} \quad \text{where} \quad (\mathcal{P}h)(\varphi,x) := h(\varphi+\omega p(\varphi),x)\,, \quad (\mathcal{P}^{-1}h)(\vartheta,x) := h(\vartheta+\omega\tilde p(\vartheta),x)\,.$$

The differential operator $\omega\cdot\partial_\varphi$ and the multiplication operator by a transform into

$$\mathcal{P}^{-1}\omega\cdot\partial_\varphi\mathcal{P} = \rho(\vartheta)\omega\cdot\partial_\varphi\,, \quad \rho(\vartheta) := \big(\mathcal{P}^{-1}[1+\omega\cdot\partial_\varphi p]\big)\,, \quad \mathcal{P}^{-1}a\mathcal{P} = (\mathcal{P}^{-1}a)\,, \tag{6.70}$$

while a space Fourier multiplier $\phi(D)$ remains clearly unchanged $\mathcal{P}^{-1}\phi(D)\mathcal{P} = \phi(D)$. Thus

$$\begin{aligned}(\mathcal{P}^{-1}\,\mathbb{I}_2)\mathcal{L}_3\,(\mathcal{P}\,\mathbb{I}_2) &= (\mathcal{P}^{-1}[1+\omega\cdot\partial_\varphi p])\omega\cdot\partial_\varphi\mathbb{I}_2 + \big(\mathcal{P}^{-1}m_3\big)\mathrm{i}\mathbf{T}(D) + (\mathcal{P}^{-1}\mathbb{I}_2\mathbf{A}_1)\partial_x \\ &\quad + \mathrm{i}(\mathcal{P}^{-1}\mathbb{I}_2)(\mathbf{A}_0^{(I)} + \mathbf{A}_0^{(II)})\mathcal{H}|D|^{\frac12} + (\mathcal{P}^{-1}\mathbb{I}_2)\big(\mathbf{R}_3^{(I)} + \mathbf{R}_3^{(II)}\big)(\mathcal{P}\mathbb{I}_2)\,.\end{aligned}$$

Splitting $m_3(\varphi) = \Pi_{K_n}m_3(\varphi) + \Pi_{K_n}^\perp m_3(\varphi)$ we solve, for all $\omega \in \mathtt{DC}_{K_n}^\gamma$ (see (1.40)), the equation

$$1 + \omega\cdot\partial_\varphi p = \mathtt{m}_3^{-1}\Pi_{K_n}m_3(\varphi)\,, \tag{6.71}$$

by defining (the function $m_3(\varphi)$ is even)

$$\mathtt{m}_3 := (2\pi)^{-\nu}\int_{\mathbb{T}^\nu}\Pi_{K_n}m_3(\varphi)\,d\varphi \overset{(6.37)}{=} (2\pi)^{-\nu}\int_{\mathbb{T}^\nu}\Big(\frac{1}{2\pi}\int_{\mathbb{T}}\sqrt{1+\eta_x^2}\,dx\Big)^{-3/2}d\varphi, \tag{6.72}$$

$$p := (\omega\cdot\partial_\varphi)^{-1}\big(\mathtt{m}_3^{-1}\Pi_{K_n}m_3(\varphi) - 1\big) \text{ which is odd in } \varphi\,. \tag{6.73}$$

Dividing $(\mathcal{P}^{-1}\,\mathbb{I}_2)\mathcal{L}_3\,(\mathcal{P}\,\mathbb{I}_2)$ by the even function $\rho := \mathcal{P}^{-1}[1+\omega\cdot\partial_\varphi p]$ we get the real, even and reversible operator

$$\begin{aligned}\mathcal{L}_4 := \rho^{-1}(\mathcal{P}^{-1}\,\mathbb{I}_2)\mathcal{L}_3\,(\mathcal{P}\,\mathbb{I}_2) &= \omega\cdot\partial_\varphi\mathbb{I}_2 + \mathrm{i}\mathtt{m}_3\mathbf{T}(D) + \mathbf{B}_1(\varphi,x)\partial_x + \mathrm{i}\big(\mathbf{B}_0^{(I)}(\varphi,x) + \mathbf{B}_0^{(II)}(\varphi,x)\big)\mathcal{H}|D|^{\frac12} \\ &\quad + \mathbf{R}_4^{(I)} + \mathbf{R}_4^{(II)} + \mathbf{R}_4^\perp\end{aligned} \tag{6.74}$$

where

$$\mathbf{B}_1 := \rho^{-1}\mathcal{P}^{-1}\mathbb{I}_2\mathbf{A}_1 = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{11}\end{pmatrix}\,, \quad a_{11} := \rho^{-1}\mathcal{P}^{-1}(a_1) \tag{6.75}$$

$$\mathbf{B}_0^{(I)} := \rho^{-1}\mathcal{P}^{-1}\mathbb{I}_2\mathbf{A}_0^{(I)} = \begin{pmatrix} a_{12} & 0 \\ 0 & -a_{12}\end{pmatrix}\,, \quad a_{12} := \rho^{-1}\mathcal{P}^{-1}(a_9) \tag{6.76}$$

$$\mathbf{B}_0^{(II)} := \rho^{-1}\mathcal{P}^{-1}\mathbb{I}_2\mathbf{A}_0^{(II)} = \begin{pmatrix} 0 & \rho^{-1}\mathcal{P}^{-1}(a_{10}) \\ -\rho^{-1}\mathcal{P}^{-1}(a_{10}) & 0\end{pmatrix} \tag{6.77}$$

$$\mathbf{R}_4^{(I)} := \begin{pmatrix} r_4^{(I)}(x,D) & 0 \\ 0 & \overline{r_4^{(I)}(x,D)}\end{pmatrix}\,, \quad r_4^{(I)}(x,D) := \rho^{-1}\mathcal{P}^{-1}r_3^{(I)}(x,D)\mathcal{P}\,, \tag{6.78}$$

$$\mathbf{R}_4^{(II)} := \begin{pmatrix} 0 & r_4^{(II)}(x,D) \\ \overline{r_4^{(II)}(x,D)} & 0\end{pmatrix}\,, \quad r_4^{(II)}(x,D) := \rho^{-1}\mathcal{P}^{-1}r_3^{(II)}(x,D)\mathcal{P} \tag{6.79}$$

and

$$\mathbf{R}_4^\perp := \mathrm{i}\rho^{-1}\Pi_{K_n}^\perp m_3(\varphi)\mathbf{T}(D)\,. \tag{6.80}$$

Lemma 6.8. *The maps $\mathcal{P}$, $\mathcal{P}^{-1}$ are $\mathcal{D}^{k_0}$-(k_0+1)-tame with tame constants satisfying the estimates*

$$\mathfrak{M}_{\mathcal{P}^{\pm1}}(s) \leq_S (1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})\,, \quad \forall s_0 \leq s \leq S\,. \tag{6.81}$$

The maps $\mathcal{P} - \mathrm{Id}$, $\mathcal{P}^{-1} - \mathrm{Id}$ are $\mathcal{D}^{k_0}$-(k_0+2)-tame and

$$\mathfrak{M}_{\mathcal{P}^{\pm1}-\mathrm{Id}}(s) \leq_S \varepsilon\gamma^{-1}(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})\,, \quad \forall s_0 \leq s \leq S\,. \tag{6.82}$$

The coefficient $\mathtt{m}_3$ defined in (6.72) *and the functions a_{11}, a_{12}, $\rho^{-1}\mathcal{P}^{-1}(a_{10})$ in* (6.75)-(6.77) *satisfy*

$$|\mathtt{m}_3 - 1|^{k_0,\gamma} \leq C\varepsilon\,, \quad |\partial_i\mathtt{m}_3[\hat\imath]| \leq C\varepsilon\|\hat\imath\|_\sigma\,, \tag{6.83}$$

$$\|a_{11}\|_s^{k_0,\gamma}, \|a_{12}\|_s^{k_0,\gamma}, \|\rho^{-1}\mathcal{P}^{-1}(a_{10})\|_s^{k_0,\gamma} \leq_S \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,,\ \forall s_0 \leq s \leq S\,, \tag{6.84}$$

and

$$
\begin{aligned}
&|r_4^{(I)}(x,D)|_{0,s,\alpha}^{k_0,\gamma}, |r_4^{(II)}(x,D)|_{0,s,\alpha}^{k_0,\gamma} \leq_{S,\alpha} \varepsilon(1+\|\mathfrak{I}_0\|_{s+\alpha+\sigma}^{k_0,\gamma}) && (6.85)\\
&\|(\partial_i \mathcal{P}^{\pm 1}[\hat\imath])h\|_{s_1} \leq_S \varepsilon\gamma^{-1}\|\hat\imath\|_{s_1+\sigma}\|h\|_{s_1+\sigma} && (6.86)\\
&\|\partial_i a_{11}[\hat\imath]\|_{s_1}, \|\partial_i a_{12}[\hat\imath]\|_{s_1}\|\partial_i\{\rho^{-1}\mathcal{P}^{-1}(a_{10})\}[\hat\imath]\|_{s_1} \leq_S \varepsilon\|\hat\imath\|_{s_1+\sigma} && (6.87)\\
&|\partial_i r_4^{(I)}(x,D)[\hat\imath]|_{0,s_1,\alpha}, |\partial_i r_4^{(II)}(x,D)[\hat\imath]|_{0,s_1,\alpha} \leq_{S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\alpha+\sigma}\,. && (6.88)
\end{aligned}
$$

Proof. The estimates (6.81), (6.84) follow by Lemmata 2.21, 2.14 and 6.6. The bound (6.82) follows since

$$
(\mathcal{P}-\mathrm{Id})h = p\int_0^1 \mathcal{P}_\tau[\omega\cdot\partial_\varphi h]\,d\tau\,, \quad \mathcal{P}_\tau[h](\varphi,x) := h(\varphi+\tau\omega p(\varphi),x)\,,
$$

and since by Lemma 6.6, using (6.73) and (6.43), (6.84), we have

$$
\|p\|_s^{k_0,\gamma} \leq_s \varepsilon\gamma^{-1}(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})\,. \tag{6.89}
$$

The estimate for $\mathcal{P}^{-1}-\mathrm{Id}$ follows similarly. Let us prove (6.85). The conjugated operator

$$
\mathcal{P}^{-1}r_3^{(I)}(x,D)\mathcal{P} = \mathrm{Op}(\tilde r_3) \qquad \text{where} \qquad \tilde r_3(\vartheta,x,\xi) := r_3^{(I)}(\vartheta+\omega\tilde p(\vartheta),x,\xi)\,. \tag{6.90}
$$

Hence for all $\alpha\geq 0$, for all $|k|\leq k_0$, for all $\xi\in\mathbb{R}$ and for all ω we have by Lemma 2.21

$$
\|\partial_\xi^\alpha \tilde r(\omega,\cdot,\xi)\|_s^{k_0,\gamma} \leq_S \|\partial_\xi^\alpha r_3^{(I)}(\cdot,\xi)\|_{s+k_0}^{k_0,\gamma} + \|p\|_{s+\sigma}^{k_0,\gamma}\|\partial_\xi^\alpha r_3^{(I)}(\cdot,\xi)\|_{s_0+k_0}^{k_0,\gamma}\,,
$$

thus using the estimate (6.89) we get

$$
|\mathcal{P}^{-1}r_3^{(I)}(x,D)\mathcal{P}|_{0,s,\alpha}^{k_0,\gamma} \leq_S |r_3^{(I)}(x,D)|_{0,s,\alpha}^{k_0,\gamma} + \|\mathfrak{I}\|_{s+\sigma}^{k_0,\gamma}|r_3^{(I)}(x,D)|_{0,s_0,\alpha}^{k_0,\gamma} \overset{(6.67)}{\leq_{S,\alpha}} \varepsilon\big(1+\|\mathfrak{I}_0\|_{s+\alpha+\sigma}^{k_0,\gamma}\big)\,,
$$

and the estimate (6.85) for $r_4^{(I)}$ follows. The estimate for $r_4^{(II)}$ is analogous. The proof of (6.86) is similar to the proof of the estimate for $\partial_i\mathcal{B}^{\pm1}$ in Lemma 6.6. The estimate (6.87) follows by differentiating the explicit expressions in (6.72), (6.75)-(6.77), using (6.81), (6.86), the estimates of Lemma 6.6 and (2.72). The estimate (6.88) follows since by (6.90) $\partial_i\mathrm{Op}(\tilde r_3)[\hat\imath] = \partial_i\tilde p[\hat\imath]\mathrm{Op}\big(\partial_\varphi r_3^{(I)}(\vartheta+\omega\tilde p(\vartheta),x,\xi)\big)$. □

In the next sections we reduce the real, even and reversible operator $\mathcal{L}_4$ neglecting the term $\mathbf{R}_4^\perp$ (for simplicity of notation we denote it in the same way). Note that the term $\mathbf{R}_4^\perp$ is in $OPS^{3/2}$. However it is supported on the high Fourier frequencies and it will contribute to remainders in (7.94)-(7.95). In other words, these terms do not need to be treated in the KAM reducibility scheme of section 7 and the estimates (7.94)-(7.95) are yet sufficient for the convergence of Nash-Moser scheme of section 8.

6.5 Block-decoupling up to smoothing remainders

The goal of this section is to conjugate the operator $\mathcal{L}_4$ in (6.74) (without $\mathbf{R}_4^\perp$) to the operator $\mathcal{L}_M$ in (6.120) which is block-diagonal up to the smoothing remainder $\mathbf{R}_M^{(II)}\in OPS^{\frac12-M}$. This is achieved by applying iteratively M-times a conjugation map which transforms the off-diagonal block operators into 1-smoother ones.

We describe the generic inductive step. We have a real, even and reversible operator

$$
\mathcal{L}_n := \omega\cdot\partial_\varphi\mathbb{I}_2 + \mathrm{i}\mathtt{m}_3\mathbf{T}(D) + \mathbf{B}_1\partial_x + \mathrm{i}\mathbf{B}_0^{(I)}\mathcal{H}|D|^{\frac12} + \mathbf{R}_n^{(I)} + \mathbf{R}_n^{(II)} \tag{6.91}
$$

with block-diagonal terms

$$
\mathbf{R}_n^{(I)} := \begin{pmatrix} r_n^{(I)}(x,D) & 0 \\ 0 & \overline{r_n^{(I)}(x,D)}\end{pmatrix}\,, \quad r_n^{(I)}(x,D)\in OPS^0\,, \tag{6.92}
$$

and smoothing off-diagonal remainders

$$\mathbf{R}_n^{(II)} := \begin{pmatrix} 0 & r_n^{(II)}(x,D) \\ \overline{r_n^{(II)}(x,D)} & 0 \end{pmatrix}, \quad r_n^{(II)}(x,D) \in OPS^{\frac{1}{2}-n}, \tag{6.93}$$

which satisfy

$$|\!|\mathbf{R}_n^{(I)}|\!|_{0,s,\alpha}^{k_0,\gamma} + |\!|\mathbf{R}_n^{(II)}|\!|_{-n+\frac{1}{2},s,\alpha}^{k_0,\gamma} \leq_{n,S,\alpha} \varepsilon(1 + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_n(\alpha)}^{k_0,\gamma}), \quad \forall s_0 \leq s \leq S, \tag{6.94}$$

$$|\!|\partial_i \mathbf{R}_n^{(I)}[\hat{\imath}]|\!|_{0,s_1,\alpha} + |\!|\partial_i \mathbf{R}_n^{(II)}[\hat{\imath}]|\!|_{-n+\frac{1}{2},s_1,\alpha} \leq_{n,S,\alpha} \varepsilon \|\hat{\imath}\|_{s_1+\sigma+\aleph_n(\alpha)}, \tag{6.95}$$

where the increasing constants $\aleph_n(\alpha)$ are defined inductively by

$$\aleph_0(\alpha) := \alpha, \qquad \aleph_{n+1}(\alpha) := \aleph_n(\alpha+1) + n + 2\alpha + 4. \tag{6.96}$$

Initialization. The real, even and reversible operator $\mathcal{L}_4$ in (6.74) satisfies the assumptions (6.91)-(6.95) where the off diagonal remainder is $\mathrm{i}\mathbf{B}_0^{(II)}(\varphi,x)\mathcal{H}|D|^{\frac{1}{2}} + \mathbf{R}_4^{(II)} \in OPS^{1/2}$ (recall that we have neglected $\mathbf{R}_4^{\perp}$).

Inductive step. We conjugate $\mathcal{L}_n$ in (6.91) by a real operator of the form

$$\Phi_n := \mathbb{I}_2 + \Psi_n, \qquad \Psi_n := \begin{pmatrix} 0 & \psi_n(x,D) \\ \overline{\psi_n(x,D)} & 0 \end{pmatrix}, \quad \psi_n(x,D) \in OPS^{-n-1}. \tag{6.97}$$

We compute

$$\begin{aligned} \mathcal{L}_n \Phi_n = \Phi_n\big(&\omega\cdot\partial_\varphi \mathbb{I}_2 + \mathrm{i}\mathtt{m}_3 \mathbf{T}(D) + \mathbf{B}_1\partial_x + \mathrm{i}\mathbf{B}_0^{(I)}\mathcal{H}|D|^{\frac{1}{2}} + \mathbf{R}_n^{(I)}\big) \\ &+ [\mathrm{i}\mathtt{m}_3 \mathbf{T}(D) + \mathbf{B}_1\partial_x + \mathrm{i}\mathbf{B}_0^{(I)}\mathcal{H}|D|^{\frac{1}{2}} + \mathbf{R}_n^{(I)}, \Psi_n] + \omega\cdot\partial_\varphi \Psi_n + \mathbf{R}_n^{(II)} + \mathbf{R}_n^{(II)}\Psi_n. \end{aligned} \tag{6.98}$$

By (6.63) and (6.97) the vector valued commutator

$$\mathrm{i}[\mathtt{m}_3 \mathbf{T}(D), \Psi_n] = \mathrm{i}\mathtt{m}_3 \begin{pmatrix} 0 & T(D)\psi_n(x,D) + \psi_n(x,D)T(D) \\ -\big(T(D)\overline{\psi_n(x,D)} + \overline{\psi_n(x,D)}T(D)\big) & \end{pmatrix} \tag{6.99}$$

is block off-diagonal.

We define a cut off function $\chi_0 \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$, even, $0 \leq \chi_0 \leq 1$, such that

$$\chi_0(\xi) = \begin{cases} 0 & \text{if} \quad |\xi| \leq \frac{1}{2} \\ 1 & \text{if} \quad |\xi| \geq \frac{3}{4}. \end{cases} \tag{6.100}$$

Lemma 6.9. *Let*

$$\psi_n(x,\xi) := \begin{cases} -\dfrac{\chi_0(\xi) r_n^{(II)}(x,\xi)}{2\mathrm{i}\mathtt{m}_3 T(\xi)} & \text{if} \quad |\xi| > \frac{1}{3}, \\ 0 & \text{if} \quad |\xi| \leq \frac{1}{3}, \end{cases} \qquad \psi_n \in S^{-n-1}. \tag{6.101}$$

Then the operator Ψ_n *in* (6.97) *solves*

$$\mathrm{i}[\mathtt{m}_3 \mathbf{T}(D), \Psi_n] + \mathbf{R}_n^{(II)} = \mathbf{R}_{T,\psi_n} \tag{6.102}$$

where

$$\mathbf{R}_{T,\psi_n} := \mathrm{i}\begin{pmatrix} 0 & r_{T,\psi_n}(x,D) \\ -\overline{r_{T,\psi_n}(x,D)} & \end{pmatrix}, \quad r_{T,\psi_n} \in S^{-n-\frac{1}{2}}, \tag{6.103}$$

satisfies for all $s_0 \leq s \leq S$

$$|\!|r_{T,\psi_n}(x,D)|\!|_{-n-\frac{1}{2},s,\alpha}^{k_0,\gamma} \leq_{S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_n(\alpha)+\alpha+4}^{k_0,\gamma}\big). \tag{6.104}$$

The map Ψ_n is real, even, reversibility preserving and

$$|\psi_n(x,D)|^{k_0,\gamma}_{-n-1,s,\alpha} \leq_{n,S,\alpha} \varepsilon(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_n(\alpha)}), \quad \forall s_0 \leq s \leq S, \tag{6.105}$$

$$|\partial_i\psi_n(x,D)[\hat\imath]|_{-n-1,s_1,\alpha} \leq_{n,S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\aleph_n(\alpha)}, \tag{6.106}$$

$$|\partial_i r_{T,\psi_n}(x,D)[\hat\imath]|_{-n-\frac12,s_1,\alpha} \leq_{n,S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\aleph_n(\alpha)+\alpha+4}. \tag{6.107}$$

Proof. By (6.99) and (6.93), in order to solve (6.102) with a remainder $\mathbf{R}_{T,\psi_n} \in OPS^{-n-\frac12}$ as in (6.103), we have to solve the equation

$$\mathtt{im}_3(T(D)\psi_n(x,D)+\psi_n(x,D)T(D)) + r_n^{(II)}(x,D) = r_{T,\psi_n}(x,D) \in OPS^{-n-\frac12}. \tag{6.108}$$

By (2.29), (2.30) (applied with $N=1$), we have

$$T(D)\psi_n(x,D)+\psi_n(x,D)T(D) = \mathrm{Op}(2T(\xi)\psi_n(x,\xi))+\mathrm{Op}(\mathfrak{r}_{T,\psi_n}(x,\xi)) \quad \text{where} \quad \mathfrak{r}_{T,\psi_n} \in S^{-n-\frac12} \tag{6.109}$$

because $T(\xi) \in S^{3/2}$ and $\psi_n(x,\xi) \in S^{-n-1}$. The symbol $\psi_n(x,\xi)$ in (6.101) is the solution of

$$2\mathtt{im}_3 T(\xi)\psi_n(x,\xi) + \chi_0(\xi) r_n^{(II)}(x,\xi) = 0 \tag{6.110}$$

where the cut-off χ_0 is defined in (6.100). Note that $T(\xi)=0$ for all $|\xi| \leq 1/3$ (see (6.57), (2.26)) and that is why we do not include in (6.110) the symbol $(1-\chi_0(\xi))r_n^{(II)}(x,\xi) \in S^{-\infty}$. Note also that $|T(\xi)| \geq c > 0$ for all $|\xi| \geq 1/2$. By (6.101) and Lemma 2.7 and (6.94), we have, for all $s_0 \leq s \leq S$,

$$|\psi_n(x,D)|^{k_0,\gamma}_{-n-1,s,\alpha} \leq_{n,\alpha} |\mathbf{R}_n^{(II)}|^{k_0,\gamma}_{-n+\frac12,s,\alpha} \leq_{n,S,\alpha} \varepsilon(1+\|\mathfrak{I}_0\|_{s+\sigma+\aleph_n(\alpha)})$$

proving (6.105). By (6.109) and (6.110) the remainder $r_{T,\psi_n}(x,\xi)$ in (6.108) is

$$r_{T,\psi_n}(x,\xi) = \mathtt{im}_3\mathfrak{r}_{T,\psi_n}(x,\xi) + (1-\chi_0(\xi))r_n^{(II)}(x,\xi) \in S^{-n-\frac12}. \tag{6.111}$$

By (2.42) (applied with $A=T(D)$, $B=\psi_n(x,D)$, $N=1$, $m=3/2$, $m'=-n-1$) we have

$$|\mathfrak{r}_{T,\psi_n}(x,D)|^{k_0,\gamma}_{-n-\frac12,s,\alpha} \leq_{n,s,\alpha} |\mathbf{R}_n^{(II)}|^{k_0,\gamma}_{-n+\frac12,s+2+\frac32+\alpha,\alpha} \overset{(6.94)}{\leq_{n,S,\alpha}} \varepsilon\big(1+\|\mathfrak{I}_0\|_{s+\sigma+\aleph_n(\alpha)+\alpha+4}\big) \tag{6.112}$$

and the estimate (6.104) for $r_{T,\psi_n}(x,D)$ follows by (6.111) using also (6.83), (6.94). The bound (6.106) is obtained differentiating the symbol (6.101) and using (6.83), (6.94), (6.95). Let us prove the estimate (6.107). By differentiating (6.111) with respect to i we get

$$\partial_i r_{T,\psi_n}(x,\xi)[\hat\imath] := \mathrm{i}\partial_i\mathtt{m}_3[\hat\imath]\mathfrak{r}_{T,\psi_n}(x,\xi) + \mathtt{im}_3\partial_i\mathfrak{r}_{T,\psi_n}(x,\xi)[\hat\imath] + (1-\chi_0(\xi))\partial_i r_n^{(II)}(x,\xi)[\hat\imath]. \tag{6.113}$$

Note that, since $T(\xi)$ does not depend on i, by formulae (2.29), (2.30) (with $A=T(D)$, $B=\psi_n(x,D)$, $N=1$), we get $\partial_i\mathfrak{r}_{T,\psi_n}(x,D)[\hat\imath] = \mathfrak{r}_{T,\partial_i\psi_n[\hat\imath]}(x,D)$ and hence by (2.42) (for $A=T(D)$, $B=\partial_i\psi_n(x,D)[\hat\imath]$, $N=1$, $m=3/2$, $m'=-n-1$) we get

$$|\partial_i\mathfrak{r}_{T,\psi_n}(x,D)[\hat\imath]|_{-n-\frac12,s_1,\alpha} \leq_{n,S,\alpha} |\partial_i\psi_n(x,D)[\hat\imath]|_{-n-1,s_1+2+\frac32+\alpha,\alpha} \overset{(6.106)}{\leq_{n,S,\alpha}} \varepsilon\|\hat\imath\|_{s_1+\sigma+\aleph_n(\alpha)+\alpha+4}.$$

The estimate (6.107) for $\partial_i r_{T,\psi_n}(x,D)[\hat\imath]$ then follows by recalling (6.113) and (6.83), (6.95), (6.112).

Finally, using Lemma 2.3 and Lemma 2.4 we see that the map Ψ_n defined by the symbol (6.101) is even and reversibility preserving because r_n is even and reversible. □

By (6.98) and (6.103) the conjugated operator is

$$\mathcal{L}_{n+1} := \Phi_n^{-1}\mathcal{L}_n\Phi_n = \omega\cdot\partial_\varphi\mathbb{I}_2 + \mathtt{im}_3\mathbf{T}(D) + \mathbf{B}_1\partial_x + \mathrm{i}\mathbf{B}_0^{(I)}\mathcal{H}|D|^{\frac12} + \mathbf{R}_n^{(I)} + \mathbf{R}_{n+1} \tag{6.114}$$

where $\mathbf{R}_{n+1} := \Phi_n^{-1}\mathbf{R}^*_{n+1}$ and

$$\mathbf{R}^*_{n+1} := \mathbf{R}_{T,\psi_n} + [\mathbf{B}_1\partial_x,\Psi_n] + \mathrm{i}[\mathbf{B}_0^{(I)}\mathcal{H}|D|^{\frac12},\Psi_n] + [\mathbf{R}_n^{(I)},\Psi_n] + \omega\cdot\partial_\varphi\Psi_n + \mathbf{R}_n^{(II)}\Psi_n. \tag{6.115}$$

Note that $\mathbf{R}_{n+1}$ is the only operator in (6.114) containing off-diagonal terms.

Lemma 6.10. *The operator $\mathbf{R}_{n+1} \in OPS^{-n-\frac{1}{2}}$ satisfies*

$$|\mathbf{R}_{n+1}|^{k_0,\gamma}_{-n-\frac{1}{2},s,\alpha} \leq_{n,S,\alpha} \varepsilon(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_{n+1}(\alpha)}), \quad \forall s_0 \leq s \leq S\,, \tag{6.116}$$

$$|\partial_i \mathbf{R}_{n+1}[\hat{\imath}]|_{-n-\frac{1}{2},s_1,\alpha} \leq_{n,S,\alpha} \varepsilon\|\hat{\imath}\|_{s_1+\sigma+\aleph_{n+1}(\alpha)} \tag{6.117}$$

where the constant $\aleph_{n+1}(\alpha)$ is defined in (6.96)*.*

Proof. Proof of (6.116). We first estimate separately all the terms of $\mathbf{R}^*_{n+1}$ in (6.115). The operator $\mathbf{R}_{T,\psi_n} \in OPS^{-n-\frac{1}{2}}$ in (6.103) satisfies (6.104). By (6.75) and since $\psi_n(x,D) \in OPS^{-n-1}$, see (6.101), we have

$$[\mathbf{B}_1\partial_x, \Psi_n] = \begin{pmatrix} 0 & [a_{11}\partial_x, \psi_n(x,D)] \\ \overline{[a_{11}\partial_x, \psi_n(x,D)]} & 0 \end{pmatrix} \in OPS^{-n-1} \subset OPS^{-n-\frac{1}{2}}\,.$$

Moreover Lemma 2.8 (with $m = 1$, $m' = -n-1$) implies

$$\begin{aligned}|[a_{11}\partial_x, \psi_n(x,D)]|^{k_0,\gamma}_{-n-\frac{1}{2},s,\alpha} &\leq |[a_{11}\partial_x, \psi_n(x,D)]|^{k_0,\gamma}_{-n-1,s,\alpha} \leq_{n,s,\alpha} \|a_{11}\|^{k_0,\gamma}_{s+n+3+\alpha}|\psi_n(x,D)|^{k_0,\gamma}_{-n-1,s_0+3+\alpha,\alpha+1} \\ &\quad + \|a_{11}\|^{k_0,\gamma}_{s_0+n+3+\alpha}|\psi_n(x,D)|^{k_0,\gamma}_{-n-1,s+3+\alpha,\alpha+1} \\ &\overset{(6.84),(6.105),(6.10)}{\leq_{n,S,\alpha}} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_n(\alpha+1)+n+\alpha+3}\big)\,.\end{aligned}$$

We also claim that $[\mathbf{B}_0^{(I)}\mathcal{H}|D|^{\frac{1}{2}}, \Psi_n] \in OPS^{-n-\frac{1}{2}}$. Indeed by (6.76) we have

$$[\mathbf{B}_0^{(I)}\mathcal{H}|D|^{\frac{1}{2}}, \Psi_n] = \begin{pmatrix} 0 & a_{12}\mathcal{H}|D|^{\frac{1}{2}}\overline{\psi_n(x,D)} + \psi_n(x,D)a_{12}\mathcal{H}|D|^{\frac{1}{2}} \\ -a_{12}\mathcal{H}|D|^{\frac{1}{2}}\psi_n(x,D) - \overline{\psi_n(x,D)}a_{12}\mathcal{H}|D|^{\frac{1}{2}} & 0 \end{pmatrix}$$

and (2.41), (6.84), (6.105) imply $|[\mathbf{B}_0^{(I)}\mathcal{H}|D|^{\frac{1}{2}}, \Psi_n]|^{k_0,\gamma}_{-n-\frac{1}{2},s,\alpha} \leq_{n,S,\alpha} \varepsilon(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_n(\alpha)+n+\alpha+1})$. In addition the operator $[\mathbf{R}_n^{(I)}, \Psi_n] \in OPS^{-n-1} \subset OPS^{-n-\frac{1}{2}}$ because (see (6.92), (6.97))

$$[\mathbf{R}_n^{(I)}, \Psi_n] = \begin{pmatrix} 0 & r_n^{(I)}(x,D)\overline{\psi_n(x,D)} - \psi_n(x,D)\overline{r_n^{(I)}(x,D)} \\ \overline{r_n^{(I)}(x,D)}\psi_n(x,D) - \overline{\psi_n(x,D)}r_n^{(I)}(x,D) & 0 \end{pmatrix}$$

and (2.41), (6.94), (6.105) imply $|[\mathbf{R}_n^{(I)}, \Psi_n]|^{k_0,\gamma}_{-n-\frac{1}{2},s,\alpha} \leq |[\mathbf{R}_n^{(I)}, \Psi_n]|^{k_0,\gamma}_{-n-1,s,\alpha} \leq_{n,S,\alpha} \varepsilon\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_n(\alpha)+n+\alpha+1}\big)$. Moreover $\omega\cdot\partial_\varphi\Psi_n \in OPS^{-n-1} \subset OPS^{-n-\frac{1}{2}}$ satisfies

$$|\omega\cdot\partial_\varphi\Psi_n|^{k_0,\gamma}_{-n-\frac{1}{2},s,\alpha} \leq |\omega\cdot\partial_\varphi\Psi_n|^{k_0,\gamma}_{-n-1,s,\alpha} \lessdot |\Psi_n|^{k_0,\gamma}_{-n-1,s+1,\alpha} \leq_{n,S,\alpha} \varepsilon(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_n(\alpha)+1})$$

by (6.105). Finally $\mathbf{R}_n^{(II)}\Psi_n \in OPS^{-2n-\frac{1}{2}} \subset OPS^{-n-\frac{1}{2}}$ and by (2.41) (applied with $m = \frac{1}{2} - n$, $m' = -n-1$), (6.94), (6.105) we have

$$|\mathbf{R}_n^{(II)}\Psi_n|^{k_0,\gamma}_{-n-\frac{1}{2},s,\alpha} \leq |\mathbf{R}_n^{(II)}\Psi_n|^{k_0,\gamma}_{-2n-\frac{1}{2},s,\alpha} \leq_{n,S,\alpha} \varepsilon(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_n(\alpha)+n+\alpha+\frac{1}{2}})\,.$$

Collecting all the previous estimates we deduce that $\mathbf{R}^*_{n+1}$ defined in (6.115) is in $OPS^{-n-\frac{1}{2}}$ and

$$|\mathbf{R}^*_{n+1}|^{k_0,\gamma}_{-n-\frac{1}{2},s,\alpha} \leq_{n,S,\alpha} \varepsilon(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_n(\alpha+1)+n+\alpha+4})\,. \tag{6.118}$$

Now (2.41) (applied with $m = 0$, $m' = -n-\frac{1}{2}$), Lemma 2.10, (6.105), (6.118) imply

$$\begin{aligned}|\mathbf{R}_{n+1}|^{k_0,\gamma}_{-n-\frac{1}{2},s,\alpha} &= |\Phi_n^{-1}\mathbf{R}^*_{n+1}|^{k_0,\gamma}_{-n-\frac{1}{2},s,\alpha} \\ &\leq_{n,s,\alpha} |\Phi_n^{-1}|^{k_0,\gamma}_{0,s,\alpha}|\mathbf{R}^*_{n+1}|^{k_0,\gamma}_{-n-\frac{1}{2},s_0+\alpha,\alpha} + |\Phi_n^{-1}|^{k_0,\gamma}_{0,s_0,\alpha}|\mathbf{R}^*_{n+1}|^{k_0,\gamma}_{-n-\frac{1}{2},s+\alpha,\alpha} \\ &\leq_{n,S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_n(\alpha+1)+n+2\alpha+4}\big)\end{aligned}$$

which is (6.116), recalling (6.96).

PROOF OF (6.117). First we estimate $\partial_i \mathbf{R}^*_{n+1}$ in (6.115). The operator $\partial_i \mathbf{R}_{T,\psi_n}$ satisfies (6.107). Then we have

$$\partial_i[\mathbf{B}_1\partial_x, \Psi_n][\hat{\imath}] = [\partial_i \mathbf{B}_1[\hat{\imath}]\partial_x, \Psi_n] + [\mathbf{B}_1\partial_x, \partial_i \Psi_n[\hat{\imath}]] \, .$$

Hence Lemma 2.8 (with $m = 1$, $m' = -n-1$), the estimates of a_{11} in (6.84), (6.87), (6.105), (6.106), imply

$$|\!|\!|\partial_i[\mathbf{B}_1\partial_x, \Psi_n][\hat{\imath}]|\!|\!|_{-n-\frac{1}{2},s_1,\alpha} \leq |\!|\!|\partial_i[\mathbf{B}_1\partial_x, \Psi_n][\hat{\imath}]|\!|\!|_{-n-1,s_1,\alpha} \leq_{n,S,\alpha} \varepsilon \|\hat{\imath}\|_{s_1+\sigma+\aleph_n(\alpha+1)+n+\alpha+3} \, .$$

The terms $\partial_i[\mathbf{B}_0^{(I)}\mathcal{H}|D|^{\frac{1}{2}}, \Psi_n]$, $\partial_i[\mathbf{R}_n^{(I)}, \Psi_n]$ may be estimated similarly. In addition

$$\begin{aligned} |\!|\!|\partial_i\big(\omega\cdot\partial_\varphi\Psi_n\big)[\hat{\imath}]|\!|\!|_{-n-\frac{1}{2},s_1,\alpha} &\leq |\!|\!|\partial_i\big(\omega\cdot\partial_\varphi\Psi_n\big)[\hat{\imath}]|\!|\!|_{-n-1,s_1,\alpha} \lessdot |\!|\!|\partial_i\Psi_n[\hat{\imath}]|\!|\!|_{-n-1,s_1+1,\alpha} \\ &\overset{(6.106)}{\leq_{n,S,\alpha}} \varepsilon \|\hat{\imath}\|_{s_1+\sigma+\aleph_n(\alpha)+1} \, . \end{aligned}$$

Finally $|\!|\!|\partial_i(\mathbf{R}_n^{(II)}\Psi_n)[\hat{\imath}] \in OPS^{-2n-\frac{1}{2}} \subset OPS^{-n-\frac{1}{2}}$. Hence applying (2.41) with $m = -n+\frac{1}{2}$, $m' = -n-1$, and using (6.94), (6.95), (6.105),(6.106) we get

$$|\!|\!|\partial_i(\mathbf{R}_n^{(II)}\Psi_n)[\hat{\imath}]|\!|\!|_{-n-\frac{1}{2},s_1,\alpha} \leq |\!|\!|\partial_i(\mathbf{R}_n^{(II)}\Psi_n)[\hat{\imath}]|\!|\!|_{-2n-\frac{1}{2},s_1,\alpha} \leq_{n,S,\alpha} \varepsilon \|\hat{\imath}\|_{s_1+\sigma+\aleph_n(\alpha)+n+\alpha+\frac{1}{2}} \, .$$

Collecting the previous bounds we conclude that $|\!|\!|\partial_i \mathbf{R}^*_{n+1}[\hat{\imath}]|\!|\!|_{-n-\frac{1}{2},s_1,\alpha} \leq_{n,S,\alpha} \varepsilon \|\hat{\imath}\|_{s_1+\sigma+\aleph_n(\alpha+1)+n+\alpha+4}$ and the estimate (6.117) follows by

$$\partial_i \mathbf{R}_{n+1}[\hat{\imath}] = \partial_i\big(\Phi_n^{-1}\mathbf{R}^*_{n+1}\big)[\hat{\imath}] = \partial_i\Phi_n^{-1}[\hat{\imath}]\mathbf{R}^*_{n+1} + \Phi_n^{-1}\partial_i\mathbf{R}^*_{n+1}[\hat{\imath}] \quad \text{and} \quad \partial_i\Phi_n^{-1}[\hat{\imath}] = -\Phi_n^{-1}\partial_i\Phi_n[\hat{\imath}]\Phi_n$$

applying (2.41) (with $m = 0$, $m' = -n-\frac{1}{2}$), Lemma 2.10 and the estimates (6.105), (6.106). □

By (6.114) and (6.116)-(6.117) the operator $\mathcal{L}_{n+1}$ has the same form (6.91)-(6.93) with $\mathbf{R}^{(I)}_{n+1}$, $\mathbf{R}^{(II)}_{n+1}$ that satisfy the estimates (6.94)-(6.95) at the step $n+1$. Hence we can repeat iteratively the procedure of Lemmata 6.9 and 6.10. Applying it M-times (M will be fixed in (7.9)) we derive the following proposition.

Proposition 6.11. *The real invertible map $\mathbf{\Phi}_M := \Phi_4 \circ \ldots \circ \Phi_{M+4}$ satisfies the estimate*

$$|\!|\!|\mathbf{\Phi}_M^{\pm 1} - \mathbb{I}_2|\!|\!|^{k_0,\gamma}_{0,s,0} \, , \, |\!|\!|(\mathbf{\Phi}_M^{\pm 1} - \mathbb{I}_2)^*|\!|\!|^{k_0,\gamma}_{0,s,0} \leq_{S,M} \varepsilon(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_M(0)}) \, , \quad \forall s_0 \leq s \leq S \, , \tag{6.119}$$

and conjugate $\mathcal{L}_4$ to the real, even and reversible operator

$$\mathcal{L}_M := \mathbf{\Phi}_M^{-1}\mathcal{L}_4\mathbf{\Phi}_M = \omega\cdot\partial_\varphi\mathbb{I}_2 + \mathrm{i}\mathtt{m}_3\mathbf{T}(D) + \mathbf{B}_1(\varphi,x)\partial_x + \mathrm{i}\mathbf{B}_0^{(I)}(\varphi,x)\mathcal{H}|D|^{\frac{1}{2}} + \mathbf{R}_M^{(I)} + \mathbf{R}_M^{(II)} \tag{6.120}$$

where the remainders

$$\mathbf{R}_M^{(I)} := \begin{pmatrix} r_M^{(I)}(\varphi,x,D) & 0 \\ 0 & \overline{r_M^{(I)}(\varphi,x,D)} \end{pmatrix} \in OPS^0 \, , \quad \mathbf{R}_M^{(II)} := \begin{pmatrix} 0 & \mathcal{R}_M^{(II)} \\ \overline{\mathcal{R}_M^{(II)}} & 0 \end{pmatrix} \in OPS^{\frac{1}{2}-M} \tag{6.121}$$

satisfy the estimates

$$|\!|\!|\mathbf{R}_M^{(I)}|\!|\!|^{k_0,\gamma}_{0,s,\alpha} + |\!|\!|\mathbf{R}_M^{(II)}|\!|\!|^{k_0,\gamma}_{-M+\frac{1}{2},s,\alpha} \leq_{S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_M(\alpha)}\big) \, , \quad \forall s_0 \leq s \leq S \, , \tag{6.122}$$

and the constant $\aleph_M(\alpha)$ is defined recursively by (6.96). Moreover

$$|\!|\!|\partial_i\mathbf{R}_M^{(I)}[\hat{\imath}]|\!|\!|_{0,s_1,\alpha} + |\!|\!|\partial_i\mathbf{R}_M^{(II)}[\hat{\imath}]|\!|\!|_{-M+\frac{1}{2},s_1,\alpha} \leq_{M,S,\alpha} \varepsilon\|\hat{\imath}\|_{s_1+\sigma+\aleph_M(\alpha)} \tag{6.123}$$

$$|\!|\!|\partial_i\mathbf{\Phi}_M^{\pm 1}[\hat{\imath}]|\!|\!|_{0,s_1,0} \, , \, |\!|\!|\partial_i(\mathbf{\Phi}_M^{\pm 1})^*[\hat{\imath}]|\!|\!|_{0,s_1,0} \leq_{M,S} \varepsilon\|\hat{\imath}\|_{s_1+\sigma+\aleph_M(0)} \, . \tag{6.124}$$

Proof. Let us prove (6.119). For all $4 \le n \le M+4$, $s_0 \le s \le S$, we have

$$|\Phi_n - \mathbb{I}_2|_{0,s,0}^{k_0,\gamma} \overset{(6.97)}{=} |\Psi_n|_{0,s,0}^{k_0,\gamma} \overset{(6.105)}{\le_S} \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_n(0)}^{k_0,\gamma}\big) \le_S \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_M(0)}^{k_0,\gamma}\big)$$

and (6.119) follows as in the proof of Corollary 4.1 in [7]. The estimate on the adjoint operator $(\mathbf{\Phi}_M^{\pm 1} - \mathbb{I}_2)^*$ follows as well since Lemma 2.9 implies $|(\Phi_n^{\pm 1} - \mathbb{I}_2)^*|_{0,s,0}^{k_0,\gamma} \le_M |\Phi_n^{\pm 1} - \mathbb{I}_2|_{0,s+s_0,0}^{k_0,\gamma}$. Also (6.124) is proved analogously. □

The operator $\mathcal{L}_M$ in (6.120) is block-diagonal up to the smoothing remainder $\mathbf{R}_M^{(II)} \in OPS^{\frac{1}{2}-M}$. The prize which has been paid is that $\mathbf{R}_M^{(II)}$ depends on $\aleph_M(\alpha)$-derivatives of the approximate solution $\mathfrak{I}$, i.e. on $\|\mathfrak{I}\|_{s+\sigma+\aleph_M(\alpha)}^{k_0,\gamma}$ in (6.122). In any case, the number of regularizing steps M is fixed (independently on s, see (7.9), (7.6)), determined by the KAM reducibility scheme in section 7.

6.6 Elimination of order ∂_x: Egorov method

The goal of this section is to remove $\mathbf{B}_1(\varphi, x)\partial_x$ from the operator $\mathcal{L}_M$ defined in (6.120). We rewrite

$$\mathcal{L}_M = \omega\cdot\partial_\varphi \mathbb{I}_2 + \mathbf{P}_0(\varphi, x, D) + \mathbf{R}_M^{(II)} \tag{6.125}$$

where we denote the whole block-diagonal part by

$$\mathbf{P}_0(\varphi, x, D) := \mathrm{i}\mathtt{m}_3 \mathbf{T}(D) + \mathbf{B}_1(\varphi, x)\partial_x + \mathrm{i}\mathbf{B}_0^{(I)}(\varphi, x)\mathcal{H}|D|^{\frac{1}{2}} + \mathbf{R}_M^{(I)} = \begin{pmatrix} \mathrm{Op}(p_0) & 0 \\ 0 & \overline{\mathrm{Op}(p_0)} \end{pmatrix} \tag{6.126}$$

and, by (6.63), (6.57), (6.75), (6.76), (6.121), the associated symbol is

$$p_0(\varphi, x, \xi) := \mathrm{i}\big(\mathtt{m}_3 T(\xi) + a_{11}(\varphi, x)\xi\big) + a_{12}(\varphi, x)\chi(\xi)\mathrm{sign}(\xi)|\xi|^{\frac{1}{2}} + r_M^{(I)}(\varphi, x, \xi) \in S^{3/2} \tag{6.127}$$

where $T(\xi) = \chi(\xi)|\xi|^{1/2}(1 + \kappa\xi^2)^{1/2}$.

Egorov approach. We transform $\mathcal{L}_M$ in (6.125) by the *flow* of the system of pseudo-PDEs

$$\partial_t \begin{pmatrix} u \\ \overline{u} \end{pmatrix} = \mathrm{i}\mathbf{a}(\varphi, x)|D|^{\frac{1}{2}} \begin{pmatrix} u \\ \overline{u} \end{pmatrix} \qquad \text{where} \qquad \mathbf{a}(\varphi, x) := \begin{pmatrix} a(\varphi, x) & 0 \\ 0 & -a(\varphi, x) \end{pmatrix} \tag{6.128}$$

and $a(\varphi, x)$ is a *real* valued function to be determined, see (6.153). The flow $\mathbf{\Phi}(\varphi, t)$ of (6.128) has the block-diagonal form

$$\mathbf{\Phi}(\varphi, t) := \begin{pmatrix} \Phi(\varphi, t) & 0 \\ 0 & \overline{\Phi}(\varphi, t) \end{pmatrix} \tag{6.129}$$

where $\Phi(\varphi, t)$ is the flow of the scalar linear pseudo-PDE

$$\partial_t u = \mathrm{i}a(\varphi, x)|D|^{\frac{1}{2}} u\,. \tag{6.130}$$

In the Appendix we prove that its flow $\Phi(\varphi, t) : H^s \mapsto H^s$ is well defined in the Sobolev spaces H^s, see Propositions 9.2, 9.5. The flow $\Phi(\varphi, t)$ solves

$$\begin{cases} \partial_t \Phi(\varphi, t) = \mathrm{i}A(\varphi)\Phi(\varphi, t)\,, & A(\varphi) := \mathfrak{a}(\varphi, x, D)\,, \quad \mathfrak{a}(\varphi, x, \xi) := a(\varphi, x)\chi(\xi)|\xi|^{\frac{1}{2}}\,, \\ \Phi(\varphi, 0) = \mathrm{Id} \end{cases} \tag{6.131}$$

and, since (6.130) is autonomous, it satisfies the group property

$$\Phi(\varphi, t_1 + t_2) = \Phi(\varphi, t_1) \circ \Phi(\varphi, t_2)\,, \qquad \Phi(\varphi, t)^{-1} = \Phi(\varphi, -t)\,. \tag{6.132}$$

Moreover, assuming that $a(\omega, \kappa, \cdot)$ is k_0-times differentiable smooth with respect to the parameters ω and κ, the flow $\Phi(\varphi, t, \omega, \kappa)$ is also k_0-times differentiable with respect to ω and κ see Proposition 9.10. If $a(\varphi, x)$ is odd(φ)-even(x) then the flow $\mathbf{\Phi}(\varphi, t)$ is even and reversibility preserving.

We denote for simplicity $\Phi := \Phi(\varphi) := \Phi(\varphi, 1)$ the time-1 flow map of (6.130) and $\mathbf{\Phi} := \mathbf{\Phi}(\varphi) := \mathbf{\Phi}(\varphi, 1)$ the time-1 flow map of the system (6.128). The transformed operator is

$$\mathcal{L}_M^{(1)} := \mathbf{\Phi}\mathcal{L}_M\mathbf{\Phi}^{-1} = \omega\cdot\partial_\varphi\mathbb{I}_2 + \mathbf{\Phi}(\varphi)\mathbf{P}_0(\varphi, x, D)\mathbf{\Phi}(\varphi)^{-1} + \mathbf{\Phi}(\varphi)\omega\cdot\partial_\varphi\{\mathbf{\Phi}(\varphi)^{-1}\} + \mathbf{\Phi}\mathbf{R}_M^{(II)}\mathbf{\Phi}^{-1} . \quad (6.133)$$

The terms $\mathbf{\Phi}(\varphi)\mathbf{P}_0(\varphi, x, D)\mathbf{\Phi}(\varphi)^{-1}$ and $\mathbf{\Phi}(\varphi)\,\omega\cdot\partial_\varphi\{\mathbf{\Phi}(\varphi)^{-1}\}$ are block-diagonal. They are classical pseudo-differential operators and shall be analyzed by an Egorov type argument. On the other hand the off-diagonal term $\mathbf{\Phi}\mathbf{R}_M^{(II)}\mathbf{\Phi}^{-1}$ is very regularizing and satisfy tame estimates. The contents of this section are summarized in Proposition 6.26.

Analysis of $\mathbf{\Phi}(\varphi)\mathbf{P}_0(\varphi, x, D)\mathbf{\Phi}(\varphi)^{-1}$ in (6.133).

We first consider $\mathbf{P}(\varphi, t) := \mathbf{\Phi}(\varphi, t)\mathbf{P}_0\mathbf{\Phi}(\varphi, t)^{-1}$. By (6.126) and (6.129) it reads

$$\mathbf{P}(\varphi, t) := \begin{pmatrix} P(\varphi, t) & 0 \\ 0 & \overline{P}(\varphi, t) \end{pmatrix} , \qquad P(\varphi, t) := \Phi(\varphi, t)p_0(\varphi, x, D)\Phi^{-1}(\varphi, t) . \quad (6.134)$$

The operator $\mathbf{P}(\varphi, t)$ solves the vector valued Heisenberg equation

$$\begin{cases} \partial_t\mathbf{P}(\varphi, t) = \mathrm{i}[\mathbf{a}(\varphi, x)|D|^{\frac{1}{2}}, \mathbf{P}(\varphi, t)] \\ \mathbf{P}(\varphi, 0) = \mathbf{P}_0(\varphi) , \end{cases}$$

namely the operator $P(\varphi, t)$ solves the usual Heisenberg equation

$$\begin{cases} \partial_t P(\varphi, t) = \mathrm{i}[A(\varphi), P(\varphi, t)] \\ P(\varphi, 0) = P_0 := p_0(\varphi, x, D) \end{cases} \qquad \text{where} \qquad A(\varphi) := \mathfrak{a}(\varphi, x, D) = a(\varphi, x)|D|^{\frac{1}{2}} . \quad (6.135)$$

We use the notation $|D|^{\frac{1}{2}} := \mathrm{Op}(\chi(\xi)|\xi|^{\frac{1}{2}})$ as in (2.25).

We look for an approximate solution $Q(\varphi, t) := q(t, \varphi, x, D)$ of (6.135) with a symbol of the form (expanded in decreasing symbols)

$$q(t, \varphi, x, \xi) = \sum_{n=0}^{M} q_n(t, \varphi, x, \xi) , \quad q_n(t, \varphi, x, \xi) \in S^{\frac{1}{2}(3-n)} , \quad \forall n = 0, \ldots, M . \quad (6.136)$$

The order of the commutator $[A(\varphi), Q(\varphi)]$ is strictly less than the order of $Q(\varphi)$. Let $\mathfrak{a} \star q$ denote the symbol of the commutator, i.e. $[A(\varphi), Q(\varphi)] := \mathrm{Op}(\mathfrak{a} \star q)$, see (2.56).

Lemma 6.12. (Commutator symbol) *If $q \in S^m$, $m \in \mathbb{R}$, then $\mathfrak{a} \star q \in S^{m-\frac{1}{2}}$ and*

$$|[A, Q]|_{m-\frac{1}{2}, s, \alpha}^{k_0,\gamma} = |\mathrm{Op}(\mathfrak{a} \star q)|_{m-\frac{1}{2}, s, \alpha}^{k_0,\gamma} \leq_{m,s,\alpha} |\mathrm{Op}(q)|_{m, s+\alpha+3, \alpha+1}^{k_0,\gamma} \|a\|_{s_0+|m|+\alpha+2}^{k_0,\gamma} + |\mathrm{Op}(q)|_{m, s_0+\alpha+3, \alpha+1}^{k_0,\gamma} \|a\|_{s+|m|+\alpha+2}^{k_0,\gamma} .$$

Proof. By Lemma 2.8 with $m' = 1/2$. □

We solve approximately the equation (6.135) in decreasing orders. We define q_0 as the solution of

$$\begin{cases} \partial_t q_0(t, \varphi, x, \xi) = 0 \\ q_0(0, \varphi, x, \xi) = p_0(\varphi, x, \xi) , \end{cases} \quad (6.137)$$

namely

$$q_0(t, \varphi, x, \xi) = p_0(\varphi, x, \xi) \in S^{\frac{3}{2}} , \qquad \forall t \in [0, 1] . \quad (6.138)$$

Then we define inductively the symbols $q_n(t, \varphi, x, \xi)$, $n \geq 1$, as the solutions of

$$\begin{cases} \partial_t q_n = \mathrm{i}\mathfrak{a} \star q_{n-1} \\ q_n(0, \varphi, x, \xi) = 0 , \end{cases} \quad (6.139)$$

namely

$$q_n(t,\varphi,x,\xi) = \mathrm{i}\int_0^t (\mathfrak{a}\star q_{n-1})(\tau,\varphi,x,\xi)\,d\tau\,. \tag{6.140}$$

Each symbol $q_n \in S^{\frac{1}{2}(3-n)}$, $\forall n = 0,\ldots,M$. Actually $q_0 \in S^{3/2}$ by (6.138). Then, by induction, if $q_{n-1} \in S^{\frac{1}{2}(3-(n-1))}$ we deduce that $\mathfrak{a}\star q_{n-1} \in S^{\frac{1}{2}(3-n)}$ by Lemma 6.12. The quantitative estimate is given in (6.190).

We now expand the symbol q in (6.136) writing explicitly the terms of order greater than 0. They come from $q_0 \in S^{\frac{3}{2}}$, $q_1 \in S^1$ and $q_2 \in S^{\frac{1}{2}}$ (all the symbols q_n, $n \geq 2$, are yet in S^0). For that we further expand as in (2.57) the symbol of the commutator as

$$(\mathfrak{a}\star q)(t,\varphi,x,\xi) = -\mathrm{i}\{\mathfrak{a},q\}(t,\varphi,x,\xi) + \mathtt{r}_2(\mathfrak{a},q)(t,\varphi,x,\xi) \tag{6.141}$$

where $\{\mathfrak{a},q\} = (\partial_x q)(\partial_\xi \mathfrak{a}) - (\partial_\xi q)(\partial_x \mathfrak{a})$ is the Poisson bracket and $\mathtt{r}_2(\mathfrak{a},q)$ is a lower order symbol.

Lemma 6.13. (Lower order commutator symbol) *If $q \in S^m$, $m \in \mathbb{R}$, then $\mathtt{r}_2(\mathfrak{a},q) \in S^{m-\frac{3}{2}}$ and*

$$\begin{aligned}|\mathrm{Op}\big(\mathtt{r}_2(\mathfrak{a},q)\big)|^{k_0,\gamma}_{m-\frac{3}{2},s,\alpha} \lesssim_{m,s,\alpha} |\mathrm{Op}(q)|^{k_0,\gamma}_{m,s+\alpha+5,\alpha+2}\|a\|^{k_0,\gamma}_{s_0+|m|+\alpha+4} \\ + |\mathrm{Op}(q)|^{k_0,\gamma}_{m,s_0+\alpha+5,\alpha+2}\|a\|^{k_0,\gamma}_{s+|m|+\alpha+4}\,.\end{aligned}$$

Proof. Apply (2.42) to $\mathrm{Op}(q)\circ\mathrm{Op}(\mathfrak{a})$ and to $\mathrm{Op}(\mathfrak{a})\circ\mathrm{Op}(q)$ with $N = 2$ and $m' = 1/2$ (and use (2.37)). □

We now get the expansion of the symbol $q_{\leq 2}(\varphi,x,\xi) := q_{\leq 2}(1,\varphi,x,\xi) = (q_0+q_1+q_2)(1,\varphi,x,\xi)$.

Lemma 6.14. (Expansion of approximate solution) *The symbol $q_{\leq 2} = q_0+q_1+q_2$ has the expansion*

$$q_{\leq 2} = \mathrm{i}\mathtt{m}_3 T(\xi) + \mathrm{i}\big(a_{11} - \frac{3}{2}\mathtt{m}_3\sqrt{\kappa}\,a_x\big)\xi + \big(\mathrm{i}a_{13} + a_{12}\,\mathrm{sign}(\xi)\big)\chi(\xi)|\xi|^{\frac{1}{2}} + r_{q_{\leq 2}} \tag{6.142}$$

where the symbol

$$r_{q_{\leq 2}} := r_{q_{\leq 2}}(\varphi,x,\xi) = r_M^{(I)} + r^{(0)}_{\mathfrak{a}p_0} + r^{(1)}_{\mathfrak{a}p_0} + r^{(2)}_{\mathfrak{a}p_0} \in S^0 \tag{6.143}$$

is defined in (6.148), (6.150), (6.152), *and $r_M^{(I)}$ in Proposition 6.11, and the function*

$$a_{13} := a_{13}(\varphi,x) := \frac{1}{2}(a_{11})_x a - a_{11}a_x - \frac{3}{8}\mathtt{m}_3\sqrt{\kappa}a_{xx}a + \frac{3}{4}\mathtt{m}_3\sqrt{\kappa}a_x^2\,. \tag{6.144}$$

Proof. By (6.140), (6.138), (6.141) we have

$$\begin{aligned} q_1(t,\varphi,x,\xi) &= \mathrm{i}\int_0^t (\mathfrak{a}\star q_0)(\tau,\varphi,x,\xi)\,d\tau = \mathrm{i}\,t\,(\mathfrak{a}\star p_0)(\varphi,x,\xi) \\ &= t\,\{\mathfrak{a},p_0\}(\varphi,x,\xi) + \mathrm{i}\,t\,\mathtt{r}_2(\mathfrak{a},p_0)(\varphi,x,\xi) \in S^1 \end{aligned} \tag{6.145}$$

and note that $\mathtt{r}_2(\mathfrak{a},p_0) \in S^0$. Similarly, using also (6.145), the symbol

$$\begin{aligned} q_2(1,\varphi,x,\xi) &= \mathrm{i}\int_0^1 (\mathfrak{a}\star q_1)(\tau,\varphi,x,\xi)d\tau = \int_0^1 \{\mathfrak{a},q_1\}(\tau,\varphi,x,\xi)d\tau + \mathrm{i}\int_0^1 \mathtt{r}_2(\mathfrak{a},q_1)(\tau,\varphi,x,\xi)d\tau \\ &= \frac{1}{2}\big(\{\mathfrak{a},\{\mathfrak{a},p_0\}\} + \mathrm{i}\{\mathfrak{a},\mathtt{r}_2(\mathfrak{a},p_0)\}\big) + \mathrm{i}\int_0^1 \mathtt{r}_2(\mathfrak{a},q_1)(\tau,\varphi,x,\xi)d\tau \in S^{1/2} \end{aligned} \tag{6.146}$$

where $\{\mathfrak{a},\mathtt{r}_2(\mathfrak{a},p_0)\}$ and $\mathtt{r}_2(\mathfrak{a},q_1) \in S^{-1/2}$. By (6.138), (6.145) at $t=1$, and (6.146) we get

$$q_{\leq 2} = q_0 + q_1 + q_2 = p_0 + \{\mathfrak{a},p_0\} + \frac{1}{2}\{\mathfrak{a},\{\mathfrak{a},p_0\}\} + r^{(0)}_{\mathfrak{a}p_0} \tag{6.147}$$

where

$$r^{(0)}_{\mathfrak{a}p_0} := \mathrm{i}\mathtt{r}_2(\mathfrak{a},p_0) + \frac{\mathrm{i}}{2}\{\mathfrak{a},\mathtt{r}_2(\mathfrak{a},p_0)\} + \mathrm{i}\int_0^1 \mathtt{r}_2(\mathfrak{a},q_1)(\tau,\varphi,x,\xi)\,d\tau \in S^0\,. \tag{6.148}$$

By (6.127) and $\partial_\xi T(\xi) = \frac{3}{2}\sqrt{\kappa}\,\mathrm{sign}(\xi)\chi(\xi)|\xi|^{\frac12} + O(|\xi|^{-\frac32})$, we get

$$
\begin{aligned}
\{\mathfrak{a}, p_0\} &= \mathrm{i}\{a\chi(\xi)|\xi|^{\frac12}, \mathtt{m}_3 T(\xi) + a_{11}\xi\} + \tilde r_{\mathfrak{a}p_0} \\
&= -\mathrm{i}\mathtt{m}_3 \partial_\xi T(\xi) a_x \chi(\xi)|\xi|^{\frac12} + \mathrm{i}\Big(\frac12 (a_{11})_x a - a_{11}a_x\Big)\chi(\xi)|\xi|^{\frac12} + \mathrm{i}(a_{11})_x a(\partial_\xi\chi(\xi))|\xi|^{\frac12}\xi + \tilde r_{\mathfrak{a}p_0} \\
&= -\mathrm{i}\frac32 \mathtt{m}_3\sqrt{\kappa}\, a_x\xi + \mathrm{i}\Big(\frac12(a_{11})_x a - a_{11}a_x\Big)\chi(\xi)|\xi|^{\frac12} + r^{(1)}_{\mathfrak{a}p_0} \qquad (6.149)
\end{aligned}
$$

where $\tilde r_{\mathfrak{a}p_0} := \{a\chi(\xi)|\xi|^{\frac12}, a_{12}\mathrm{sign}(\xi)\chi(\xi)|\xi|^{\frac12} + r_M^{(I)}\} \in S^0$ and

$$
\begin{aligned}
r^{(1)}_{\mathfrak{a}p_0} &:= \tilde r_{\mathfrak{a}p_0} - \mathrm{i}\mathtt{m}_3\Big(\partial_\xi T(\xi) - \frac32\sqrt{\kappa}\,\mathrm{sign}(\xi)\chi(\xi)|\xi|^{\frac12}\Big)a_x\chi|\xi|^{1/2} \\
&\quad + \mathrm{i}\frac32\mathtt{m}_3\sqrt{\kappa}a_x(1-\chi^2(\xi))\xi + \mathrm{i}(a_{11})_x\, a\,(\partial_\xi\chi(\xi))|\xi|^{\frac12}\xi \in S^0\,.
\end{aligned} \qquad (6.150)
$$

Furthermore, using (6.149), we compute

$$
\frac12\{\mathfrak{a},\{\mathfrak{a},p_0\}\} = -\mathrm{i}\frac34\mathtt{m}_3\sqrt{\kappa}\Big(\frac12 a_{xx}a - a_x^2\Big)\chi(\xi)|\xi|^{\frac12} + r^{(2)}_{\mathfrak{a}p_0} \qquad (6.151)
$$

where

$$
r^{(2)}_{\mathfrak{a}p_0} := \Big\{a\chi(\xi)|\xi|^{1/2}, \mathrm{i}\Big(\frac12(a_{11})_x a - a_{11}a_x\Big)\chi(\xi)|\xi|^{1/2} + r^{(1)}_{\mathfrak{a}p_0}\Big\} - \mathrm{i}\frac34\sqrt{\kappa}\mathtt{m}_3 a_{xx}a(\partial_\xi\chi(\xi))|\xi|^{\frac12}\xi \in S^0\,. \qquad (6.152)
$$

Finally (6.147), (6.127), (6.149), (6.151) imply (6.142)-(6.143). $\square$

Choice of the function $a(\varphi,x)$**.** We now choose the function $a(\varphi,x)$ so that the first order term in (6.142) vanishes, namely such that $a_{11}(\varphi,x) - \frac32\mathtt{m}_3\sqrt{\kappa}a_x(\varphi,x) = 0$. Since the function $a_{11}(\varphi,x)$ is odd in x (see (6.75) and remark 6.4) such equation may be solved. Its solution is

$$
a(\varphi,x) := \tilde a(\varphi,x) + a_0(\varphi) \qquad \text{where} \qquad \tilde a(\varphi,x) := \frac{2}{3\mathtt{m}_3\sqrt{\kappa}}\partial_x^{-1}a_{11}(\varphi,x) \qquad (6.153)
$$

and the function $a_0(\varphi)$ will be determined later, see (6.169). In this way (by (6.142))

$$
q_{\le 2} = \mathrm{i}\mathtt{m}_3 T(\xi) + \big(\mathrm{i}a_{13} + a_{12}\,\mathrm{sign}(\xi)\big)\chi(\xi)|\xi|^{\frac12} + r_{q_{\le 2}} \qquad (6.154)
$$

where $r_{q_{\le 2}} \in S^0$. The next lemma proves that we have found an approximate solution of (6.135).

Lemma 6.15. (Approximate solution of (6.135)**)** *The operator* $Q(\varphi,t) = q(t,\varphi,x,D)$ *where* $q = \sum_{n=0}^M q_n$ *with* q_0 *defined in* (6.138) *and* q_n, $n = 1,\dots,M$ *in* (6.140)*, solves the approximate Heisenberg equation*

$$
\begin{cases}
\partial_t Q(\varphi,t) = \mathrm{i}[A(\varphi), Q(\varphi,t)] + R_M(\varphi,t) \\
Q(0) = P_0
\end{cases} \qquad (6.155)
$$

where $R_M(\varphi,t) := -\mathrm{i}\mathrm{Op}(\mathfrak{a}\star q_M) \in OPS^{1-\frac{M}{2}}$*. The quantitative estimate is given in* (6.192)*.*

Proof. By (6.137) and (6.139) the initial symbol $q(0,\varphi,x,\xi) = q_0(0,\varphi,x,\xi) + \sum_{n=1}^M q_n(0,\varphi,x,\xi) = p_0(\varphi,x,\xi)$. Hence $Q(0) = P_0$. Moreover (6.137) and (6.139) imply

$$
\partial_t q = \sum_{n=0}^M \partial_t q_n = \mathrm{i}\sum_{n=1}^M \mathfrak{a}\star q_{n-1} = \mathrm{i}\sum_{n=0}^{M-1}\mathfrak{a}\star q_n = \mathrm{i}\sum_{n=0}^M \mathfrak{a}\star q_n - \mathrm{i}\mathfrak{a}\star q_M = \mathrm{i}\mathfrak{a}\star q - \mathrm{i}\mathfrak{a}\star q_M
$$

because $\mathfrak{a}\star q$ is linear in q. Since $[A(\varphi),Q] = \mathrm{Op}(\mathfrak{a}\star q)$ we get (6.155) with $R_M(\varphi,t) := -\mathrm{i}\mathrm{Op}(\mathfrak{a}\star q_M)$. The operator $R_M \in OPS^{1-\frac{M}{2}}$ since $q_M \in S^{\frac12(3-M)}$, see after (6.139)-(6.140). $\square$

The next lemma expresses the difference between $P(\varphi, t)$ and the approximate solution $Q(\varphi, t)$ of (6.135) in terms of the remainder R_M in (6.155) and the flow $\Phi(\varphi, t)$ of (6.130).

Lemma 6.16. *We have*

$$W(\varphi, t) := Q(\varphi, t) - P(\varphi, t) = \int_0^t \Phi(\varphi, t-\tau) R_M(\varphi, \tau) \Phi(\varphi, \tau - t)\, d\tau\,. \tag{6.156}$$

Proof. Recalling (6.134) we write $W(\varphi, t) = \big(Q(\varphi, t)\Phi(\varphi, t) - \Phi(\varphi, t) P_0\big)\Phi(\varphi, t)^{-1}$. By (6.131) and (6.155) we deduce that $V(\varphi, t) := Q(\varphi, t)\Phi(\varphi, t) - \Phi(\varphi, t) P_0$ solves the non-homogeneous equation

$$\partial_t V(\varphi, t) = \mathrm{i} A(\varphi) V(\varphi, t) + R_M(\varphi, t)\Phi(\varphi, t)\,, \quad V(\varphi, t)(\varphi, 0) = 0\,.$$

By Duhamel principle (variation of constants method) and (6.132) we get

$$V(\varphi, t) := \int_0^t \Phi(\varphi, t-\tau) R_M(\varphi, \tau)\Phi(\varphi, \tau)\, d\tau$$

and thus (6.156) using again (6.132). □

Analysis of $\mathbf{\Phi}(\varphi)\omega\cdot\partial_\varphi\{\mathbf{\Phi}(\varphi)^{-1}\}$ in (6.133)**.**
Set for brevity (recall (6.129))

$$\mathbf{\Psi}(\varphi, t) := \mathbf{\Phi}(\varphi, t)\omega\cdot\partial_\varphi\{\mathbf{\Phi}(\varphi, t)^{-1}\} = \begin{pmatrix} \Psi(\varphi, t) & 0 \\ 0 & \overline{\Psi}(\varphi, t)\end{pmatrix} \quad \text{where} \quad \Psi(\varphi, t) := \Phi(\varphi, t)\omega\cdot\partial_\varphi\{\Phi(\varphi, t)^{-1}\}\,.$$

The term $\Psi(\varphi, t)$ can be computed in terms of the flow Φ of (6.130) and $A(\varphi) = a(\varphi, x)|D|^{\frac{1}{2}}$.

Lemma 6.17. *The operator*

$$\Psi(\varphi, t) = -\mathrm{i}\int_0^t S_\omega(\varphi, \tau)\, d\tau \qquad \text{where} \qquad S_\omega(\varphi, t) := \Phi(\varphi, t)(\omega\cdot\partial_\varphi A(\varphi))\Phi(\varphi, t)^{-1}\,.$$

Proof. By (6.132) the flow $\Phi^{-1}(t) = \Phi(-t)$ and $\partial_t \Phi(t)^{-1} = -\mathrm{i}A\Phi(t)^{-1}$. Thus $\Psi(\varphi, t)$ solves

$$\begin{aligned}\partial_t\Psi(\varphi, t) &= (\partial_t\Phi)\omega\cdot\partial_\varphi\Phi^{-1} + \Phi\omega\cdot\partial_\varphi(\partial_t\Phi^{-1}) = -\Phi(\partial_t\Phi^{-1})\Phi\omega\cdot\partial_\varphi\Phi^{-1} - \mathrm{i}\Phi\omega\cdot\partial_\varphi(A\Phi^{-1}) \\ &= \mathrm{i}\Phi A\omega\cdot\partial_\varphi\Phi^{-1} - \mathrm{i}\Phi A\omega\cdot\partial_\varphi\Phi^{-1} - \mathrm{i}\Phi(\omega\cdot\partial_\varphi A)\Phi^{-1} = -\mathrm{i}\Phi(\omega\cdot\partial_\varphi A)\Phi^{-1}\,.\end{aligned}$$

Moreover $\Psi(\varphi, 0) = 0$ (as $\Phi(\varphi, 0) = \mathrm{Id}$, $\forall\varphi\in\mathbb{T}^\nu$, see (6.131)). The lemma follows by integration. □

The operator $S_\omega(\varphi, t)$ has the same conjugation structure of $P(\varphi, t)$ in (6.134) and therefore it solves the Heisenberg equation

$$\begin{cases}\partial_t S_\omega(\varphi, t) = \mathrm{i}[A(\varphi), S_\omega(\varphi, t)] \\ S_\omega(\varphi, 0) = (\omega\cdot\partial_\varphi a)|D|^{\frac{1}{2}}\,.\end{cases} \tag{6.157}$$

Following the same procedure used for $P(\varphi, t)$, we look for an approximate solution of (6.157) of the form (expansion in decreasing symbols)

$$S_{\omega, M}(\varphi, t) := s(t, \varphi, x, D)\,, \qquad s = \sum\nolimits_{n=0}^M s_n\,, \qquad s_n \in S^{\frac{1}{2}(1-n)}\,. \tag{6.158}$$

We define the principal symbol s_0 to be the solution of

$$\begin{cases}\partial_t s_0(t, \varphi, x, \xi) = 0 \\ s_0(0, \varphi, x, \xi) = (\omega\cdot\partial_\varphi a)\chi(\xi)|\xi|^{\frac{1}{2}}\,,\end{cases} \qquad i.e. \quad s_0(t, \varphi, x, \xi) = (\omega\cdot\partial_\varphi a)\chi(\xi)|\xi|^{\frac{1}{2}} \in S^{1/2}\,. \tag{6.159}$$

Then we define inductively the symbols s_n, $n \geq 1$, as the solutions of

$$\begin{cases}\partial_t s_n = \mathrm{i}\mathfrak{a}\star s_{n-1} \\ s_n(0, \varphi, x, \xi) = 0\,,\end{cases} \qquad i.e. \quad s_n(t, \varphi, x, \xi) = \mathrm{i}\int_0^t (\mathfrak{a}\star s_{n-1})(\tau, \varphi, x, \xi)\, d\tau\,. \tag{6.160}$$

It turns out that $s_n \in S^{\frac{1}{2}(1-n)}$, in particular each $s_n \in S^0$, $\forall n \geq 1$.

Lemma 6.18. (Approximate solution of (6.157)**)** *The pseudo-differential operator* $S_{\omega,M}(\varphi,t) = s(\varphi,t,x,D)$ *in* (6.158) *with* $s_0 \in S^{\frac{1}{2}}$ *defined in* (6.159) *and* $s_n \in S^{\frac{1}{2}(1-n)}$, $n = 1,\ldots,M$ *in* (6.160)*, solves the approximate Heisenberg equation*

$$\begin{cases} \partial_t S_{\omega,M}(\varphi,t) = \mathrm{i}[A(\varphi), S_{\omega,M}(\varphi,t)] + R_{\omega,M}(\varphi,t) \\ S_{\omega,M}(\varphi,0) = (\omega\cdot\partial_\varphi a)|D|^{\frac{1}{2}} \end{cases} \tag{6.161}$$

where $R_{\omega,M}(\varphi,t) := -\mathrm{i}\mathrm{Op}(\mathfrak{a}\star s_M) \in OPS^{-\frac{M}{2}}$. *Moreover*

$$W_\omega(\varphi,t) := S_{\omega,M}(\varphi,t) - S_\omega(\varphi,t) = \int_0^t \Phi(\varphi,t-\tau)R_{\omega,M}(\varphi,\tau)\Phi(\varphi,\tau-t)\,d\tau \tag{6.162}$$

where $\Phi(\varphi,t)$ *denotes the flow of* (6.130).

Proof. The equation (6.161) follows as in Lemma 6.15. Then (6.162) follows as in Lemma 6.16. □

Sub-principal symbol of $\mathcal{L}_M^{(1)}$. By Lemma 6.14 and the choice of $a(\varphi,x)$ in (6.153), the principal and subprincipal symbols of $\mathbf{\Phi}(\varphi)\mathbf{P}_0(\varphi,x,D)\mathbf{\Phi}(\varphi)^{-1}$ are given by (6.154). Also $\mathbf{\Phi}(\varphi)\omega\cdot\partial_\varphi\{\mathbf{\Phi}(\varphi)^{-1}\}$ contributes to the subprincipal symbol of $\mathcal{L}_M^{(1)}$, i.e to $OPS^{1/2}$. By Lemmata 6.17, 6.18 and the expression of $s_0 = (\omega\cdot\partial_\varphi a)\chi(\xi)|\xi|^{\frac{1}{2}}$ in (6.159) we find that the conjugated operator $\mathcal{L}_M^{(1)}$ in (6.133) has the expansion

$$\mathcal{L}_M^{(1)} = \omega\cdot\partial_\varphi \mathbb{I}_2 + \mathrm{i}\mathtt{m}_3\mathbf{T}(D) + \mathrm{i}\big(\mathbf{C}_1(\varphi,x) + \mathbf{C}_0(\varphi,x)\mathcal{H}\big)|D|^{\frac{1}{2}} + \ldots \tag{6.163}$$

where

$$\mathbf{C}_1(\varphi,x) := \begin{pmatrix} a_{14} & 0 \\ 0 & -a_{14} \end{pmatrix},\ a_{14} := a_{13} - \omega\cdot\partial_\varphi a\,,\quad \mathbf{C}_0(\varphi,x) := \begin{pmatrix} a_{12} & 0 \\ 0 & -a_{12} \end{pmatrix}, \tag{6.164}$$

and the functions a_{13}, a_{12} are defined respectively in (6.144), (6.76).

In the next sections we reduce the operator $\mathcal{L}_M^{(1)}$ neglecting the term

$$\mathbf{R}_M^{(1),\perp} := \mathrm{i}\Pi_{K_n}^{\perp}\mathbf{C}_1|D|^{\frac{1}{2}} := \mathrm{i}\begin{pmatrix} \Pi_{K_n}^{\perp}a_{14}(\varphi,x) & 0 \\ 0 & -\Pi_{K_n}^{\perp}a_{14}(\varphi,x) \end{pmatrix}|D|^{\frac{1}{2}} \tag{6.165}$$

which is supported on the high Fourier frequencies and which will contribute to the remainders in (7.94)-(7.95) (as we did with the similar terms at the end of section 6.4). For simplicity of notation we still denote it by $\mathcal{L}_M^{(1)}$.

Choice of the function $a_0(\varphi)$. In view of the reduction of $\mathrm{i}\Pi_{K_n}\mathbf{C}_1|D|^{\frac{1}{2}}$ in section 6.7, we choose the function $a_0(\varphi)$ in (6.153) in such a way that, for all $\varphi\in\mathbb{T}^\nu$, the integral

$$\frac{1}{2\pi}\int_{\mathbb{T}} \Pi_{K_n}a_{14}(\varphi,x)\,dx = \mathtt{m}_{1,K_n}\,,\quad \forall\varphi\in\mathbb{T}^\nu\,, \tag{6.166}$$

is a constant. Since $a = \tilde a + a_0$ (see (6.153)) we write the function a_{14} in (6.164) as

$$a_{14}(\varphi,x) = \tilde a_{14}(\varphi,x) - \omega\cdot\partial_\varphi a_0(\varphi) \qquad \text{where} \qquad \tilde a_{14} = a_{13} - \omega\cdot\partial_\varphi \tilde a\,. \tag{6.167}$$

The function $a_{13}(\varphi,x)$ in (6.144) depends on a, and thus also on $a_0(\varphi)$, but the integral $\int_{\mathbb{T}} a_{13}(\varphi,x)dx$, and thus $\int_{\mathbb{T}}\tilde a_{14}(\varphi,x)dx$, does not depend on $a_0(\varphi)$. For solving (6.166) we look for $a_0(\varphi) = \Pi_{K_n}a_0(\varphi)$ such that $\frac{1}{2\pi}\int_{\mathbb{T}}\Pi_{K_n}\tilde a_{14}(\varphi,x)\,dx - (\omega\cdot\partial_\varphi a_0)(\varphi) = \mathtt{m}_{1,K_n}$. For all $\omega\in\mathtt{DC}^\gamma_{K_n}$ (see (1.40)) such equation is solved by

$$\mathtt{m}_{1,K_n} := (2\pi)^{-(\nu+1)}\int_{\mathbb{T}^{\nu+1}}\Pi_{K_n}\tilde a_{14}(\varphi,x)\,d\varphi\,dx = (2\pi)^{-(\nu+1)}\int_{\mathbb{T}^{\nu+1}}\tilde a_{14}(\varphi,x)\,d\varphi\,dx\,, \tag{6.168}$$

$$a_0(\varphi) := -(\omega\cdot\partial_\varphi)^{-1}\Big(\mathtt{m}_{1,K_n} - \frac{1}{2\pi}\int_{\mathbb{T}}\Pi_{K_n}\tilde a_{14}(\varphi,x)\,dx\Big)\,. \tag{6.169}$$

Note that $a_0(\varphi)$ is odd in φ. Since also $\tilde a(\varphi,x)$ defined in (6.153) is odd in φ, and even in x, the flow $\mathbf{\Phi}(\varphi,t)$ of (6.128) is even and reversibility preserving.

Lemma 6.19. (Coefficient $\mathtt{m}_{1,K_n}$) *The coefficient*

$$\mathtt{m}_{1,K_n} = -\frac{(2\pi)^{-\nu-\frac{5}{2}}}{2\sqrt{\kappa}} \int_{\mathbb{T}^{\nu+1}} (1+\beta_x)[\omega\cdot\partial_\varphi\beta + V(1+\beta_x)]^2 \Pi_{K_n}\Big(\int_{\mathbb{T}} \sqrt{1+\eta_y^2}\,dy\Big)^{3/2} d\varphi\,dx \tag{6.170}$$

where the function V is defined in (2.117) *and β in* (6.33)*. The coefficient $\mathtt{m}_{1,K_n}$ satisfies*

$$|\mathtt{m}_{1,K_n}|^{k_0,\gamma} \le C\varepsilon\,, \qquad |\partial_i \mathtt{m}_{1,K_n}[\hat\imath]| \le C\varepsilon\|\hat\imath\|_\sigma\,. \tag{6.171}$$

Proof. By (6.168), (6.167), (6.144), (6.153) the coefficient

$$\begin{aligned}
\mathtt{m}_{1,K_n} &= \frac{1}{(2\pi)^{\nu+1}}\int_{\mathbb{T}^{\nu+1}} \tilde a_{14}(\varphi,x)\,d\varphi dx = \frac{1}{(2\pi)^{\nu+1}}\int_{\mathbb{T}^{\nu+1}} a_{13}(\varphi,x)\,d\varphi dx \\
&= \frac{1}{(2\pi)^{\nu+1}}\int_{\mathbb{T}^{\nu+1}} \frac12 (a_{11})_x\tilde a - a_{11}\tilde a_x - \frac38 \mathtt{m}_3\sqrt{\kappa}\tilde a_{xx}\tilde a + \frac34 \mathtt{m}_3\sqrt{\kappa}\tilde a_x^2\,d\varphi dx \\
&= -\frac{(2\pi)^{-\nu-1}}{2\mathtt{m}_3\sqrt{\kappa}}\int_{\mathbb{T}^{\nu+1}} a_{11}^2(\varphi,x)\,d\varphi dx\,.
\end{aligned} \tag{6.172}$$

By (6.75), (6.70), $d\vartheta = (1+\omega\cdot\partial_\varphi p)d\varphi$ (by (6.69)), (6.71), (6.29), (6.23) we have

$$\int_{\mathbb{T}^{\nu+1}} a_{11}^2(\varphi,x)\,d\varphi dx = \int_{\mathbb{T}^{\nu+1}} \frac{a_1^2(\varphi,x)}{1+\omega\cdot\partial_\varphi p}\,d\varphi dx = \mathtt{m}_3\int_{\mathbb{T}^{\nu+1}} \frac{(\omega\cdot\partial_\varphi\beta + V(1+\beta_x))^2}{\Pi_{K_n} m_3(\varphi)}(1+\beta_x)\,d\varphi dx\,. \tag{6.173}$$

By (6.172), (6.173), (6.37) we deduce (6.170). □

Lemmata 6.14, 6.16, 6.17, 6.18, imply that

$$\mathcal{L}_M^{(1)} = \omega\cdot\partial_\varphi \mathbb{I}_2 + \mathrm{i}\mathtt{m}_3\mathbf{T}(D) + \mathrm{i}\big(\mathbf{C}_1(\varphi,x) + \mathbf{C}_0(\varphi,x)\mathcal{H}\big)|D|^{\frac12} + \mathbf{R}_M^{(1)} + \mathbf{Q}_M^{(1)}$$

with remainders

$$\mathbf{R}_M^{(1)} := \begin{pmatrix} \mathcal{R}_M^{(1)} & 0 \\ 0 & \overline{\mathcal{R}}_M^{(1)} \end{pmatrix}, \quad \mathbf{Q}_M^{(1)} := \begin{pmatrix} 0 & \mathcal{Q}_M^{(1)} \\ \overline{\mathcal{Q}}_M^{(1)} & 0 \end{pmatrix}$$

$$\mathcal{R}_M^{(1)} := \mathrm{Op}(r_M^{(1)}) - W(\varphi,1) + \int_0^1 W_\omega(\varphi,\tau)d\tau\,, \qquad \mathcal{Q}_M^{(1)} := \Phi\mathcal{R}_M^{(II)}\overline{\Phi}^{-1}\,, \tag{6.174}$$

$$r_M^{(1)}(\varphi,x,\xi) := r_{q_{\le 2}}(\varphi,x,\xi) + \sum\nolimits_{n=3}^M q_n(1,\varphi,x,\xi) + \mathrm{i}\sum\nolimits_{n=1}^M \int_0^1 s_n(\tau,\varphi,x,\xi)d\tau \in S^0$$

where $r_{q_{\le 2}}$ is defined in (6.143), q_n in (6.140), s_n in (6.160), the operator W is defined in (6.156), W_ω in (6.162) and $\mathcal{R}_M^{(II)}$ in Proposition 6.11.

In the final part of this section we prove that $\mathbf{R}_M^{(1)}$ and $\mathbf{Q}_M^{(1)}$ are tame operators and (6.212) holds.

Lemma 6.20. *For all $s_0 \le s \le S$, we have*

$$\|a_{12}\|_s^{k_0,\gamma}, \|a_{13}\|_s^{k_0,\gamma}, \|a_{14}\|_s^{k_0,\gamma}, \|\tilde a\|_s^{k_0,\gamma} \le_S \varepsilon(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})\,, \quad \|a_0\|_s^{k_0,\gamma} \le_S \varepsilon\gamma^{-1}(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})\,, \tag{6.175}$$

$$\|\partial_i a_{12}[\hat\imath]\|_{s_1}, \|\partial_i a_{13}[\hat\imath]\|_{s_1}, \|\partial_i a_{14}[\hat\imath]\|_{s_1}, \|\partial_i\tilde a[\hat\imath]\|_{s_1} \le_S \varepsilon\|\hat\imath\|_{s_1+\sigma}, \|\partial_i a_0[\hat\imath]\|_{s_1} \le_S \varepsilon\gamma^{-1}\|\hat\imath\|_{s_1+\sigma}\,. \tag{6.176}$$

Lemma 6.21. *The remainder $r_{q_{\le 2}} \in S^0$ in* (6.154) *(see* (6.143)*) satisfies, for some $\sigma := \sigma(\tau,\nu) > 0$,*

$$|r_{q_{\le 2}}(x,D)|_{0,s,\alpha}^{k_0,\gamma} \le_{S,\alpha} \varepsilon\big(1+\|\mathfrak{I}_0\|_{s+\sigma+\aleph_M(\alpha+4)+2\alpha}^{k_0,\gamma}\big)\,, \quad \forall s_0 \le s \le S\,. \tag{6.177}$$

Moreover, if the constant μ in (6.10) *satisfies*

$$s_1 + \sigma + \aleph_M(\alpha+4) + 2\alpha \le s_0 + \mu\,, \tag{6.178}$$

then

$$|\partial_i r_{q_{\le 2}}(x,D)[\hat\imath]|_{0,s_1,\alpha} \le_{S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\aleph_M(\alpha+4)+2\alpha}\,. \tag{6.179}$$

Proof. We rely on the Lemmata 6.12 and 6.13. We prove that each term of $r_{q_{\leq 2}} = r_M^{(I)} + r^{(0)}_{\mathtt{a}p_0} + r^{(1)}_{\mathtt{a}p_0} + r^{(2)}_{\mathtt{a}p_0}$ defined in (6.148), (6.150), (6.152) satisfies (6.177). The term $\mathrm{Op}(r_M^{(I)})$ satisfies (6.177), (6.179) by Proposition 6.11. Then we consider $r^{(0)}_{\mathtt{a}p_0}$ in (6.148). Lemma 6.13 (with $m = 3/2$), the definition of p_0 in (6.127), the estimates of Proposition 6.11, and (6.175), imply

$$|\mathtt{r}_2(\mathtt{a}, p_0)(x, D)|^{k_0,\gamma}_{0,s,\alpha} \leq_{S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_M(\alpha+2)+\alpha}\big)\,. \tag{6.180}$$

In the same way, using $\partial_i \mathtt{r}_2(\mathtt{a}, p_0)[\hat\imath] = \mathtt{r}_2(\partial_i \mathtt{a}[\hat\imath], p_0) + \mathtt{r}_2(\mathtt{a}, \partial_i p_0[\hat\imath])$ and (6.10), (6.178), we deduce that

$$|\partial_i \mathtt{r}_2(\mathtt{a}, p_0)(x, D)[\hat\imath]|_{0,s_1,\alpha} \leq_{S,\alpha} \varepsilon \|\hat\imath\|_{s_1+\sigma+\aleph_M(\alpha+2)+\alpha}\,. \tag{6.181}$$

Lemma 2.59, (6.180) and (6.175) imply

$$\begin{aligned}|\{\mathtt{a}, \mathtt{r}_2(\mathtt{a}, p_0)\}(x, D)|^{k_0,\gamma}_{0,s,\alpha} &\leq_{s,\alpha} |\mathtt{r}_2(\mathtt{a}, p_0)(x, D)|^{k_0,\gamma}_{0,s+1,\alpha+1}\|a\|^{k_0,\gamma}_{s_0+1} + |\mathtt{r}_2(\mathtt{a}, p_0)(x, D)|^{k_0,\gamma}_{0,s_0+1,\alpha+1}\|a\|^{k_0,\gamma}_{s+1} \\ &\leq_{S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_M(\alpha+2)+\alpha}\big)\end{aligned} \tag{6.182}$$

for some $\sigma := \sigma(\tau, \nu) > 0$. Moreover $\partial_i\{\mathtt{a}, \mathtt{r}_2(\mathtt{a}, p_0)\}[\hat\imath] = \{\partial_i \mathtt{a}[\hat\imath], \mathtt{r}_2(\mathtt{a}, p_0)\} + \{\mathtt{a}, \partial_i \mathtt{r}_2(\mathtt{a}, p_0)[\hat\imath]\}$. Hence (2.59), (6.175), (6.176), (6.180), (6.181), (6.10), (6.178) imply that

$$|\partial_i\{\mathtt{a}, \mathtt{r}_2(\mathtt{a}, p_0)\}(x, D)[\hat\imath]|_{0,s_1,\alpha} \leq_{S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\aleph_M(\alpha+2)+\alpha}\,. \tag{6.183}$$

Moreover by (6.145), (6.180), (6.181), (2.59) and Proposition 6.11 (and (6.10), (6.178)) we get

$$|q_1(x, D)|^{k_0,\gamma}_{1,s,\alpha} \leq_{S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_M(\alpha+2)+\alpha}\big)\,, \tag{6.184}$$

$$|\partial_i q_1(x, D)[\hat\imath]|_{1,s_1,\alpha} \leq_{S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\aleph_M(\alpha+2)+\alpha}\,, \tag{6.185}$$

and using Lemma 6.13 (with $m = 1$), by the same arguments used to deduce (6.180), (6.181), we get

$$|\mathtt{r}_2(\mathtt{a}, q_1)(x, D)|^{k_0,\gamma}_{0,s,\alpha} \leq_{S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_M(\alpha+4)+2\alpha}\big)\,, \tag{6.186}$$

$$|\partial_i \mathtt{r}_2(\mathtt{a}, q_1)(x, D)[\hat\imath]|_{0,s_1,\alpha} \leq_{S,\alpha} \varepsilon\|\hat\imath\|_{s+\sigma+\aleph_M(\alpha+4)+2\alpha} \tag{6.187}$$

for some $\sigma := \sigma(\tau, \nu) > 0$. The estimates (6.180), (6.181), (6.182), (6.183), (6.186), (6.187) imply

$$\begin{aligned}&|r^{(0)}_{\mathtt{a}p_0}(x, D)|^{k_0,\gamma}_{0,s,\alpha} \leq_{S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_M(\alpha+4)+2\alpha}\big) \\ &|\partial_i r^{(0)}_{\mathtt{a}p_0}(x, D)[\hat\imath]|_{0,s_1,\alpha} \leq_{S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\aleph_M(\alpha+4)+2\alpha}\end{aligned}$$

for some $\sigma := \sigma(\tau, \nu) > 0$. The symbol $\tilde r_{\mathtt{a}p_0}$ defined in (6.150) satisfies

$$|\tilde r_{\mathtt{a}p_0}(x, D)|^{k_0,\gamma}_{0,s,\alpha} \leq_{S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_M(\alpha+1)}\big)\,, \tag{6.188}$$

$$|\partial_i \tilde r_{\mathtt{a}p_0}(x, D)[\hat\imath]|_{0,s_1,\alpha} \leq_{S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\aleph_M(\alpha+1)}\,, \tag{6.189}$$

by (6.122), (6.123), Lemma 6.6 and (6.64). Also the symbols $r^{(1)}_{\mathtt{a}p_0}$ in (6.150) and $r^{(2)}_{\mathtt{a}p_0}$ in (6.152) satisfy (6.188), (6.189). □

Lemma 6.22. *For all $n \in \{1, \ldots, M\}$ the symbols $q_n \in S^{\frac12(3-n)}$ defined in* (6.140) *satisfy*

$$|\mathrm{Op}(q_n)|^{k_0,\gamma}_{\frac12(3-n),s,\alpha} \leq_{n,S,\alpha} \varepsilon\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\daleth_n(M,\alpha)}\big)\,, \quad \forall s_0 \leq s \leq S\,, \tag{6.190}$$

where the constants $\daleth_n(M, \alpha)$, $n \in \{3, \ldots, M\}$ are defined inductively by

$$\daleth_1(M, \alpha) := \aleph_M(\alpha + 2) + \alpha\,, \quad \daleth_{n+1}(M, \alpha) := \alpha + \frac{n}{2} + \frac{3}{2} + \daleth_n(M, \alpha + 1)\,. \tag{6.191}$$

The operator $R_M(\varphi,t) := -\mathrm{i}\mathrm{Op}(\mathfrak{a}\star q_M) \in OPS^{1-\frac{M}{2}}$ *satisfies*

$$|R_M(\varphi,t)|^{k_0,\gamma}_{1-\frac{M}{2},s,\alpha} \leq_{M,S,\alpha} \varepsilon\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\beth_{M+1}(M,\alpha)}\big)\,,\quad \forall s_0\leq s\leq S\,. \tag{6.192}$$

Moreover if the constant μ *in* (6.10) *satisfies*

$$s_1+\sigma+\beth_{M+1}(M,\alpha)\leq s_0+\mu\,, \tag{6.193}$$

then for all $n\in\{3,\dots,M\}$

$$|\partial_i \mathrm{Op}(q_n)[\hat\imath]|_{\frac12(3-n),s_1,\alpha} \leq_{n,S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\beth_n(M,\alpha)}\,, \tag{6.194}$$

$$|\partial_i R_M(\varphi,t)[\hat\imath]|_{1-\frac{M}{2},s_1,\alpha} \leq_{M,S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\beth_{M+1}(M,\alpha)}\,. \tag{6.195}$$

Proof. For $n=1$ the estimates (6.190), (6.194) for $\mathrm{Op}(q_1)$ have been proved in (6.184), (6.185) in Lemma 6.21. Then we argue by induction supposing that $q_n\in S^{\frac12(3-n)}$ satisfies (6.190), (6.194). Then, recalling (6.140), Lemma 6.12 and (6.175) imply

$$|\mathrm{Op}(q_{n+1})|^{k_0,\gamma}_{\frac12(3-(n+1)),s,\alpha} \leq_{n,S,\alpha} \varepsilon\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\beth_{n+1}(M,\alpha)}\big)$$

where $\beth_{n+1}(M,\alpha)$ is defined in (6.191). By (6.140)

$$\partial_i\mathrm{Op}(q_{n+1})[\hat\imath] = \mathrm{i}\mathrm{Op}\Big(\int_0^t(\partial_i\mathfrak{a}[\hat\imath]\star q_{n-1})(\tau,\varphi,x,\xi)\,d\tau\Big)+\mathrm{i}\mathrm{Op}\Big(\int_0^t(\mathfrak{a}\star\partial_i q_{n-1})(\tau,\varphi,x,\xi)[\hat\imath]\,d\tau\Big)\,.$$

Then (6.175), (6.176), (6.190), (6.194), (6.10), (6.193) imply

$$|\partial_i\mathrm{Op}(q_{n+1})[\hat\imath]|_{\frac12(3-(n+1)),s_1,\alpha} \leq_{n,S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\beth_{n+1}(M,\alpha)}\,.$$

In the same way (6.192), (6.195) follow. □

Remark 6.23. We need (6.192) only for $\alpha=0$. □

We now estimate the difference $W(\varphi,t)$ in (6.156) between the approximate solution $Q(\varphi,t)$ and the exact solution $P(\varphi,t)$ of the equation (6.135).

Lemma 6.24. *For all* $\beta\in\mathbb{N}$ *with* $\beta+k_0+4\leq M$, *the operators* $\partial^\beta_{\varphi_j}W(\varphi,t)$, $\partial^\beta_{\varphi_j}[W(\varphi,t),\partial_x]$, $j=1,\dots,\nu$, *are* $\mathcal{D}^{k_0}$*-tame with tame constants*

$$\mathfrak{M}_{\partial^\beta_{\varphi_j}W(\varphi,t)}(s),\mathfrak{M}_{\partial^\beta_{\varphi_j}[W(\varphi,t),\partial_x]}(s) \leq_{S,M} \varepsilon(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\frac32 M+\daleth(M)+\beta})\,,\quad \forall s_0\leq s\leq S\,, \tag{6.196}$$

for some $\sigma:=\sigma(\tau,\nu,k_0)>0$ *and (the constants* $\beth_n(M,\alpha)$ *are defined in Lemma 6.22)*

$$\daleth(M):=\beth_{M+1}(M,0)\,. \tag{6.197}$$

Moreover if the constant μ *in* (6.10) *satisfies*

$$s_1+\sigma+\chi M+\daleth(M)+\beta\leq s_0+\mu\,, \tag{6.198}$$

then

$$\|\partial^\beta_{\varphi_j}[\partial_i W(\varphi,t)[\hat\imath],\partial_x]\|_{\mathcal{L}(H^{s_1})},\|\partial^\beta_{\varphi_j}\partial_i W(\varphi,t)[\hat\imath]\|_{\mathcal{L}(H^{s_1})} \leq_{M,S} \varepsilon\|\hat\imath\|_{s_1+\sigma+\frac32 M+\daleth(M)+\beta}\,. \tag{6.199}$$

Proof. To simplify $\partial_\varphi:=\partial_{\varphi_j}$, $j=1,\dots,\nu$. We prove that $\partial^\beta_\varphi[W(\varphi,t),\partial_x]=\partial^\beta_\varphi W(\varphi,t)\partial_x-\partial_x\partial^\beta_\varphi W(\varphi,t)$ is $\mathcal{D}^{k_0}$-tame. We first consider $\partial^\beta_\varphi W(\varphi,t)\partial_x$. Recalling (6.156) it is sufficient to estimate $\forall t,\tau\in[0,1]$

$$\partial^\beta_\varphi\partial^k_\lambda\Big(\Phi(t-\tau)R_M(\tau)\Phi(\tau-t)\Big)=\sum_{\substack{\beta_1+\beta_2+\beta_3=\beta\\ k_1+k_2+k_3=k}} C(\beta_1,\dots,k_3)\partial^{\beta_1}_\varphi\partial^{k_1}_\lambda\Phi(t-\tau)\partial^{\beta_2}_\varphi\partial^{k_2}_\lambda R_M(\tau)\partial^{\beta_3}_\varphi\partial^{k_3}_\lambda\Phi(\tau-t)$$

where $\beta_1, \beta_2, \beta_3 \in \mathbb{N}$ and $k_1, k_2, k_3 \in \mathbb{N}^{\nu+1}$. We write each term as

$$\partial_\varphi^{\beta_1}\partial_\lambda^{k_1}\Phi(t-\tau)\partial_\varphi^{\beta_2}\partial_\lambda^{k_2}R_M(\tau)\partial_\varphi^{\beta_3}\partial_\lambda^{k_3}\Phi(\tau-t)\partial_x = \partial_\varphi^{\beta_1}\partial_\lambda^{k_1}\Phi(t-\tau)\langle D\rangle^{-\frac{\beta_1+|k_1|}{2}} \tag{6.200}$$

$$\langle D\rangle^{\frac{\beta_1+|k_1|}{2}}\partial_\varphi^{\beta_2}\partial_\lambda^{k_2}R_M(\tau)\langle D\rangle^{\frac{\beta_3+|k_3|}{2}+1} \tag{6.201}$$

$$\langle D\rangle^{-\frac{\beta_3+|k_3|}{2}-1}\partial_\varphi^{\beta_3}\partial_\lambda^{k_3}\Phi(\tau-t)\partial_x\,. \tag{6.202}$$

Propositions 9.7 and 9.10 and (6.175) provide the estimates for (6.200) and (6.202): for some $\sigma := \sigma(\tau,\nu,k_0) > 0$,

$$\|\partial_\varphi^{\beta_1}\partial_\lambda^{k_1}\Phi(t-\tau)\langle D\rangle^{-\frac{\beta_1+|k_1|}{2}}h\|_s \le_s \gamma^{-|k_1|}\big(\|h\|_s + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\beta_1+\sigma}\|h\|_{s_0}\big)\,, \tag{6.203}$$

$$\|\langle D\rangle^{-\frac{\beta_3+|k_3|}{2}-1}\partial_\varphi^{\beta_3}\partial_\lambda^{k_3}\Phi(\tau-t)\partial_x h\|_s \le_s \gamma^{-|k_3|}\big(\|h\|_s + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\frac32\beta_3+\sigma}\|h\|_{s_0}\big)\,. \tag{6.204}$$

We now estimate the norm of the pseudo-differential operator in (6.201) where $R_M \in OPS^{1-\frac{M}{2}}$, see (6.192). By (2.37), $\beta_0 + k_0 + 4 \le M$, Lemmata 2.7 and 2.6, (2.40), we get

$$\begin{aligned}
&|\langle D\rangle^{\frac{\beta_1+|k_1|}{2}}\partial_\varphi^{\beta_2}\partial_\lambda^{k_2}R_M(\tau)\langle D\rangle^{\frac{\beta_3+|k_3|}{2}+1}|_{0,s,0} \le_s \\
&|\langle D\rangle^{\frac{\beta_1+|k_1|}{2}}\partial_\varphi^{\beta_2}\partial_\lambda^{k_2}R_M(\tau)\langle D\rangle^{\frac{\beta_3+|k_3|}{2}+1}|_{\frac{\beta_1+|k_1|}{2}+1-\frac{M}{2}+\frac{\beta_3+|k_3|}{2}+1,s,0} \le_s \\
&|\langle D\rangle^{\frac{\beta_1+|k_1|}{2}}\partial_\varphi^{\beta_2}\partial_\lambda^{k_2}R_M(\tau)|_{\frac{\beta_1+|k_1|}{2}+1-\frac{M}{2},s,0} \le_s |\partial_\varphi^{\beta_2}\partial_\lambda^{k_2}R_M(\tau)|_{1-\frac{M}{2},s+\frac{\beta_1+|k_1|}{2},0} \\
&\le_{s,M} \gamma^{-|k_2|}|R_M(\tau)|^{k_0,\gamma}_{1-\frac{M}{2},s+\frac32\beta+\frac{k_0}{2},0} \overset{(6.192)}{\le_{S,M}} \varepsilon\gamma^{-|k_2|}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\daleth(M)+\frac32\beta+\frac{k_0}{2}}\big)
\end{aligned} \tag{6.205}$$

where $\daleth(M) := \beth_{M+1}(M,0)$, see (6.197). Then (6.203), (6.204), (6.205) and Lemma 2.13 imply that $\partial_\varphi^\beta W(\varphi,t)\partial_x$ is $\mathcal{D}^{k_0}$-tame with tame constant $\le C(S)\varepsilon(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\frac32 M+\daleth(M)+\beta})$. The operator $\partial_x\partial_\varphi^\beta W(\varphi,t)$ satisfies a similar estimate and so (6.196) is proved.

The estimate (6.199) follows by differentiating the operator $W(\varphi,t)$ with respect to the torus i, using the same strategy as above, applying (6.10), (6.198), the estimate (6.195) for $\partial_i R_M(\tau)[\hat\imath]$, Proposition 9.10 and the estimates for $\partial_i\Phi$ in Propositions 9.13-9.14. □

The following lemma can be proved as Lemmata 6.22 and 6.24.

Lemma 6.25. *For all $n \in \{1,\ldots,M\}$ the symbols $s_n \in S^{\frac12(1-n)}$ defined in* (6.160) *satisfy*

$$|\mathrm{Op}(s_n)|^{k_0,\gamma}_{\frac12(1-n),s,\alpha} \le_{n,S,\alpha} \varepsilon\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\beth_{n+2}(M,\alpha)}\big)\,, \quad \forall s_0 \le s \le S \tag{6.206}$$

where the constants $\beth_n(M,\alpha)$ are defined in (6.191)*. The operator $R_{\omega,M}(\varphi,t) := -\mathrm{i}\mathrm{Op}(\mathfrak{a}\star s_M) \in OPS^{-\frac{M}{2}}$ satisfies*

$$|R_{\omega,M}(\varphi,t)|^{k_0,\gamma}_{-\frac{M}{2},s,\alpha} \le_{M,S,\alpha} \varepsilon\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\beth_{M+3}(M,\alpha)}\big)\,, \quad \forall s_0 \le s \le S\,.$$

For all $\beta \in \mathbb{N}$, $\beta + k_0 + 4 \le M$, the operators $\partial_{\varphi_j}^\beta W_\omega(\varphi,t)$, $\partial_{\varphi_j}^\beta[W_\omega(\varphi,t),\partial_x]$, $j = 1,\ldots,\nu$ (recall (6.162)*) are $\mathcal{D}^{k_0}$-tame where the tame constant satisfies*

$$\mathfrak{M}_{\partial_{\varphi_j}^\beta[W_\omega(\varphi,t),\partial_x]}(s)\,,\mathfrak{M}_{\partial_{\varphi_j}^\beta W_\omega(\varphi,t)}(s) \le_{M,S} \varepsilon(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\frac32 M+\daleth(M+2)+\beta})\,, \quad \forall s_0 \le s \le S\,. \tag{6.207}$$

Moreover if the constant μ in (6.10) *satisfies $s_1 + \sigma + \frac32 M + \daleth(M+2) + \beta \le s_0 + \mu$ then*

$$|\partial_i\mathrm{Op}(s_n)[\hat\imath]|_{\frac12(1-n),s_1,\alpha} \le_{n,S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\beth_{n+2}(M,\alpha)}\,, \tag{6.208}$$

$$|\partial_i R_{\omega,M}(\varphi,t)[\hat\imath]|_{-\frac{M}{2},s_1,\alpha} \le_{M,S,\alpha} \varepsilon\|\hat\imath\|_{s_1+\sigma+\beth_{M+3}(M,\alpha)}\,, \tag{6.209}$$

$$\|\partial_{\varphi_j}^\beta[\partial_i W_\omega(\varphi,t)[\hat\imath],\partial_x]\|_{\mathcal{L}(H^{s_1})}\,,\ \|\partial_{\varphi_j}^\beta\partial_i W_\omega(\varphi,t)[\hat\imath]\|_{\mathcal{L}(H^{s_1})} \le_{M,S} \varepsilon\|\hat\imath\|_{s_1+\sigma+\frac32 M+\daleth(M+2)+\beta}\,. \tag{6.210}$$

We summarize the whole section in the next proposition:

Proposition 6.26. *Let $a(\varphi,x)$ be as in (6.153) and $a_0(\varphi)$ in (6.169). Then the conjugated operator $\mathcal{L}_M^{(1)}$ in (6.133) is real, even, reversible and has the form*

$$\mathcal{L}_M^{(1)} = \omega\cdot\partial_\varphi \mathbb{I}_2 + \mathrm{i}\mathtt{m}_3 \mathbf{T}(D) + \mathrm{i}\big(\mathbf{C}_1(\varphi,x) + \mathbf{C}_0(\varphi,x)\mathcal{H}\big)|D|^{\frac12} + \mathbf{R}_M^{(1)} + \mathbf{Q}_M^{(1)} \tag{6.211}$$

where $\mathbf{C}_1(\varphi,x)$, $\mathbf{C}_0(\varphi,x)$ are defined in (6.164), the function a_{14} satisfies (6.166), and

$$\mathbf{R}_M^{(1)} := \begin{pmatrix} \mathcal{R}_M^{(1)} & 0 \\ 0 & \overline{\mathcal{R}_M^{(1)}} \end{pmatrix}, \quad \mathbf{Q}_M^{(1)} := \begin{pmatrix} 0 & \mathcal{Q}_M^{(1)} \\ \overline{\mathcal{Q}_M^{(1)}} & 0 \end{pmatrix}.$$

For all $\beta\in\mathbb{N}$, $\beta + k_0 + 4 \le M$, the operators $\partial_{\varphi_j}^\beta \mathcal{R}_M^{(1)}$, $\partial_{\varphi_j}^\beta[\mathcal{R}_M^{(1)},\partial_x]$, $\partial_{\varphi_j}^\beta \mathcal{Q}_M^{(1)}$, $\partial_{\varphi_j}^\beta[\mathcal{Q}_M^{(1)},\partial_x]$, $j=1,\dots,\nu$ are $\mathcal{D}^{k_0}$-tame with tame constants satisfying for all $s_0\le s\le S$

$$\mathfrak{M}_{\partial_{\varphi_j}^\beta[\mathcal{R},\partial_x]}(s)\,,\, \mathfrak{M}_{\partial_{\varphi_j}^\beta\mathcal{R}}(s) \le_{M,S} \varepsilon(1+\|\mathfrak{I}_0\|_{s+\sigma+\frac32 M+\daleth(M+2)+\beta}^{k_0,\gamma})\,, \quad \mathcal{R}\in\{\mathcal{R}_M^{(1)},\mathcal{Q}_M^{(1)}\} \tag{6.212}$$

where the constant $\daleth(M+2)$ is defined by (6.197). Moreover if the constant μ in (6.10) satisfies

$$s_1 + \sigma + \chi M + \daleth(M+2) + \beta \le s_0 + \mu\,, \tag{6.213}$$

then each $\mathcal{R}\in\{\mathcal{R}_M^{(1)},\mathcal{Q}_M^{(1)}\}$ satisfies

$$\|\partial_{\varphi_j}^\beta[\partial_i\mathcal{R}[\hat\imath],\partial_x]\|_{\mathcal{L}(H^{s_1})}\,,\, \|\partial_{\varphi_j}^\beta\partial_i\mathcal{R}[\hat\imath]\|_{\mathcal{L}(H^{s_1})} \le_{M,S} \varepsilon\|\hat\imath\|_{s_1+\sigma+\frac32 M+\daleth(M+2)+\beta}\,. \tag{6.214}$$

Proof. It remains only to prove (6.212) and (6.214).

Proof of (6.212). We estimate each term in (6.174). Let $\partial_\varphi := \partial_{\varphi_j}$, $j=1,\dots,\nu$. The estimates (6.177), (6.190), (6.206) imply $|r_M^{(1)}(x,D)|_{0,s,\alpha}^{k_0,\gamma} \le_{S,\alpha} \varepsilon(1+\|\mathfrak{I}_0\|_{s+\sigma+\beth_{M+2}(M,\alpha)}^{k_0,\gamma})$. Now since $\partial_\varphi^\beta[\partial_\lambda^k \mathrm{Op}(r_M^{(1)}),\partial_x] = \partial_\lambda^k\mathrm{Op}(\partial_\varphi^\beta\partial_x r_M^{(1)})$, we get

$$\begin{aligned}|\partial_\varphi^\beta[\partial_\lambda^k\mathrm{Op}(r_M^{(1)}),\partial_x]|_{0,s,0} &\lessdot \gamma^{-|k|}|\mathrm{Op}(\partial_\varphi^\beta(\partial_x r_M^{(1)}))|_{0,s,0}^{k_0,\gamma} \lessdot \gamma^{-|k|}|\mathrm{Op}(r_M^{(1)})|_{0,s+\beta+1,0}^{k_0,\gamma} \\ &\le_S \varepsilon(1+\|\mathfrak{I}_0\|_{s+\sigma+\daleth(M+2)+\beta}^{k_0,\gamma})\,.\end{aligned}$$

Hence the operator $r_M^{(1)}(\varphi,x,D)$ satisfies the estimate (6.212).

The lemma follows by the estimates (6.196), (6.207). The proof of (6.212) for $\mathcal{Q}_M^{(1)}$ is similar. It follows by (6.122) (for $\alpha=0$) and Lemma 9.10 using the same strategy for proving (6.196) in Lemma 6.24.

Proof of (6.214). It follows by differentiating with respect to i the expression of $\mathcal{R}_M^{(1)}$ in (6.174) and by applying the estimates (6.179), (6.194), (6.199), (6.208), (6.210). ☐

6.7 Space reduction of the order $|D|^{\frac12}$

The aim of this section is to eliminate the x-dependence of the coefficient in front of $|D|^{\frac12}$ in the operator $\mathcal{L}_M^{(1)}$ in (6.211) (where we have neglected the term (6.165)) and $\Pi_{K_n}\mathbf{C}_1 := \begin{pmatrix}\Pi_{K_n}a_{14} & 0 \\ 0 & -\Pi_{K_n}a_{14}\end{pmatrix}$.

We conjugate $\mathcal{L}_M^{(1)}$ by means of a real operator of the form

$$\mathbf{V} := \begin{pmatrix}\mathcal{V} & 0 \\ 0 & \overline{\mathcal{V}}\end{pmatrix}, \qquad \mathcal{V} := \mathrm{Op}(v)\,, \qquad v := v(\varphi,x,\xi)\in S^0\,. \tag{6.215}$$

Setting $\mathbf{\Sigma} := \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$ and recalling that $\mathtt{m}_{1,K_n}$ is defined by (6.166), we compute

$$\begin{aligned}\mathcal{L}_M^{(1)}\mathbf{V} - \mathbf{V}\big(\omega\cdot\partial_\varphi\mathbb{I}_2 + \mathrm{i}\mathtt{m}_3\mathbf{T}(D) + \mathrm{i}\mathtt{m}_{1,K_n}\mathbf{\Sigma}|D|^{\frac12}\big) &= \mathrm{i}\mathtt{m}_3[\mathbf{T}(D),\mathbf{V}] + \mathrm{i}\big((\Pi_{K_n}\mathbf{C}_1 + \mathbf{C}_0\mathcal{H})|D|^{\frac12}\mathbf{V} - \mathtt{m}_{1,K_n}\mathbf{V}\mathbf{\Sigma}|D|^{\frac12}\big) \\ &\quad + (\omega\cdot\partial_\varphi\mathbf{V}) + \big(\mathbf{R}_M^{(1)} + \mathbf{Q}_M^{(1)}\big)\mathbf{V}\,.\end{aligned} \tag{6.216}$$

By (6.63), (6.57) and (2.28), the commutator has the expansion

$$\mathtt{im}_3[\mathbf{T}(D),\mathbf{V}] = \begin{pmatrix} \mathtt{im}_3[T(D),\mathcal{V}] & 0 \\ 0 & -\mathtt{im}_3[T(D),\overline{\mathcal{V}}] \end{pmatrix}, \quad \mathtt{im}_3[T(D),\mathcal{V}] = \mathtt{m}_3 \mathrm{Op}\big(\partial_\xi T(\xi) v_x\big) + r_{T,\mathcal{V}}(x,D)$$

with $r_{T,\mathcal{V}}(x,D) \in OPS^{-\frac12}$. Similarly (recall (6.164)) the operator

$$\mathrm{i}\big(\Pi_{K_n}\mathbf{C}_1 + \mathbf{C}_0\mathcal{H}\big)|D|^{\frac12}\mathbf{V} = \begin{pmatrix} \mathrm{i}(\Pi_{K_n}a_{14} + a_{12}\mathcal{H})|D|^{\frac12}\mathcal{V} & 0 \\ 0 & -\mathrm{i}(\Pi_{K_n}a_{14} + a_{12}\mathcal{H})|D|^{\frac12}\overline{\mathcal{V}} \end{pmatrix}$$

has the expansion

$$\mathrm{i}(\Pi_{K_n}a_{14} + a_{12}\mathcal{H})|D|^{\frac12}\mathcal{V} = \mathrm{Op}\big(\big(\mathrm{i}\Pi_{K_n}a_{14} + a_{12}\mathrm{sign}(\xi)\big)|\xi|^{\frac12}\chi(\xi)v\big) + \mathfrak{r}_{\mathcal{V}}(x,D) \tag{6.217}$$

with $\mathfrak{r}_{\mathcal{V}}(x,D) \in OPS^{-\frac12}$. In addition

$$\mathtt{im}_{1,K_n}\mathbf{V}\mathbf{\Sigma}|D|^{\frac12} = \begin{pmatrix} \mathrm{Op}\big(\mathtt{im}_{1,K_n} v\,\chi(\xi)|\xi|^{\frac12}\big) & 0 \\ 0 & \overline{\mathrm{Op}\big(\mathtt{im}_{1,K_n} v\,\chi(\xi)|\xi|^{\frac12}\big)} \end{pmatrix}. \tag{6.218}$$

By (6.217), (6.218) and decomposing the cut-off function $\chi(\xi) = \chi_0(\xi) + (\chi(\xi) - \chi_0(\xi))$ where χ_0 is the cut-off function defined in (6.100), we get

$$\mathrm{i}\big((\Pi_{K_n}a_{14} + a_{12}\mathcal{H})|D|^{\frac12}\mathcal{V} - \mathtt{m}_{1,K_n}\mathcal{V}|D|^{\frac12}\big) = \mathrm{Op}\big(\big(\mathrm{i}(\Pi_{K_n}a_{14} - \mathtt{m}_{1,K_n}) + a_{12}\mathrm{sign}(\xi)\big)|\xi|^{\frac12}\chi_0(\xi)v\big) + r_{\mathcal{V}}(x,D)$$

where

$$r_{\mathcal{V}}(x,D) := \mathfrak{r}_{\mathcal{V}}(x,D) + \mathrm{Op}\big(\big(\mathrm{i}\Pi_{K_n}a_{14} + a_{12}\mathrm{sign}(\xi) - \mathtt{im}_{1,K_n}\big)|\xi|^{\frac12}(\chi(\xi) - \chi_0(\xi))v\big) \in OPS^{-\frac12}$$

noting that $\big(\mathrm{i}\Pi_{K_n}a_{14} + a_{12}\mathrm{sign}(\xi) - \mathtt{im}_{1,K_n}\big)|\xi|^{\frac12}(\chi(\xi) - \chi_0(\xi))v \in S^{-\infty}$ because $\chi(\xi) - \chi_0(\xi) = 0$ for $|\xi| \geq 3/4$. Therefore we have to solve the equation

$$\mathtt{m}_3\partial_\xi T(\xi)v_x + \big(\mathrm{i}(\Pi_{K_n}a_{14} - \mathtt{m}_{1,K_n}) + a_{12}\mathrm{sign}(\xi)\big)\chi_0(\xi)|\xi|^{\frac12}v = 0\,. \tag{6.219}$$

We look for a solution of (6.219) of the form

$$v := v(\varphi,x,\xi) := \exp(p(\varphi,x,\xi))\,, \qquad p := p(\varphi,x,\xi) \in S^0\,. \tag{6.220}$$

Thus, from (6.219), the symbol p has to solve

$$\mathtt{m}_3\partial_\xi T(\xi)p_x(\varphi,x,\xi) = -\big(\mathrm{i}(\Pi_{K_n}a_{14}(\varphi,x) - \mathtt{m}_{1,K_n}) + a_{12}(\varphi,x)\mathrm{sign}(\xi)\big)\chi_0(\xi)|\xi|^{\frac12}\,. \tag{6.221}$$

The right hand side in (6.221) has zero average in x by (6.166) and because a_{12} is odd in x, by (6.76), (6.64) and remark 6.5. By (6.57) the derivative

$$\partial_\xi T(\xi) = \begin{cases} \dfrac{\chi(\xi)\,\mathrm{sign}(\xi)(1+3\kappa\xi^2)}{2|\xi|^{1/2}(1+\kappa\xi^2)^{1/2}} + \partial_\xi\chi(\xi)|\xi|^{\frac12}(1+\kappa|\xi|^2)^{\frac12} \in S^{1/2} & \text{if} \quad |\xi| > \frac13 \\ 0 & \text{if} \quad |\xi| \leq \frac13\,. \end{cases}$$

Since the symbol $T(\xi)$ is even in ξ, the derivative $\partial_\xi T(\xi)$ is odd. Moreover, by (2.26), $\partial_\xi\chi(\xi) > 0$ for all $1/3 < \xi < 2/3$, and so $|\partial_\xi T(\xi)| > 0$ for all $|\xi| > 1/3$ and $|\partial_\xi T(\xi)| > c > 0$ for all $|\xi| \geq 1/2$. Therefore (6.221) admits the solution

$$p(\varphi,x,\xi) := \begin{cases} -\dfrac{|\xi|^{\frac12}\chi_0(\xi)}{\mathtt{m}_3\partial_\xi T(\xi)}\partial_x^{-1}\Big(\mathrm{i}(\Pi_{K_n}a_{14}(\varphi,x) - \mathtt{m}_{1,K_n}) + a_{12}(\varphi,x)\mathrm{sign}(\xi)\Big) & \text{if} \quad |\xi| > \frac12 \\ 0 & \text{if} \quad |\xi| \leq \frac12\,. \end{cases} \tag{6.222}$$

Since $p(-\varphi, x, -\xi) = \overline{p(\varphi, x, \xi)}$ and $p(\varphi, -x, -\xi) = p(\varphi, x, \xi)$, then $\mathbf{V}$ is reversibility preserving and $\mathbf{V}$ is even, by Lemma 2.4. As a consequence (6.216)-(6.219) imply that

$$\mathbf{V}^{-1}\mathcal{L}_M^{(1)}\mathbf{V} = \omega\cdot\partial_\varphi \mathbb{I}_2 + \mathrm{i}\mathtt{m}_3\mathbf{T}(D) + \mathrm{i}\mathtt{m}_{1,K_n}\mathbf{\Sigma}|D|^{\frac{1}{2}} + \mathbf{R}_M^{(2)} + \mathbf{Q}_M^{(2)} \tag{6.223}$$

with block-diagonal terms

$$\mathbf{R}_M^{(2)} := \begin{pmatrix} \mathcal{R}_M^{(2)} & 0 \\ 0 & \overline{\mathcal{R}}_M^{(2)} \end{pmatrix}, \quad \mathbf{Q}_M^{(2)} := \begin{pmatrix} 0 & \mathcal{Q}_M^{(2)} \\ \overline{\mathcal{Q}}_M^{(2)} & 0 \end{pmatrix}$$

$$\mathcal{R}_M^{(2)} := \mathcal{V}^{-1}\big(r_{T,\mathcal{V}}(x,D) + r_{\mathcal{V}}(x,D) + \omega\cdot\partial_\varphi \mathcal{V} + \mathcal{R}_M^{(1)}\mathcal{V}\big), \quad \mathcal{Q}_M^{(2)} := \mathcal{V}^{-1}\mathcal{Q}_M^{(1)}\overline{\mathcal{V}}. \tag{6.224}$$

Finally we define the real, even and reversible operator

$$\mathcal{L}_M^{(2)} := \omega\cdot\partial_\varphi \mathbb{I}_2 + \mathrm{i}\mathtt{m}_3\mathbf{T}(D) + \mathrm{i}\mathtt{m}_1\mathbf{\Sigma}|D|^{\frac{1}{2}} + \mathbf{R}_M^{(2)} + \mathbf{Q}_M^{(2)} \tag{6.225}$$

where the coefficient

$$\mathtt{m}_1 := -\frac{(2\pi)^{-\nu-\frac{5}{2}}}{2\sqrt{\kappa}} \int_{\mathbb{T}^{\nu+1}} (1+\beta_x)[\omega\cdot\partial_\varphi\beta + V(1+\beta_x)]^2 \Big(\int_{\mathbb{T}} \sqrt{1+\eta_y^2}\, dy\Big)^{3/2} d\varphi\, dx \tag{6.226}$$

substitutes $\mathtt{m}_{1,K_n}$ in (6.223), i.e.

$$\mathbf{V}^{-1}\mathcal{L}_M^{(1)}\mathbf{V} = \mathcal{L}_M^{(2)} + \mathbf{R}_{\mathtt{m}_1}^{\perp}, \quad \mathbf{R}_{\mathtt{m}_1}^{\perp} := \mathrm{i}(\mathtt{m}_{1,K_n} - \mathtt{m}_1)\mathbf{\Sigma}|D|^{\frac{1}{2}}. \tag{6.227}$$

The term $\mathbf{R}_{\mathtt{m}_1}^{\perp}$ will contribute to the remainder $\mathbf{R}_{\omega}^{\perp}$ in the estimates (7.94)-(7.95).

Lemma 6.27. *$|\mathtt{m}_1 - \mathtt{m}_{1,K_n}|^{k_0,\gamma} \leq C\varepsilon K_n^{-b}$, $\forall b > 0$.*

Proof. By (6.170), (6.226) one has

$$\mathtt{m}_1 - \mathtt{m}_{1,K_n} = \frac{(2\pi)^{-\nu-\frac{5}{2}}}{2\sqrt{\kappa}} \int_{\mathbb{T}^{\nu}} (1+\beta_x)[\omega\cdot\partial_\varphi\beta + V(1+\beta_x)]^2 \Pi_{K_n}^{\perp}\Big(\int_{\mathbb{T}} \sqrt{1+\eta_y^2}\, dy\Big)^{3/2} d\varphi dx\,.$$

Then the lemma follows by (6.18), (6.33), (6.43), (6.13), (6.10), using the smoothing property (2.8). □

Lemma 6.28. *The coefficient $\mathtt{m}_1$ defined in (6.226) satisfies, for some $\sigma := \sigma(\tau,\nu,k_0) > 0$, the estimates*

$$|\mathtt{m}_1|^{k_0,\gamma} \leq C\varepsilon, \quad |\partial_i \mathtt{m}_1[\hat{\imath}]| \leq C\varepsilon\|\hat{\imath}\|_\sigma. \tag{6.228}$$

The operator $\mathbf{V}$ defined in (6.215) is real, even, reversibility preserving and $\mathcal{V} = \mathrm{Op}(v(\varphi,x,\xi)) \in OPS^0$ with symbol $v(\varphi,x,\xi) \in S^0$ defined in (6.220) and (6.222), satisfies, for all $s_0 \leq s \leq S$,

$$|\!|\!|\mathcal{V}^{\pm 1} - \mathrm{Id}|\!|\!|_{0,s,0}^{k_0,\gamma}, |\!|\!|(\mathcal{V}^{\pm 1} - \mathrm{Id})^*|\!|\!|_{0,s,0}^{k_0,\gamma} \leq_S \varepsilon(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}). \tag{6.229}$$

For all $\beta \in \mathbb{N}$, $\beta + k_0 + 4 \leq M$, the operators $\partial_{\varphi_j}^\beta \mathcal{R}_M^{(2)}$, $\partial_{\varphi_j}^\beta[\mathcal{R}_M^{(2)}, \partial_x]$, $\partial_{\varphi_j}^\beta \mathcal{Q}_M^{(2)}$, $\partial_{\varphi_j}^\beta[\mathcal{Q}_M^{(2)}, \partial_x]$ are $\mathcal{D}^{k_0}$-tame and the tame constants $\mathfrak{M}_{\partial_{\varphi_j}^\beta[\mathcal{R},\partial_x]}(s)$, $\mathfrak{M}_{\partial_{\varphi_j}^\beta \mathcal{R}}(s)$, $\mathcal{R} \in \{\mathcal{R}_M^{(2)}, \mathcal{Q}_M^{(2)}\}$ satisfy (6.212) (with a possibly larger $\sigma := \sigma(\tau,\nu,k_0) > 0$).
Moreover if the constant μ in (6.10) satisfies (6.213) (with a possibly larger $\sigma := \sigma(\tau,\nu,k_0) > 0$) then

$$|\!|\!|\partial_i \mathcal{V}^{\pm 1}[\hat{\imath}]|\!|\!|_{0,s_1,0}, |\!|\!|\partial_i(\mathcal{V}^{\pm 1})^*[\hat{\imath}]|\!|\!|_{0,s_1,0} \leq_S \varepsilon\|\hat{\imath}\|_{s_1+\sigma}, \tag{6.230}$$

and the remainders $\mathcal{R}_M^{(2)}$, $\mathcal{Q}_M^{(2)}$ satisfy the estimates (6.214). The operators $\mathcal{R}_M^{(2)}$, $\mathcal{Q}_M^{(2)}$ are reversible.

Proof. The estimate (6.228) follows by (6.226), (6.18), (6.19), (6.33), (6.43), (6.46), (6.13), (6.10). The estimates (6.229), (6.230) for $\mathcal{V}^{\pm 1}$ follows by (6.215), (6.220), (6.222) and Lemma 2.10. The estimates for $(\mathcal{V}^{\pm 1} - \mathrm{Id})^*$ and $\partial_i(\mathcal{V}^{\pm 1})^*$ follow by Lemma 2.9. Using Lemma 2.6 we get $|\!|\!|r_{T,\mathcal{V}}(x,D)|\!|\!|_{0,s,0}^{k_0,\gamma}, |\!|\!|r_{\mathcal{V}}(x,D)|\!|\!|_{0,s,0}^{k_0,\gamma} \leq_S \varepsilon(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})$, $|\!|\!|\partial_i r_{T,\mathcal{V}}(x,D)[\hat{\imath}]|\!|\!|_{0,s_1,0}, |\!|\!|\partial_i r_{\mathcal{V}}(x,D)[\hat{\imath}]|\!|\!|_{0,s_1,0} \leq_S \varepsilon\|\hat{\imath}\|_{s_1+\sigma}$, for some $\sigma := \sigma(\tau,\nu,k_0) > 0$. The term $\mathcal{V}^{-1}\mathcal{R}_M^{(1)}\mathcal{V}$ in (6.224) is estimated following the same strategy of Lemma 6.24. □

6.8 Conclusion: partial reduction of $\mathcal{L}_\omega$

By sections 6.1-6.7 the linear operator $\mathcal{L}$ in (6.8) is semi conjugated to the real, even and reversible operator $\mathcal{L}_M^{(2)}$ defined in (6.225), up to operators which are supported on high Fourier frequencies, namely

$$\mathcal{L}_M^{(2)} = \mathcal{W}_2^{-1}\mathcal{L}\mathcal{W}_1 + \mathbf{R}_M^{(2),\perp} + \mathbf{R}_{\pi_0} \tag{6.231}$$

$$\mathbf{R}_M^{(2),\perp} := -\mathbf{V}^{-1}\mathbf{\Phi}\mathbf{\Phi}_M^{-1}\mathbf{R}_4^\perp\mathbf{\Phi}_M\mathbf{\Phi}^{-1}\mathbf{V} - \mathbf{V}^{-1}\mathbf{R}_M^{(1),\perp}\mathbf{V} - \mathbf{R}_{\mathtt{m}_1}^\perp\,, \tag{6.232}$$

$$\mathbf{R}_{\pi_0} := -\mathbf{V}^{-1}\mathbf{\Phi}\mathbf{\Phi}_M^{-1}\rho^{-1}(\mathcal{P}^{-1}\mathbb{I}_2)(\mathrm{i}m_3(\varphi)\Pi_0)(\mathcal{P}\mathbb{I}_2)\mathbf{\Phi}_M\mathbf{\Phi}^{-1}\mathbf{V} \tag{6.233}$$

where

$$\mathcal{W}_1 := \mathcal{Z}\mathcal{B}\mathcal{Q}\mathcal{S}(\mathcal{P}\mathbb{I}_2)\mathbf{\Phi}_M\mathbf{\Phi}^{-1}\mathbf{V}\,,\ \ \mathcal{W}_2 := \mathcal{Z}\mathcal{B}\mathcal{Q}\mathcal{S}(\mathcal{P}\mathbb{I}_2)\rho\,\mathbf{\Phi}_M\mathbf{\Phi}^{-1}\mathbf{V}\,, \tag{6.234}$$

and $\mathbf{R}_4^\perp, \mathbf{R}_M^{(1),\perp}, \mathbf{R}_{\mathtt{m}_1}^\perp$ are defined respectively in (6.80), (6.165), (6.227) (they will contribute to the remainders in (7.94)-(7.95)) and the operator Π_0 is defined in (6.66). The maps $\mathcal{W}_1$, $\mathcal{W}_2$ are real, even and reversibility preserving.
Let $\mathbb{S} = \mathbb{S}^+ \cup (-\mathbb{S}^+)$ and $\mathbb{S}_0 := \mathbb{S} \cup \{0\}$. We denote by $\Pi_{\mathbb{S}_0}$ the corresponding L^2-orthogonal projection and $\Pi_{\mathbb{S}_0}^\perp := \mathrm{Id} - \Pi_{\mathbb{S}_0}$.

Lemma 6.29. *Assume* (6.10). *For $\varepsilon\gamma^{-1}$ small enough, the operators*

$$\mathcal{W}_1^\perp := \Pi_{\mathbb{S}_0}^\perp \mathcal{W}_1 \Pi_{\mathbb{S}_0}^\perp\,, \qquad \mathcal{W}_2^\perp := \Pi_{\mathbb{S}_0}^\perp \mathcal{W}_2 \Pi_{\mathbb{S}_0}^\perp\,, \tag{6.235}$$

are invertible and for all $s_0 \le s \le S$ they satisfy the tame estimates

$$\|\mathcal{W}_n^\perp h\|_s^{k_0,\gamma} + \|(\mathcal{W}_n^\perp)^{-1}h\|_s^{k_0,\gamma} \le_{M,S} \|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_M(0)}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\,, \qquad n = 1,2\,, \tag{6.236}$$

for some $\sigma := \sigma(\tau,\nu) > 0$.
Moreover if the constant μ in (6.10) *satisfies $s_1 + \sigma + \aleph_M(0) \le s_0 + \mu$ for some $\sigma := \sigma(\tau,\nu,k_0) > 0$, then*

$$\|\partial_i \mathcal{W}_n^{\pm 1}[\hat\imath]h\|_{s_1}\,,\ \|\partial_i(\mathcal{W}_n^\perp)^{\pm 1}[\hat\imath]h\|_{s_1} \le_{M,S} \|\hat\imath\|_{s_1+\sigma+\aleph_M(0)}\|h\|_{s_1+\sigma}\,. \tag{6.237}$$

Proof. By Lemmata 2.12, 2.14 and by the estimates of sections 6.1-6.7, the operators $\mathcal{W}_1, \mathcal{W}_2$ are invertible and satisfy tame estimates $\|\mathcal{W}_1^{\pm 1}h\|_s^{k_0,\gamma} \le_S \|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_M(0)}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}$ where $\aleph_M(0)$ is given in Proposition 6.11. In order to prove that $\mathcal{W}_1^\perp$ is invertible, it is sufficient to prove that $\Pi_{\mathbb{S}_0}\mathcal{W}_1\Pi_{\mathbb{S}_0}$ is invertible. This follows by a perturbative argument, for $\varepsilon\gamma^{-1}$ small, as in [8] using that $\Pi_{\mathbb{S}_0}$ is a finite dimensional projector. ☐

Finally, the operator $\mathcal{L}_\omega$ defined in (5.40) (i.e. (6.7)) is semi-conjugated to

$$(\mathcal{W}_2^\perp)^{-1}\mathcal{L}_\omega\mathcal{W}_1^\perp = \Pi_{\mathbb{S}_0}^\perp\mathcal{L}_M^{(2)}\Pi_{\mathbb{S}_0}^\perp - \Pi_{\mathbb{S}_0}^\perp\mathbf{R}_M^{(2),\perp}\Pi_{\mathbb{S}_0}^\perp + R_M$$

where $\Pi_{\mathbb{S}_0}^\perp\mathbf{R}_M^{(2),\perp}\Pi_{\mathbb{S}_0}^\perp$ is supported on the high Fourier modes and

$$R_M := (\mathcal{W}_2^\perp)^{-1}\Pi_{\mathbb{S}_0}^\perp\big(\mathcal{W}_2\Pi_{\mathbb{S}_0}\mathcal{L}_M^{(2)}\Pi_{\mathbb{S}_0}^\perp - \mathcal{W}_2\Pi_{\mathbb{S}_0}\mathbf{R}_M^{(2),\perp}\Pi_{\mathbb{S}_0}^\perp - \mathcal{L}\Pi_{\mathbb{S}_0}\mathcal{W}_1\Pi_{\mathbb{S}_0}^\perp - \mathcal{W}_2\mathbf{R}_{\pi_0}\Pi_{\mathbb{S}_0}^\perp + \varepsilon R\mathcal{W}_1^\perp\big) \tag{6.238}$$

is a finite dimensional operator.

Lemma 6.30. *The operator R_M has the finite dimensional form* (6.3)-(6.4).

Proof. We analyze the term $(\mathcal{W}_2^\perp)^{-1}R\mathcal{W}_1^\perp$ in (6.238). The others are similar. Since R has the form (6.3), it is sufficient to prove that, given $\mathcal{R} : h \to (h,g)_{L_x^2}\chi$, the operator $(\mathcal{W}_2^\perp)^{-1}\mathcal{R}\mathcal{W}_1^\perp$ has the form (6.3) as well. We use the following property: given a scalar function $a : \mathbb{T}^\nu \to \mathbb{C}$ and $\chi := \chi(\varphi,\cdot) \in H_{\mathbb{S}_0}^\perp$, we have

$$(\mathcal{W}_i^\perp)^{\pm 1}[a(\varphi)\chi] = (\mathcal{P}^{\pm 1}a)(\varphi)(\mathcal{W}_i^\perp)^{\pm 1}[\chi]\,. \tag{6.239}$$

Let us prove (6.239) for $\mathcal{W}_2^\perp$. We write (recall (6.235) and (6.234))

$$\mathcal{W}_2^\perp = \Pi_{\mathbb{S}_0}^\perp\big(\mathbf{\Gamma}_1\mathcal{P}\mathbb{I}_2\rho\mathbf{\Gamma}_2\big)\Pi_{\mathbb{S}_0}^\perp \qquad \text{where} \qquad \mathbf{\Gamma}_1 := \mathcal{Z}\mathcal{B}\mathcal{Q}\mathcal{S}\,, \quad \mathbf{\Gamma}_2 := \mathbf{\Phi}_M\mathbf{\Phi}^{-1}\mathbf{V}\,,$$

are, for any $\varphi \in \mathbb{T}^\nu$, linear operators $\mathbf{\Gamma}_i(\varphi) : H_{\mathbb{S}_0}^\perp \to H_{\mathbb{S}_0}^\perp$ of the phase space. Then

$$\begin{aligned}\mathcal{W}_2^\perp[a(\varphi)\chi] &= \Pi_{\mathbb{S}_0}^\perp\big(\mathbf{\Gamma}_1\mathcal{P}\mathbb{I}_2\rho\mathbf{\Gamma}_2\big)\Pi_{\mathbb{S}_0}^\perp[a(\varphi)\chi] = \Pi_{\mathbb{S}_0}^\perp\mathbf{\Gamma}_1\mathcal{P}\mathbb{I}_2[a(\varphi)\rho\mathbf{\Gamma}_2\Pi_{\mathbb{S}_0}^\perp[\chi]] \\ &= \Pi_{\mathbb{S}_0}^\perp\mathbf{\Gamma}_1[(\mathcal{P}a)(\varphi)(\mathcal{P}\mathbb{I}_2\rho\mathbf{\Gamma}_2\Pi_{\mathbb{S}_0}^\perp[\chi])] = (\mathcal{P}a)(\varphi)\Pi_{\mathbb{S}_0}^\perp\mathbf{\Gamma}_1\mathcal{P}\mathbb{I}_2\rho\mathbf{\Gamma}_2\Pi_{\mathbb{S}_0}^\perp[\chi] = (\mathcal{P}a)(\varphi)\mathcal{W}_2^\perp[\chi]\,.\end{aligned} \tag{6.240}$$

Then (6.239) follows also for $(\mathcal{W}_2^\perp)^{-1}$. Denoting $\tilde a := \mathcal{P}^{-1}a$ and $\tilde\chi := (\mathcal{W}_2^\perp)^{-1}[\chi]$, we have

$$(\mathcal{W}_2^\perp)^{-1}[a(\varphi)\chi] = (\mathcal{W}_2^\perp)^{-1}[(\mathcal{P}\tilde a)(\varphi)(\mathcal{W}_2^\perp\tilde\chi)] \overset{(6.240)}{=} (\mathcal{W}_2^\perp)^{-1}\mathcal{W}_2^\perp[\tilde a(\varphi)\tilde\chi] = (\mathcal{P}^{-1}a)(\varphi)(\mathcal{W}_2^\perp)^{-1}[\chi]\,.$$

Now for any $h(\varphi,\cdot) \in H_{\mathbb{S}_0}^\perp$ one has

$$(\mathcal{W}_2^\perp)^{-1}\mathcal{R}\mathcal{W}_1^\perp[h] = (\mathcal{W}_2^\perp)^{-1}\big[(\mathcal{W}_1^\perp[h],g)_{L^2_x}\chi\big] \overset{(6.239)}{=} \big(\mathcal{P}^{-1}(\mathcal{W}_1^\perp[h],g)_{L^2_x}\big)\chi_* \tag{6.241}$$

with $\chi_* := (\mathcal{W}_2^\perp)^{-1}[\chi]$ and

$$\begin{aligned}\mathcal{P}^{-1}(\mathcal{W}_1^\perp[h],g)_{L^2_x} &= \mathcal{P}^{-1}(\Pi_{\mathbb{S}_0}^\perp\mathbf{\Gamma}_1\mathcal{P}\mathbb{I}_2\mathbf{\Gamma}_2\Pi_{\mathbb{S}_0}^\perp[h],g)_{L^2_x} = \mathcal{P}^{-1}(\mathcal{P}\mathbb{I}_2\mathbf{\Gamma}_2\Pi_{\mathbb{S}_0}^\perp[h],\mathbf{\Gamma}_1^*\Pi_{\mathbb{S}_0}^\perp g)_{L^2_x} \\ &= (\mathbf{\Gamma}_2\Pi_{\mathbb{S}_0}^\perp[h],\mathcal{P}^{-1}\mathbf{\Gamma}_1^*\Pi_{\mathbb{S}_0}^\perp g)_{L^2_x} = (h,\Pi_{\mathbb{S}_0}^\perp\mathbf{\Gamma}_2^*\mathcal{P}^{-1}\mathbf{\Gamma}_1^*\Pi_{\mathbb{S}_0}^\perp g)_{L^2_x} = (h,g_*)_{L^2_x}\end{aligned} \tag{6.242}$$

with $g_* := \Pi_{\mathbb{S}_0}^\perp\mathbf{\Gamma}_2^*\mathcal{P}^{-1}\mathbf{\Gamma}_1^*\Pi_{\mathbb{S}_0}^\perp g$. By (6.241) and (6.242) the lemma follows. □

In conclusion we write

$$\mathcal{L}_\omega = \mathcal{W}_2^\perp\mathcal{L}_M^{(3)}(\mathcal{W}_1^\perp)^{-1} + \mathbf{R}_M^{(3),\perp}\,, \quad \mathcal{L}_M^{(3)} := \mathcal{L}_M^{(2)} + R_M\,, \quad \mathbf{R}_M^{(3),\perp} := -\mathcal{W}_2^\perp\mathbf{R}_M^{(2),\perp}(\mathcal{W}_1^\perp)^{-1} \tag{6.243}$$

where $\mathcal{L}_M^{(2)}$ is defined in (6.225), $\mathbf{R}_M^{(2),\perp}$ is defined in (6.232) and R_M in (6.238). The remainder $\mathbf{R}_M^{(3),\perp}$ satisfies tame estimates: there is $\sigma := \sigma(\tau,\nu,k_0) > 0$ such that

$$\|\mathbf{R}_M^{(3),\perp}h\|_{s_0}^{k_0,\gamma} \le_S \varepsilon K_n^{-b}\big(\|h\|_{s_0+\sigma+b}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s_0+\sigma+\aleph_M(0)+b}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,, \quad \forall b > 0\,, \tag{6.244}$$

$$\|\mathbf{R}_M^{(3),\perp}h\|_{s}^{k_0,\gamma} \le_S \varepsilon\big(\|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_M(0)}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,, \quad \forall s_0 \le s \le S\,. \tag{6.245}$$

The estimates (6.244), (6.245) follow by (6.243), (6.231), (6.80), (6.165), (6.227), using the estimates (6.43), (6.175), (6.236), (6.229), (6.119), (2.8), Lemma 6.27 and Proposition 9.11.

Proposition 6.31. *Assume* (6.10). *For all* $(\omega,\kappa) \in \mathtt{DC}_{K_n}^\gamma \times [\kappa_1,\kappa_2]$ *(see* (1.40)*) the operator* $\mathcal{L}_\omega$ *defined in* (5.40) *(i.e.* (6.7)*) is semiconjugated to the real, even and reversible operator* $\mathcal{L}_M^{(3)}$ *in* (6.243) *up to the remainder* $\mathbf{R}_M^{(3),\perp}$ *which satisfies* (6.244)-(6.245). *The operator*

$$\mathcal{L}_M^{(3)} = \Pi_{\mathbb{S}_0}^\perp\big(\omega\cdot\partial_\varphi\mathbb{I}_2 + \mathrm{i}\mathtt{m}_3\mathbf{T}(D) + \mathrm{i}\mathtt{m}_1\mathbf{\Sigma}|D|^{\frac12} + \mathbf{R}_M^{(3)} + \mathbf{Q}_M^{(3)}\big)\Pi_{\mathbb{S}_0}^\perp \tag{6.246}$$

where the constant coefficients $\mathtt{m}_3 := \mathtt{m}_3(\omega,\kappa) \in \mathbb{R}$, $\mathtt{m}_1 := \mathtt{m}_1(\omega,\kappa) \in \mathbb{R}$, *are defined in* (6.72), (6.226) *for all* $(\omega,\kappa) \in \mathbb{R}^\nu \times [\kappa_1,\kappa_2]$, *and satisfy* (6.83), (6.228). *The operator* $\mathbf{T}(D)$ *is defined in* (6.63), (6.57) *and the matrix* $\mathbf{\Sigma} := \begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}$. *The remainders*

$$\mathbf{R}_M^{(3)} := \begin{pmatrix} \mathcal{R}_M^{(3)} & 0 \\ 0 & \overline{\mathcal{R}_M^{(3)}}\end{pmatrix}\,, \quad \mathbf{Q}_M^{(3)} := \begin{pmatrix} 0 & \mathcal{Q}_M^{(3)} \\ \overline{\mathcal{Q}_M^{(3)}} & 0\end{pmatrix} \tag{6.247}$$

satisfy the following tame properties: for all $\beta \in \mathbb{N}$, $\beta + k_0 + 4 \le M$, *the operators* $\partial_{\varphi_j}^\beta\mathcal{R}_M^{(3)}$, $\partial_{\varphi_j}^\beta[\mathcal{R}_M^{(3)},\partial_x]$, $\partial_{\varphi_j}^\beta\mathcal{Q}_M^{(3)}$, $\partial_{\varphi_j}^\beta[\mathcal{Q}_M^{(3)},\partial_x]$, $j = 1,\dots,\nu$, *are* $\mathcal{D}^{k_0}$*-tame and their tame constants satisfy, for all* $s_0 \le s \le S$,

$$\max_{\mathcal{R}\in\{\mathcal{R}_M^{(3)},\mathcal{Q}_M^{(3)}\}}\big\{\mathfrak{M}_{\partial_{\varphi_j}^\beta\mathcal{R}}(s)\,,\mathfrak{M}_{\partial_{\varphi_j}^\beta[\mathcal{R},\partial_x]}(s)\big\} \le_{M,S} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\frac32 M+\daleth(M+2)+\aleph_M(0)+\beta}^{k_0,\gamma}\big) \tag{6.248}$$

for some $\sigma := \sigma(\tau, \nu, k_0) > 0$ where the constant $\aleph_M(0)$, $\daleth(M)$ are defined in (6.96), (6.197)*.*
Moreover if the constant μ in (6.10) *satisfies*

$$s_1 + \sigma + \chi M + \daleth(M+2) + \aleph_M(0) + M - k_0 - 4 \le s_0 + \mu \,, \tag{6.249}$$

then each $\mathcal{R} \in \{\mathcal{R}_M^{(3)}, \mathcal{Q}_M^{(3)}\}$ satisfies, for all $\beta \in \mathbb{N}$, $\beta + k_0 + 4 \le M$,

$$\|\partial_{\varphi_j}^{\beta}[\partial_i \mathcal{R}[\hat\imath], \partial_x]\|_{\mathcal{L}(H^{s_1})}, \|\partial_{\varphi_j}^{\beta}\partial_i \mathcal{R}[\hat\imath]\|_{\mathcal{L}(H^{s_1})} \le_{M,S} \varepsilon\gamma^{-1}\|\hat\imath\|_{s_1+\sigma+\frac{3}{2}M+\daleth(M+2)+\aleph_M(0)+\beta} \,. \tag{6.250}$$

Proof. Note that the coefficients $\mathtt{m}_3$, $\mathtt{m}_1$ in (6.72), (6.226) are actually defined for all the parameters $(\omega, \kappa) \in \mathbb{R}^\nu \times [\kappa_1, \kappa_2]$ since the approximate solution (η, ψ) is defined for all $(\omega, \kappa) \in \mathbb{R}^\nu \times [\kappa_1, \kappa_2]$ at each step of the Nash-Moser iteration in section 8, see the extension Lemma 8.5.

By (6.243), (6.225) and Lemma 6.28, it is enough to prove the estimates (6.248), (6.250) for the operator R_M defined in (6.238). We estimate the term $(\mathcal{W}_2^\perp)^{-1}\Pi_{\mathbb{S}_0}^\perp \mathcal{W}_2 \mathbf{R}_{\pi_0}\Pi_{\mathbb{S}_0}^\perp$, the others are analogous. By (6.233), setting

$$\mathbf{\Gamma}_2 := \mathbf{\Phi}_M \mathbf{\Phi}^{-1}\mathbf{V} \,, \quad \mathbf{\Gamma}_3 := (\mathcal{W}_2^\perp)^{-1}\Pi_{\mathbb{S}_0}^\perp \mathcal{W}_2 \mathbf{V}^{-1}\mathbf{\Phi}\mathbf{\Phi}_M^{-1}\rho^{-1} \,,$$

and recalling (6.69) we write

$$(\mathcal{W}_2^\perp)^{-1}\Pi_{\mathbb{S}_0}^\perp \mathcal{W}_2 \mathbf{R}_{\pi_0}\Pi_{\mathbb{S}_0}^\perp = \mathbf{\Gamma}_3(\mathrm{i}\mathtt{m}_3\Pi_0)\mathbf{\Gamma}_2\Pi_{\mathbb{S}_0}^\perp \quad \text{where} \quad \mathtt{m}_3(\vartheta) := \mathcal{P}^{-1}m_3(\vartheta) = m_3(\vartheta + \omega\tilde p(\vartheta)) \,.$$

Writing $\mathbf{\Gamma}_m = \begin{pmatrix} \Gamma_m^{(1)} & \Gamma_m^{(2)} \\ \overline{\Gamma}_m^{(2)} & \overline{\Gamma}_m^{(1)} \end{pmatrix}$, $m = 2, 3$, and recalling the definition (6.66) of Π_0 and using that $\Pi_0\Pi_{\mathbb{S}_0}^\perp = 0$, we get

$$\mathbf{R} := \mathbf{\Gamma}_3(\mathrm{i}\mathtt{m}_3\Pi_0)\mathbf{\Gamma}_2\Pi_{\mathbb{S}_0}^\perp = \mathbf{\Gamma}_3(\mathrm{i}\mathtt{m}_3\Pi_0)\big(\mathbf{\Gamma}_2 - \mathrm{Id}\big)\Pi_{\mathbb{S}_0}^\perp$$

and then for all $h \in H_{\mathbb{S}_0}^\perp$ we get

$$\mathbf{R}h = \chi(\varphi, x)\big(h(\varphi, \cdot)\,,\, g(\varphi, \cdot)\big)_{L_x^2} \,, \quad \chi := \mathrm{i}\mathbf{\Gamma}_3[\mathtt{m}_3] \in H_{\mathbb{S}_0}^\perp \,, \quad g := \Pi_{\mathbb{S}_0}^\perp(\mathbf{\Gamma}_2 - \mathrm{Id})^*[1] \in H_{\mathbb{S}_0}^\perp \,.$$

Lemma 6.29, the estimates of sections 6.1-6.7 and of Propositions 9.17, 9.18 imply that for some $\sigma := \sigma(k_0, \tau, \nu) > 0$, for all $s \in [s_0, S]$,

$$\|g\|_s^{k_0,\gamma} \le_{S,M} \varepsilon\gamma^{-1}(1 + \|\mathfrak{I}_0\|_{s+\aleph_M(0)+\sigma}^{k_0,\gamma}) \,, \quad \|\chi\|_s^{k_0,\gamma} \le_{S,M} 1 + \|\mathfrak{I}_0\|_{s+\aleph_M(0)+\sigma}^{k_0,\gamma} \,,$$
$$\|\partial_i g[\widehat\imath]\|_{s_1} \le_{S,M} \varepsilon\gamma^{-1}\|\widehat\imath\|_{s_1+\aleph_M(0)+\sigma} \,, \quad \|\partial_i \chi[\widehat\imath]\|_{s_1} \le_{S,M} \|\widehat\imath\|_{s_1+\aleph_M(0)+\sigma},$$

provided (6.249) is satisfied. Then the estimates (6.248), (6.250) for the operator $\mathcal{R}$ follow since for all $j = 1, \ldots, \nu$, $\beta \in \mathbb{N}$, $k \in \mathbb{N}^{\nu+1}$,

$$\partial_{\varphi_j}^{\beta}\partial_\lambda^k[\mathbf{R}, \partial_x]h = -\sum_{\beta_1+\beta_2=\beta, k_1+k_2=k} \big(\partial_\lambda^{k_1}\partial_{\varphi_j}^{\beta_1}\chi\big(h, \partial_\lambda^{k_2}\partial_{\varphi_j}^{\beta_2}g_x\big)_{L_x^2} + \partial_\lambda^{k_1}\partial_{\varphi_j}^{\beta_1}\chi_x\big(h, \partial_\lambda^{k_2}\partial_{\varphi_j}^{\beta_2}g\big)_{L_x^2}\big)$$

and the operators $\partial_{\varphi_j}^{\beta}\partial_\lambda^k\mathbf{R}$, $\partial_{\varphi_j}^{\beta}[\partial_i\mathbf{R}[\widehat\imath], \partial_x]$, $\partial_{\varphi_j}^{\beta}\partial_i\mathbf{R}[\widehat\imath]$ have similar expressions. □

In the next section we diagonalize the operator $\mathcal{L}_M^{(3)}$. We neglect the term $\mathbf{R}_M^{(3),\perp}$ in (6.243), which will contribute to the remainders in (7.94)-(7.95).

7 Almost diagonalization and invertibiility of $\mathcal{L}_\omega$

We have a linear real operator acting on $H_{\mathbb{S}_0}^\perp$,

$$\mathbf{L}_0 := \mathbf{L}_0(i) := \omega\cdot\partial_\varphi \mathbb{I}_2^\perp + \mathrm{i}\mathbf{D}_0 + \mathbf{R}_0 + \mathbf{Q}_0 \,, \qquad \mathbb{I}_2^\perp := \mathbb{I}_2\Pi_{\mathbb{S}_0}^\perp \,, \tag{7.1}$$

defined for all $(\omega,\kappa) \in \mathtt{DC}^{\gamma}_{K_n} \times [\kappa_1,\kappa_2]$(see (1.40)), with diagonal part (with respect to the exponential basis)

$$\mathbf{D}_0 := \begin{pmatrix} \mathcal{D}_0 & 0 \\ 0 & -\mathcal{D}_0 \end{pmatrix}, \qquad \mathcal{D}_0 := \mathrm{diag}_{j\in\mathbb{S}_0^c}\mu_j^{(0)}, \qquad \mu_j^{(0)} := \mathtt{m}_3|j|^{\frac12}(1+\kappa|j|^2)^{\frac12} + \mathtt{m}_1|j|^{\frac12}, \tag{7.2}$$

where $\mathbb{S}_0^c := \mathbb{Z}\setminus\mathbb{S}_0$ (see (1.42)), $\mathtt{m}_3 := \mathtt{m}_3(\omega,\kappa) \in \mathbb{R}$, $\mathtt{m}_1 := \mathtt{m}_1(\omega,\kappa) \in \mathbb{R}$ are defined for all $(\omega,\kappa) \in \mathbb{R}^\nu \times [\kappa_1,\kappa_2]$, and

$$\mathbf{R}_0, \mathbf{Q}_0 : H^{\perp}_{\mathbb{S}_0} \to H^{\perp}_{\mathbb{S}_0}, \qquad \mathbf{R}_0 := \begin{pmatrix} \mathcal{R}_0 & 0 \\ 0 & \overline{\mathcal{R}_0} \end{pmatrix}, \quad \mathbf{Q}_0 := \begin{pmatrix} 0 & \mathcal{Q}_0 \\ \overline{\mathcal{Q}_0} & 0 \end{pmatrix} \tag{7.3}$$

are real, even and reversible. The operators $\mathbf{R}_0$, $\mathbf{Q}_0$ satisfy also the following tame estimates:

- **(Smallness assumption on $\mathbf{R}_0$ and $\mathbf{Q}_0$).** *The operators*

$$\mathcal{R}_0\,,\ [\mathcal{R}_0,\partial_x]\,,\ \partial_{\varphi_m}^{s_0}\mathcal{R}_0\,,\ \partial_{\varphi_m}^{s_0}[\mathcal{R}_0,\partial_x]\,,\ \mathcal{Q}_0,\ [\mathcal{Q}_0,\partial_x]\,,\ \partial_{\varphi_m}^{s_0}\mathcal{Q}_0\,,\ \partial_{\varphi_m}^{s_0}[\mathcal{Q}_0,\partial_x]\,,\ \forall m=1,\ldots,\nu\,,$$

are $\mathcal{D}^{k_0}$-tame with tame constants, defined for all $s_0 \le s \le S$,

$$\mathbb{M}_0(s) := \max_{m=1,\ldots,\nu,\mathcal{R}\in\{\mathcal{R}_0,\mathcal{Q}_0\}} \big\{\mathfrak{M}_{\mathcal{R}}(s), \mathfrak{M}_{[\mathcal{R},\partial_x]}(s), \mathfrak{M}_{\partial_{\varphi_m}^{s_0}\mathcal{R}}(s), \mathfrak{M}_{\partial_{\varphi_m}^{s_0}[\mathcal{R},\partial_x]}(s)\big\}\,. \tag{7.4}$$

In addition the operators

$$\partial_{\varphi_m}^{s_0+\mathtt{b}}\mathcal{R}_0\,,\ \partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathcal{R}_0,\partial_x]\,,\ \partial_{\varphi_m}^{s_0+\mathtt{b}}\mathcal{Q}_0\,,\ \partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathcal{Q}_0,\partial_x]\,,\ m=1,\ldots,\nu\,,$$

are $\mathcal{D}^{k_0}$-tame with tame constants, defined for all $s_0 \le s \le S$,

$$\mathbb{M}_0(s,\mathtt{b}) := \max_{m=1,\ldots,\nu,\mathcal{R}\in\{\mathcal{R}_0,\mathcal{Q}_0\}} \big\{\mathfrak{M}_{\partial_{\varphi_m}^{s_0+\mathtt{b}}\mathcal{R}}(s), \mathfrak{M}_{\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathcal{R},\partial_x]}(s)\big\} \tag{7.5}$$

where $\mathtt{b} \in \mathbb{N}$ satisfies

$$\mathtt{b} := [\mathtt{a}] + 2 \in \mathbb{N}\,, \quad \mathtt{a} := 3\tau_1\,, \quad \chi = 3/2\,, \quad \tau_1 := \tau + (\tau+1)k_0\,. \tag{7.6}$$

We assume that the tame constants satisfy

$$\mathfrak{M}_0(s_0,\mathtt{b}) := \max\{\mathbb{M}_0(s_0), \mathbb{M}_0(s_0,\mathtt{b})\} \le C(S)\varepsilon\gamma^{-1} \tag{7.7}$$

and moreover, there is $\sigma(\mathtt{b}) > 0$ (we take $\sigma(\mathtt{b}) = \mu(\mathtt{b}) + \sigma$ in Lemma 7.2), such that, for all $m = 1,\ldots,\nu$, $\beta \in \mathbb{N}$, $\beta \le \mathtt{b} + s_0$,

$$\max_{\mathcal{R}\in\{\mathcal{R}_0,\mathcal{Q}_0\}} \big\{\|\partial_{\varphi_m}^{\beta}\partial_i\mathcal{R}[\hat\imath]\|_{\mathcal{L}(H^{s_0})}, \|\partial_{\varphi_m}^{\beta}[\partial_i\mathcal{R}[\hat\imath],\partial_x]\|_{\mathcal{L}(H^{s_0})}\big\} \le C(S)\varepsilon\gamma^{-1}\|\hat\imath\|_{s_0+\sigma(\mathtt{b})}\,. \tag{7.8}$$

Remark 7.1. The conditions $\mathtt{b} > \mathtt{a} + \chi^{-1}$ and $\mathtt{a} > 3\tau_1 = \tau_1\chi/(2-\chi)$ arise for the convergence of the iterative scheme (7.74)-(7.75), see Lemma 7.10. We take an integer $\mathtt{b} := [\mathtt{a}] + 2 \in \mathbb{N}$ so that $\partial_{\varphi_m}^{s_0+\mathtt{b}}$ are differential operators (recall also that $s_0 \in \mathbb{N}$ by (1.20)). Note also that $\mathtt{a} > \chi k_0(\tau+2) + 1$ (as $\tau \ge 1$) which is used in the extension procedure in $(\mathbf{S2})_\nu$, see e.g. (7.27). Moreover $\mathtt{a} > \chi(\tau + k_0(\tau+2))$ which is used in Lemma 8.7. □

Proposition 6.31 implies that the operators $\mathbf{R}_M^{(3)}$, $\mathbf{Q}_M^{(3)}$ in (6.247) satisfy the above tame estimates by fixing the constant M in section 6.5 large enough (this means to perform sufficiently many regularizing steps in Proposition 6.11), namely

$$M := \mathtt{b} + s_0 + k_0 + 4\,. \tag{7.9}$$

Set (recall (6.197), (6.96))

$$\mathbf{c}(\mathtt{b}) := \chi(\mathtt{b} + s_0 + k_0 + 4) + \daleth(\mathtt{b} + s_0 + k_0 + 6) + \aleph_{\mathtt{b}+s_0+k_0+4}(0)\,, \quad \mu(\mathtt{b}) := s_0 + \mathbf{c}(\mathtt{b}) + \mathtt{b}\,. \tag{7.10}$$

Lemma 7.2. (Tame estimates of $\mathbf{R}_M^{(3)}$, $\mathbf{Q}_M^{(3)}$) *Assume* (6.10) *with $\mu \geq \mu(\mathtt{b}) + \sigma$. Then the operators $\mathbf{R}_0 := \mathbf{R}_M^{(3)}$, $\mathbf{Q}_0 := \mathbf{Q}_M^{(3)}$ in* (6.247) *satisfy, for all $s_0 \leq s \leq S$, the tame estimates* (7.4)-(7.5) *with*

$$\mathbb{M}_0(s) \leq_S \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+s_0+\sigma+\mathtt{c}(\mathtt{b})}\big)\,, \qquad \mathbb{M}_0(s,\mathtt{b}) \leq_S \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu(\mathtt{b})+\sigma}\big) \tag{7.11}$$

and (7.7) *holds. Moreover, for all $m = 1, \ldots, \nu$, $\beta \in \mathbb{N}$, $\beta \leq \mathtt{b} + s_0$, the operators $\partial^\beta_{\varphi_m}\partial_i \mathcal{R}[\hat\imath]$, $\partial^\beta_{\varphi_m}[\partial_i\mathcal{R}[\hat\imath], \partial_x]$, $\mathcal{R} \in \{\mathcal{R}_0, \mathcal{Q}_0\}$ satisfy the bounds* (7.8) *with $\sigma(\mathtt{b}) = \mu(\mathtt{b}) + \sigma$.*

Proof. The estimates (7.11) follow by (6.248) and by the definitions (7.9), (7.10). Moreover with the choice of $\mu := \mu(\mathtt{b}) + \sigma$ in (7.10) (see also (7.9)) the condition (6.249) holds with $s_1 = s_0$ and so (7.8) holds by (6.250), with $\sigma(\mathtt{b}) = \mu(\mathtt{b}) + \sigma$. ☐

By (7.11), (7.10), we have verified that, for all $s_0 \leq s \leq S$,

$$\mathfrak{M}_0(s,\mathtt{b}) := \max\{\mathbb{M}_0(s), \mathbb{M}_0(s,\mathtt{b})\} \leq_S \varepsilon\gamma^{-1}(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu(\mathtt{b})+\sigma})\,. \tag{7.12}$$

We perform the almost reducibility of $\mathbf{L}_0$ along the scale

$$N_{-1} := 1\,, \quad N_\nu := N_0^{\chi^\nu}\,, \quad \forall \nu \geq 0\,, \quad \chi := 3/2\,, \tag{7.13}$$

requiring inductively at each step the second order Melnikov non-resonance conditions in (7.19).

Theorem 7.3. (Almost reducibility) *There exists $\tau_0 := \tau_0(\tau, \nu) > 0$ such that, for all $S > s_0$, there is $N_0 := N_0(S,\mathtt{b}) \in \mathbb{N}$ such that, if*

$$N_0^{\tau_0}\mathfrak{M}_0(s_0,\mathtt{b})\gamma^{-1} \leq 1\,, \tag{7.14}$$

(see (7.7)*), then, for all $n \in \mathbb{N}$, $\nu = 0, 1, \ldots, n$:*

(S1)$_\nu$ *There exists a real, even and reversible operator*

$$\mathbf{L}_\nu := \omega\cdot\partial_\varphi \mathbb{I}_2^\perp + \mathrm{i}\mathbf{D}_\nu + \mathbf{R}_\nu + \mathbf{Q}_\nu\,, \quad \mathbf{D}_\nu := \begin{pmatrix} \mathcal{D}_\nu & 0 \\ 0 & -\mathcal{D}_\nu \end{pmatrix}\,, \quad \mathcal{D}_\nu := \mathrm{diag}_{j\in\mathbb{S}_0^c}\mu_j^\nu\,, \tag{7.15}$$

-which acts on the space of functions even in x- defined for $(\omega,\kappa) \in \mathtt{DC}^\gamma_{K_n} \times [\kappa_1,\kappa_2]$ for $\nu = 0$, and for all (ω,κ) in

$$\mathcal{N}(\Lambda^\gamma_\nu, \gamma N_{\nu-1}^{-\tau-2}) \subset \Lambda^{\gamma/2}_\nu\,, \quad \text{for } \nu \geq 1\,, \tag{7.16}$$

(recall the definition (1.41)*) where μ_j^ν are k_0-times differentiable functions of the form*

$$\mu_j^\nu(\omega,\kappa) := \mu_j^0(\omega,\kappa) + r_j^\nu(\omega,\kappa)\,, \quad \mu_j^0 := \mathtt{m}_3|j|^{\frac12}(1+\kappa j^2)^{\frac12} + \mathtt{m}_1|j|^{\frac12}\,, \tag{7.17}$$

satisfying

$$\mu_j^\nu = \mu_{-j}^\nu\,, \quad \text{i.e. } r_j^\nu = r_{-j}^\nu\,, \quad |r_j^\nu|^{k_0,\gamma} \leq C(S)\varepsilon\gamma^{-1}\,, \quad \forall j \in \mathbb{S}_0^c\,. \tag{7.18}$$

The sets Λ^γ_ν are defined by $\Lambda^\gamma_0 := \Omega \times [\kappa_1,\kappa_2]$, and, for all $\nu \geq 1$,

$$\begin{aligned}\Lambda^\gamma_\nu := \Lambda^\gamma_\nu(i) := \Big\{ &\lambda = (\omega,\kappa) \in \Lambda^\gamma_{\nu-1} \cap \big([\mathtt{DC}^\gamma_{K_n} \cap \mathtt{DC}^\gamma_{N_{\nu-1}}] \times [\kappa_1,\kappa_2]\big) \ : \\ &|\omega\cdot\ell + \mu_j^{\nu-1} - \varsigma\mu_{j'}^{\nu-1}| \geq \gamma|j^{\frac32} - \varsigma j'^{\frac32}|\langle\ell\rangle^{-\tau}, \forall |\ell| \leq N_{\nu-1}, j, j' \in \mathbb{N}\setminus\mathbb{S}^+, \varsigma \in \{+,-\}\Big\}\end{aligned} \tag{7.19}$$

(recall (1.40) *and that the tangential sites $\mathbb{S} = \mathbb{S}^+ \cup (-\mathbb{S}^+) \subset \mathbb{Z}$ with $\mathbb{S}^+ \subset \mathbb{N}$). The remainders*

$$\mathbf{R}_\nu := \begin{pmatrix} \mathcal{R}_\nu & 0 \\ 0 & \overline{\mathcal{R}_\nu} \end{pmatrix}\,, \qquad \mathbf{Q}_\nu := \begin{pmatrix} 0 & \mathcal{Q}_\nu \\ \overline{\mathcal{Q}_\nu} & 0 \end{pmatrix} \tag{7.20}$$

are $\mathcal{D}^{k_0}$-modulo-tame: more precisely the operators $\mathcal{R}_\nu, \mathcal{Q}_\nu$, respectively $\langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{R}_\nu, \langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{Q}_\nu$, are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constants respectively

$$\mathfrak{M}^\sharp_\nu(s) := \max\{\mathfrak{M}^\sharp_{\mathcal{R}_\nu}(s), \mathfrak{M}^\sharp_{\mathcal{Q}_\nu}(s)\}\,, \quad \mathfrak{M}^\sharp_\nu(s,\mathtt{b}) := \max\{\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{R}_\nu}(s), \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{Q}_\nu}(s)\}, \tag{7.21}$$

satisfying for all $s \in [s_0, S]$,

$$\mathfrak{M}^\sharp_\nu(s) \le \mathfrak{M}_0(s,\mathtt{b}) N_{\nu-1}^{-\mathtt{a}}\,, \quad \mathfrak{M}^\sharp_\nu(s,\mathtt{b}) \le \mathfrak{M}_0(s,\mathtt{b}) N_{\nu-1}\,. \tag{7.22}$$

Moreover, for $\nu \ge 1$, *there exists a real, even and reversibility preserving map*

$$\mathbf{\Phi}_{\nu-1} := \mathbb{I}_2^\perp + \mathbf{\Psi}_{\nu-1}\,, \quad \mathbf{\Psi}_{\nu-1} := \begin{pmatrix} \Psi_{\nu-1,1} & \Psi_{\nu-1,2} \\ \overline{\Psi}_{\nu-1,2} & \overline{\Psi}_{\nu-1,1} \end{pmatrix}\,, \tag{7.23}$$

such that

$$\mathbf{L}_\nu := \mathbf{\Phi}_{\nu-1}^{-1} \mathbf{L}_{\nu-1} \mathbf{\Phi}_{\nu-1}\,. \tag{7.24}$$

The operators $\Psi_{\nu-1,m}$ *and* $\langle\partial_\varphi\rangle^{\mathtt{b}}\Psi_{\nu-1,m}$, $m=1,2$, *are* $\mathcal{D}^{k_0}$*-modulo-tame with modulo-tame constants satisfying, for all* $s \in [s_0,S]$, *(*$\tau_1, \mathtt{a}$ *are defined in (7.6))*

$$\mathfrak{M}^\sharp_{\Psi_{\nu-1,m}}(s) \le \frac{C(k_0)}{\gamma} N_{\nu-1}^{\tau_1} N_{\nu-2}^{-\mathtt{a}} \mathfrak{M}_0(s,\mathtt{b})\,,\ \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\Psi_{\nu-1,m}}(s) \le \frac{C(k_0)}{\gamma} N_{\nu-1}^{\tau_1} N_{\nu-2} \mathfrak{M}_0(s,\mathtt{b})\,. \tag{7.25}$$

$(\mathbf{S2})_\nu$ *For all* $j \in \mathbb{S}_0^c$ *there exists a* k_0*-times differentiable extension* $\tilde\mu_j^\nu : \Omega\times[\kappa_1,\kappa_2] \mapsto \mathbb{R}$ *such that* $\tilde\mu_j^\nu = \mu_j^\nu$ *on* Λ_ν^γ, *and*

$$\tilde\mu_j^\nu(\omega,\kappa) := \mu_j^0(\omega,\kappa) + \tilde r_j^\nu(\omega,\kappa) \in \mathbb{R}\,,\ \tilde r_j^\nu = \tilde r_{-j}^\nu\,,\ |\tilde r_j^\nu|^{k_0,\gamma} \le C(S)\varepsilon\gamma^{-1} N_0^{k_0(\tau+2)}\,,\ \forall j \in \mathbb{S}_0^c\,, \tag{7.26}$$

and for all $\nu \ge 1$

$$|\tilde\mu_j^\nu - \tilde\mu_j^{\nu-1}|^{k_0,\gamma} \le C(k_0) N_{\nu-1}^{k_0(\tau+2)} \mathfrak{M}^\sharp_{\nu-1}(s_0) \le C(k_0,S)\varepsilon\gamma^{-1} N_{\nu-1}^{k_0(\tau+2)} N_{\nu-2}^{-\mathtt{a}}\,. \tag{7.27}$$

$(\mathbf{S3})_\nu$ *Let* $i_1(\omega,\kappa)$, $i_2(\omega,\kappa)$ *such that* $\mathbf{R}_0(i_1)$, $\mathbf{Q}_0(i_1)$, $\mathbf{R}_0(i_2)$, $\mathbf{Q}_0(i_2)$ *satisfy (7.7). Assume also (7.8). Then for all* $\nu = 0,\dots n$, *for all* $(\omega,\kappa) \in \Lambda_\nu^{\gamma_1}(i_1) \cap \Lambda_\nu^{\gamma_2}(i_2)$ *with* $\gamma_1,\gamma_2 \in [\gamma/2, 2\gamma]$, *there exists* $\sigma := \sigma(\tau,\nu,k_0) > 0$ *such that*

$$|||\mathcal{R}_\nu(i_1) - \mathcal{R}_\nu(i_2)|||_{\mathcal{L}(H^{s_0})}, |||\mathcal{Q}_\nu(i_1) - \mathcal{Q}_\nu(i_2)|||_{\mathcal{L}(H^{s_0})} \le_{S,\mathtt{b}} \varepsilon\gamma^{-1} N_{\nu-1}^{-\mathtt{a}} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma}, \tag{7.28}$$

$$|||\langle\partial_\varphi\rangle^{\mathtt{b}}(\mathcal{R}_\nu(i_1) - \mathcal{R}_\nu(i_2))|||_{\mathcal{L}(H^{s_0})}, |||\langle\partial_\varphi\rangle^{\mathtt{b}}(\mathcal{Q}_\nu(i_1) - \mathcal{Q}_\nu(i_2))|||_{\mathcal{L}(H^{s_0})} \le_{S,\mathtt{b}} \frac{\varepsilon}{\gamma} N_{\nu-1} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma}\,. \tag{7.29}$$

Moreover for all $\nu = 1,\dots,n$, *for all* $j \in \mathbb{S}_0^c$,

$$\big|(r_j^\nu(i_1) - r_j^\nu(i_2)) - (r_j^{\nu-1}(i_1) - r_j^{\nu-1}(i_2))\big| \le C |||\mathcal{R}_\nu(i_1) - \mathcal{R}_\nu(i_2)|||_{\mathcal{L}(H^{s_0})}\,, \tag{7.30}$$

$$|r_j^\nu(i_1) - r_j^\nu(i_2)| \le C(S)\varepsilon\gamma^{-1} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma}\,. \tag{7.31}$$

$(\mathbf{S4})_\nu$ *Let* i_1, i_2 *be like in* $(\mathbf{S3})_\nu$ *and* $0 < \rho < \gamma/2$. *Then*

$$\varepsilon\gamma^{-1} C(S) N_{\nu-1}^{\tau} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma} \le \rho \quad \Longrightarrow \quad \Lambda_\nu^\gamma(i_1) \subseteq \Lambda_\nu^{\gamma-\rho}(i_2)\,.$$

Remark 7.4. Note that (7.30)-(7.31) are sufficient to prove $(\mathbf{S4})_\nu$ about the inclusion of the Cantor sets $\Lambda_\nu^\gamma(i_1)$, $\Lambda_\nu^{\gamma-\rho}(i_2)$ corresponding to two nearby approximate solutions: a smallness condition in $|\ |^{k_0,\gamma}$ is not required. This is sufficient to prove Lemma 8.6, and thus Lemma 8.7. The bounds (7.30)-(7.31) are implied just by the estimate (7.28), which is in s_0 norm and there is no control of the derivatives with respect to (ω,κ). This is why we do not need to estimate the derivatives with respect to (ω,κ) of the operators $\partial_i\mathcal{R}$ in (7.8). □

An important point of Theorem 7.3 is to require only the bound (7.14) for $\mathfrak{M}_0(s_0,\mathtt{b})$ in low norm, which is verified in Lemma 7.2, as well as the estimate (7.8) (which is still in low norm). On the other hand Theorem 7.3 provides the smallness (7.22) of the tame constants $\mathfrak{M}^\sharp_\nu(s)$ and proves that $\mathfrak{M}^\sharp_\nu(s,\mathtt{b})$, $\nu \ge 0$, do not diverge too much. Theorem 7.3 implies that the invertible operator

$$\mathbf{U}_n := \mathbf{\Phi}_0 \circ \ldots \circ \mathbf{\Phi}_n \tag{7.32}$$

has almost diagonalized $\mathbf{L}_0$, i.e. (7.35) below holds. We have the following corollary:

Theorem 7.5. (KAM almost-reducibility) *Assume* (6.10) *with $\mu \geq \mu(\mathtt{b}) + \sigma$. For all $S > s_0$ there exists $N_0 := N_0(S, \mathtt{b}) > 0$, $\delta_0 := \delta_0(S) > 0$ such that, if the smallness condition*

$$N_0^{\tau_0} \varepsilon \gamma^{-2} \leq \delta_0 \tag{7.33}$$

holds, where the constant $\tau_0 := \tau_0(\tau, \nu)$ is defined in Theorem 7.3, then, for all $n \in \mathbb{N}$, for all $\lambda = (\omega, \kappa)$ in

$$\Lambda_{n+1}^{\gamma} := \Lambda_{n+1}^{\gamma}(i) = \bigcap_{\nu=0}^{n+1} \Lambda_{\nu}^{\gamma} \tag{7.34}$$

where the sets Λ_{ν}^{γ} are defined in (7.19)*, the operator $\mathbf{U}_n$ in* (7.32) *is well defined and*

$$\mathbf{L}_n := \mathbf{U}_n^{-1} \mathbf{L}_0 \mathbf{U}_n = \omega \cdot \partial_{\varphi} \mathbb{I}_2^{\perp} + \mathrm{i} \mathbf{D}_n + \mathbf{R}_n + \mathbf{Q}_n \tag{7.35}$$

where $\mathbf{D}_n$ is defined in (7.15) *and $\mathbf{R}_n$, $\mathbf{Q}_n$ in* (7.20) *(with $\nu = n$). The operators $\mathcal{R}_n$, $\mathcal{Q}_n$ are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constants*

$$\mathfrak{M}^{\sharp}_{\mathcal{R}_n}(s), \mathfrak{M}^{\sharp}_{\mathcal{Q}_n}(s) \leq_S \varepsilon \gamma^{-1} N_{n-1}^{-\mathtt{a}} (1 + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}), \qquad \forall s_0 \leq s \leq S. \tag{7.36}$$

Moreover the operators $\mathbf{U}_n^{\pm 1} - \mathbb{I}_2^{\perp}$ are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constants

$$\mathfrak{M}^{\sharp}_{\mathbf{U}_n^{\pm 1} - \mathbb{I}_2^{\perp}}(s) \leq_S \varepsilon \gamma^{-2} N_0^{\tau_1} (1 + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}), \quad \forall s_0 \leq s \leq S, \tag{7.37}$$

where τ_1 is defined in (7.6)*. The operators $\mathbf{U}_n$, $\mathbf{U}_n^{-1}$ are real, even and reversibility preserving. $\mathbf{L}_n$ is real, even and reversible.*

Proof. The assumption (7.14) of Theorem 7.3 holds by (7.12), (6.10) with $\mu \geq \mu(\mathtt{b}) + \sigma$, and (7.33). The estimate (7.36) follows by (7.22) (for $\nu = n$) and (7.12). It remains to prove (7.37). By Lemma 2.16 the composition of $\mathcal{D}^{k_0}$-modulo-tame operators is $\mathcal{D}^{k_0}$-modulo-tame. To estimate the modulo-tame constant $\mathfrak{M}^{\sharp}_{\mathbf{U}_{\nu+1}}(s)$ of $\mathbf{U}_{\nu+1} = \mathbf{U}_{\nu} \circ \mathbf{\Phi}_{\nu+1} = \mathbf{U}_{\nu} \circ (\mathbb{I}_2^{\perp} + \mathbf{\Psi}_{\nu+1})$, we use the following inductive inequalities, which are deduced by Lemma 2.16 and (7.25),

$$\mathfrak{M}^{\sharp}_{\mathbf{U}_{\nu+1}}(s_0) \leq \mathfrak{M}^{\sharp}_{\mathbf{U}_{\nu}}(s_0) \big(1 + C(k_0) \varepsilon_{\nu}(s_0)\big), \tag{7.38}$$

$$\mathfrak{M}^{\sharp}_{\mathbf{U}_{\nu+1}}(s) \leq \mathfrak{M}^{\sharp}_{\mathbf{U}_{\nu}}(s) (1 + C(k_0) \varepsilon_{\nu}(s_0)) + C(k_0) \mathfrak{M}^{\sharp}_{\mathbf{U}_{\nu}}(s_0) \varepsilon_{\nu}(s) \tag{7.39}$$

where $\varepsilon_{\nu}(s) := \mathfrak{M}_0(s, \mathtt{b}) \gamma^{-1} N_{\nu+1}^{\tau_1} N_{\nu}^{-\mathtt{a}}$.

Iterating (7.38), setting $\varepsilon_{\nu} := C(k_0) \varepsilon_{\nu}(s_0)$, and using (7.7), (7.25), (7.33) we get

$$\mathfrak{M}^{\sharp}_{\mathbf{U}_{\nu+1}}(s_0) \leq \mathfrak{M}^{\sharp}_{\mathbf{U}_0}(s_0) \prod_{\nu \geq 0} (1 + \varepsilon_{\nu}) \leq \mathfrak{M}^{\sharp}_{\mathbf{U}_0}(s_0) \exp(C(S) \varepsilon \gamma^{-2}) \leq 2, \quad \forall \nu \geq 0. \tag{7.40}$$

Iterating (7.39), using (7.40) and $\prod_{\nu \geq 0}(1 + \varepsilon_{\nu}) \leq 2$, we get

$$\mathfrak{M}^{\sharp}_{\mathbf{U}_{\nu+1}}(s) \leq_{k_0} \sum\nolimits_{\nu \geq 0} \varepsilon_{\nu}(s) + \mathfrak{M}^{\sharp}_{\mathbf{U}_0}(s) \leq C(k_0) \big(1 + N_0^{\tau_1} \mathfrak{M}_0(s, \mathtt{b}) \gamma^{-1}\big), \quad \forall \nu \geq 0, \tag{7.41}$$

since $\mathbf{U}_0 = \mathbf{\Phi}_0 = \mathbb{I}_2^{\perp} + \mathbf{\Psi}_0$ and $\mathfrak{M}^{\sharp}_{\mathbf{U}_0}(s) \leq 1 + C(k_0) N_0^{\tau_1} \mathfrak{M}_0(s, \mathtt{b}) \gamma^{-1}$ by (7.25). Finally

$$\mathbf{U}_n - \mathbb{I}_2^{\perp} = (\mathbf{U}_n - \mathbf{\Phi}_0) + (\mathbf{\Phi}_0 - \mathbb{I}_2^{\perp}) = \sum_{\nu=0}^{n-1} (\mathbf{U}_{\nu+1} - \mathbf{U}_{\nu}) + \mathbf{\Psi}_0 = \sum_{\nu=0}^{n-1} \mathbf{U}_{\nu} \mathbf{\Psi}_{\nu+1} + \mathbf{\Psi}_0.$$

Hence Lemma 2.16, (7.40), (7.41), (7.12), (6.10), (7.25), (7.33), imply (7.37) for $\mathbf{U}_n - \mathbb{I}_2^{\perp}$. The estimate for $\mathbf{U}_n^{-1} - \mathbb{I}_2^{\perp}$ follows by Lemma 2.17. □

7.1 Proof of Theorem 7.3

Proof of $(\mathbf{S1})_0$. Properties (7.15)-(7.20) for $\nu = 0$ follow by the assumptions (7.1)-(7.3) with $r_j^0(\omega,\kappa) = 0$. We now prove that also (7.22) for $\nu = 0$ holds:

Lemma 7.6. *$\mathfrak{M}_0^\sharp(s)$, $\mathfrak{M}_0^\sharp(s,\mathtt{b}) \leq_{s_0,\mathtt{b}} \mathfrak{M}_0(s,\mathtt{b})$.*

Proof. Let $\mathcal{R} \in \{\mathcal{R}_0, \mathcal{Q}_0\}$ and set $\lambda := (\omega,\kappa)$. The matrix elements of the commutator $[\mathcal{R},\partial_x]$ are $\mathrm{i}(j'-j)(\mathcal{R})_j^{j'}(\ell-\ell')$, of $\partial_{\varphi_m}^b \mathcal{R}$, $m = 1,\ldots,\nu$, are $\mathrm{i}^b(\ell_m - \ell'_m)^b \mathcal{R}_j^{j'}(\ell-\ell')$, and of $\partial_{\varphi_m}^b[\mathcal{R},\partial_x]$ are $\mathrm{i}^{b+1}(\ell_m - \ell'_m)^b(j'-j)(\mathcal{R}_0)_j^{j'}(\ell-\ell')$. Then, recalling (2.67) with $\sigma = 0$, the assumptions (7.4)-(7.5) imply that $\forall |k| \leq k_0$, $s_0 \leq s \leq S$, $\ell' \in \mathbb{Z}^\nu$, $j' \in \mathbb{S}_0^c$,

$$\gamma^{2|k|}\sum_{\ell,j}\langle \ell,j\rangle^{2s}|\partial_\lambda^k \mathcal{R}_j^{j'}(\ell-\ell')|^2 \leq 2\mathbb{M}_0^2(s_0)\langle \ell',j'\rangle^{2s} + 2\mathbb{M}_0^2(s)\langle \ell',j'\rangle^{2s_0} \tag{7.42}$$

$$\gamma^{2|k|}\sum_{\ell,j}\langle \ell,j\rangle^{2s}|j-j'|^2|\partial_\lambda^k \mathcal{R}_j^{j'}(\ell-\ell')|^2 \leq 2\mathbb{M}_0^2(s_0)\langle \ell',j'\rangle^{2s} + 2\mathbb{M}_0^2(s)\langle \ell',j'\rangle^{2s_0} \tag{7.43}$$

$$\gamma^{2|k|}\sum_{\ell,j}\langle \ell,j\rangle^{2s}|\ell_m-\ell'_m|^{2s_0}|\partial_\lambda^k \mathcal{R}_j^{j'}(\ell-\ell')|^2 \leq 2\mathbb{M}_0^2(s_0)\langle \ell',j'\rangle^{2s} + 2\mathbb{M}_0^2(s)\langle \ell',j'\rangle^{2s_0} \tag{7.44}$$

$$\gamma^{2|k|}\sum_{\ell,j}\langle \ell,j\rangle^{2s}|\ell_m-\ell'_m|^{2s_0}|j-j'|^2|\partial_\lambda^k \mathcal{R}_j^{j'}(\ell-\ell')|^2 \leq 2\mathbb{M}_0^2(s_0)\langle \ell',j'\rangle^{2s} + 2\mathbb{M}_0^2(s)\langle \ell',j'\rangle^{2s_0} \tag{7.45}$$

$$\gamma^{2|k|}\sum_{\ell,j}\langle \ell,j\rangle^{2s}|\ell_m-\ell'_m|^{2(s_0+\mathtt{b})}|\partial_\lambda^k \mathcal{R}_j^{j'}(\ell-\ell')|^2 \leq 2\mathbb{M}_0^2(s_0,\mathtt{b})\langle \ell',j'\rangle^{2s} + 2\mathbb{M}_0^2(s,\mathtt{b})\langle \ell',j'\rangle^{2s_0} \tag{7.46}$$

$$\gamma^{2|k|}\sum_{\ell,j}\langle \ell,j\rangle^{2s}|\ell_m-\ell'_m|^{2(s_0+\mathtt{b})}|j-j'|^2|\partial_\lambda^k \mathcal{R}_j^{j'}(\ell-\ell')|^2 \leq 2\mathbb{M}_0^2(s_0,\mathtt{b})\langle \ell',j'\rangle^{2s} + 2\mathbb{M}_0^2(s,\mathtt{b})\langle \ell',j'\rangle^{2s_0}. \tag{7.47}$$

Using the inequality

$$\langle \ell-\ell'\rangle^{2s_1}\langle j-j'\rangle^2 \leq_{s_1} 1 + |j-j'|^2 + \max_{m=1,\ldots,\nu}|\ell_m-\ell'_m|^{2s_1} + |j-j'|^2\max_{m=1,\ldots,\nu}|\ell_m-\ell'_m|^{2s_1} \tag{7.48}$$

for $s_1 = s_0$, $s = s_0+\mathtt{b}$, the estimates (7.42)-(7.47) imply, recalling also (7.7),

$$\gamma^{2|k|}\sum_{\ell,j}\langle \ell,j\rangle^{2s}\langle \ell-\ell'\rangle^{2s_0}\langle j-j'\rangle^2|\partial_\lambda^k \mathcal{R}_j^{j'}(\ell-\ell')|^2 \leq_{\mathtt{b}} \mathfrak{M}_0^2(s_0,\mathtt{b})\langle \ell',j'\rangle^{2s} + \mathfrak{M}_0^2(s,\mathtt{b})\langle \ell',j'\rangle^{2s_0} \tag{7.49}$$

$$\gamma^{2|k|}\sum_{\ell,j}\langle \ell,j\rangle^{2s}\langle \ell-\ell'\rangle^{2(s_0+\mathtt{b})}\langle j-j'\rangle^2|\partial_\lambda^k \mathcal{R}_j^{j'}(\ell-\ell')|^2 \leq_{\mathtt{b}} \mathfrak{M}_0^2(s_0,\mathtt{b})\langle \ell',j'\rangle^{2s} + \mathfrak{M}_0^2(s,\mathtt{b})\langle \ell',j'\rangle^{2s_0}. \tag{7.50}$$

We can now prove that $\langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{R}$ is $\mathcal{D}^{k_0}$-modulo-tame. $\forall |k| \leq k_0$, by Cauchy-Schwartz inequality, we get

$$\begin{aligned}
\||\langle\partial_\varphi\rangle^{\mathtt{b}}\partial_\lambda^k\mathcal{R}|h\|_s^2 &\leq \sum_{\ell,j}\langle \ell,j\rangle^{2s}\Big(\sum_{\ell',j'}|\langle \ell-\ell'\rangle^{\mathtt{b}}\partial_\lambda^k\mathcal{R}_j^{j'}(\ell-\ell')||h_{\ell',j'}|\Big)^2 \\
&= \sum_{\ell,j}\langle \ell,j\rangle^{2s}\Big(\sum_{\ell',j'}\langle \ell-\ell'\rangle^{s_0+\mathtt{b}}\langle j'-j\rangle|\partial_\lambda^k\mathcal{R}_j^{j'}(\ell-\ell')||h_{\ell',j'}|\frac{1}{\langle \ell-\ell'\rangle^{s_0}\langle j'-j\rangle}\Big)^2 \\
&\leq_{s_0} \sum_{\ell,j}\langle \ell,j\rangle^{2s}\sum_{\ell',j'}\langle \ell-\ell'\rangle^{2(s_0+\mathtt{b})}\langle j'-j\rangle^2|\partial_\lambda^k\mathcal{R}_j^{j'}(\ell-\ell')|^2|h_{\ell',j'}|^2 \\
&= \sum_{\ell',j'}|h_{\ell',j'}|^2\sum_{\ell,j}\langle \ell,j\rangle^{2s}\langle \ell-\ell'\rangle^{2(s_0+\mathtt{b})}\langle j'-j\rangle^2|\partial_\lambda^k\mathcal{R}_j^{j'}(\ell-\ell')|^2 \\
&\overset{(7.50)}{\leq_{s_0,\mathtt{b}}} \gamma^{-2|k|}\sum_{\ell',j'}|h_{\ell',j'}|^2\big(\mathfrak{M}_0^2(s_0,\mathtt{b})\langle \ell',j'\rangle^{2s} + \mathfrak{M}_0^2(s,\mathtt{b})\langle \ell',j'\rangle^{2s_0}\big) \\
&\leq_{s_0,\mathtt{b}} \gamma^{-2|k|}\big(\mathfrak{M}_0^2(s_0,\mathtt{b})\|h\|_s^2 + \mathfrak{M}_0^2(s,\mathtt{b})\|h\|_{s_0}^2\big)\,.
\end{aligned} \tag{7.51}$$

Therefore (recall (2.73)) the modulo-tame constant $\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{R}}(s) \leq_{s_0,\mathtt{b}} \mathfrak{M}_0(s,\mathtt{b})$. Since $\mathcal{R}$ is both $\{\mathcal{R}_0, \mathcal{Q}_0\}$ we have proved that (see (7.21))

$$\mathfrak{M}_0^\sharp(s,\mathtt{b}) := \max\{\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{R}_0}(s), \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{Q}_0}(s)\} \leq_{s_0,\mathtt{b}} \mathfrak{M}_0(s,\mathtt{b})\,.$$

The inequality $\mathfrak{M}_0^\sharp(s) \leq_{s_0} \mathfrak{M}_0(s,\mathtt{b})$ follows similarly by (7.49). $\square$

Proof of $(\mathbf{S2})_0$. It follows since the functions $\mathtt{m}_3(\omega,\kappa)$ and $\mathtt{m}_1(\omega,\kappa)$ are k_0-times differentiable on all $\Omega \times [\kappa_1,\kappa_2]$ (they depend on the torus $i_\delta(\omega,\kappa)$ which is k_0-times differentiable with respect to (ω,κ) on all $\Omega \times [\kappa_1,\kappa_2]$).
Proof of $(\mathbf{S3})_0$. We prove (7.29) at $\nu = 0$, namely that, for $\mathcal{R} \in \{\mathcal{R}_0, \mathcal{Q}_0\}$,

$$\| |\langle \partial_\varphi \rangle^{\mathtt{b}} \Delta_{12} \mathcal{R}| h \|_{s_0}^2 \leq C(S,\mathtt{b}) \varepsilon^2 \gamma^{-2} \| i_1 - i_2 \|_{s_0+\mu(\mathtt{b})+\sigma}^2 \|h\|_{s_0}^2 , \quad \forall h \in H^{s_0} , \tag{7.52}$$

where we denote $\Delta_{12}\mathcal{R} := \mathcal{R}(i_1) - \mathcal{R}(i_2)$. By (7.8) and the mean value theorem we get

$$\|\Delta_{12}\mathcal{R}\|_{\mathcal{L}(H^{s_0})}, \|[\Delta_{12}\mathcal{R}, \partial_x]\|_{\mathcal{L}(H^{s_0})}, \|\partial_{\varphi_m}^{s_0+\mathtt{b}} \Delta_{12}\mathcal{R}\|_{\mathcal{L}(H^{s_0})}, \|\partial_{\varphi_m}^{s_0+\mathtt{b}} [\Delta_{12}\mathcal{R}, \partial_x]\|_{\mathcal{L}(H^{s_0})} \leq_{S,\mathtt{b}} \varepsilon\gamma^{-1} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma}$$

for all $m = 1, \ldots, \nu$. We deduce as in (7.42)-(7.47) (with $k = 0$) and (7.48) that, for all $\ell' \in \mathbb{Z}^\nu, j' \in \mathbb{S}_0^c$,

$$\sum\nolimits_{\ell,j} \langle \ell, j \rangle^{2s_0} \langle j - j' \rangle^2 \langle \ell - \ell' \rangle^{2(s_0+\mathtt{b})} |(\Delta_{12}\mathcal{R})_j^{j'}(\ell - \ell')|^2 \leq C(S,\mathtt{b}) \varepsilon^2 \gamma^{-2} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma}^2 \langle \ell', j' \rangle^{2s_0}$$

which, arguing as in (7.51), proves (7.52). The proof of (7.28) at $\nu = 0$ is analogous.
Proof of $(\mathbf{S4})_0$. It is trivial because by definition $\Omega_0^\gamma(i_1) = \Omega = \Omega_0^{\gamma-\rho}(i_2)$.

7.1.1 The reducibility step

In this section we describe the generic inductive step, showing how to define $\mathbf{L}_{\nu+1}$ (and $\mathbf{\Phi}_\nu$, $\mathbf{\Psi}_\nu$ etc). To simplify notations we drop the index ν and we write $+$ instead of $\nu + 1$, so that we write $\mathbf{L} := \mathbf{L}_\nu$, $\mathbf{D} := \mathbf{D}_\nu$, $\mathbf{R} := \mathbf{R}_\nu$, $\mathcal{R} := \mathcal{R}_\nu$, $\mathbf{Q} := \mathbf{Q}_\nu$, $\mathcal{Q} := \mathcal{Q}_\nu$, $\mathcal{D} := \mathcal{D}_\nu$, $\mu_j = \mu_j^\nu$, etc ...

We conjugate $\mathbf{L}$ by a transformation of the form (see (7.23))

$$\mathbf{\Phi} := \mathbb{I}_2^\perp + \mathbf{\Psi} , \qquad \mathbf{\Psi} := \begin{pmatrix} \Psi_1 & \Psi_2 \\ \overline{\Psi_2} & \overline{\Psi_1} \end{pmatrix} . \tag{7.53}$$

We have

$$\mathbf{L\Phi} = \mathbf{\Phi}\big(\omega\cdot\partial_\varphi \mathbb{I}_2^\perp + \mathrm{i}\mathbf{D}\big) + \big(\omega\cdot\partial_\varphi \mathbf{\Psi} + \mathrm{i}[\mathbf{D}, \mathbf{\Psi}] + \Pi_N \mathbf{R} + \Pi_N \mathbf{Q}\big) + \Pi_N^\perp \mathbf{R} + \Pi_N^\perp \mathbf{Q} + \mathbf{R\Psi} + \mathbf{Q\Psi} \tag{7.54}$$

where the projector Π_N is defined in (2.13) and $\Pi_N^\perp := \mathbb{I}_2 - \Pi_N$. We want to solve the homological equation

$$\omega\cdot\partial_\varphi \mathbf{\Psi} + \mathrm{i}[\mathbf{D}, \mathbf{\Psi}] + \Pi_N \mathbf{R} + \Pi_N \mathbf{Q} = [\mathbf{R}] \tag{7.55}$$

where

$$[\mathbf{R}] := \begin{pmatrix} [\mathcal{R}] & 0 \\ 0 & \overline{[\mathcal{R}]} \end{pmatrix} , \quad [\mathcal{R}] := \mathrm{diag}_{j \in \mathbb{S}_0^c} (\mathcal{R})_j^j(0) . \tag{7.56}$$

By (7.15), (7.20), (7.53) the equation (7.55) is equivalent to the two scalar homological equations

$$\omega\cdot\partial_\varphi \Psi_1 + \mathrm{i}[\mathcal{D}, \Psi_1] + \Pi_N \mathcal{R} = [\mathcal{R}] , \qquad \omega\cdot\partial_\varphi \Psi_2 + \mathrm{i}\big(\mathcal{D}\Psi_2 + \Psi_2 \mathcal{D}\big) + \Pi_N \mathcal{Q} = 0 . \tag{7.57}$$

The solutions of (7.57) are

$$(\Psi_1)_j^{j'}(\ell) := \begin{cases} -\dfrac{(\mathcal{R})_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + \mu_j - \mu_{j'})} & \forall (\ell, j, j') \neq (0, \pm j, \pm j) , |\ell| \leq N , \\ 0 \quad \text{otherwise} \end{cases} \tag{7.58}$$

$$(\Psi_2)_j^{j'}(\ell) := -\frac{(\mathcal{Q})_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + \mu_j + \mu_{j'})} , \qquad \forall (\ell, j, j') \in \mathbb{Z}^\nu \times \mathbb{S}_0^c \times \mathbb{S}_0^c , |\ell| \leq N . \tag{7.59}$$

Note that, since $\mu_j = \mu_{-j}$, $\forall j \in \mathbb{S}_0^c$ (see (7.18)) the denominators in (7.58), (7.59) are different from zero for $(\omega,\kappa) \in \Lambda_{\nu+1}^\gamma$ (see (7.19) with $\nu \rightsquigarrow \nu + 1$) and the maps Ψ_1, Ψ_2 are well defined.

Lemma 7.7. (Homological equations) *For all* $(\omega,\kappa)\in\Lambda^{\gamma/2}_{\nu+1}$ *the solutions* Ψ_1, Ψ_2 *in* (7.58), (7.59) *of the homological equations* (7.57) *are* $\mathcal{D}^{k_0}$*-modulo-tame operators with modulo-tame constants satisfying*

$$\mathfrak{M}^\sharp_{\Psi_1}(s),\mathfrak{M}^\sharp_{\Psi_2}(s)\le_{k_0} N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp(s),\quad \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\Psi_1}(s),\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\Psi_2}(s)\le_{k_0} N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp(s,\mathtt{b}) \tag{7.60}$$

where $\tau_1:=\tau(k_0+1)+k_0$.
Given i_1, i_2 *denote* $\Delta_{12}\Psi_1:=\Psi_1(i_2)-\Psi_1(i_1)$. *If* $\gamma/2\le\gamma_1,\gamma_2\le 2\gamma$ *then, for all* $(\omega,\kappa)\in\Lambda^{\gamma_1}_{\nu+1}(i_1)\cap\Lambda^{\gamma_2}_{\nu+1}(i_2)$,

$$|||\Delta_{12}\Psi_1|||_{\mathcal{L}(H^{s_0})}\le CN^{2\tau}\gamma^{-1}\big(|||\mathcal{R}(i_2)|||_{\mathcal{L}(H^{s_0})}\|i_1-i_2\|_{2s_0+\sigma+\mu(\mathtt{b})}+|||\Delta_{12}\mathcal{R}|||_{\mathcal{L}(H^{s_0})}\big)\,, \tag{7.61}$$

$$|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\Psi_1|||_{\mathcal{L}(H^{s_0})}\le_{\mathtt{b}} N^{2\tau}\gamma^{-1}\big(|||\langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{R}(i_2)|||_{\mathcal{L}(H^{s_0})}\|i_1-i_2\|_{2s_0+\sigma+\mu(\mathtt{b})}+|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\mathcal{R}|||_{\mathcal{L}(H^{s_0})}\big) \tag{7.62}$$

and a similar estimate holds for Ψ_2, *replacing* $\mathcal{R}$ *by* $\mathcal{Q}$. *Moreover* $\mathbf{\Psi}$ *is real, even and reversibility preserving.*

Proof. We make the proof for $\Psi:=\Psi_1$, for Ψ_2 is analogous.
Proof of (7.60). Let $(\omega,\kappa)\in\Lambda^{\gamma/2}_{\nu+1}$. By (7.19) with $\nu\rightsquigarrow\nu+1$, and the definition of Ψ_1 in (7.58), we have, for all $(\ell,j,j')\in\mathbb{Z}^\nu\times\mathbb{S}_0^c\times\mathbb{S}_0^c$, with $|\ell|\le N$, $(\ell,j,j')\ne(0,\pm j,\pm j)$, $|\Psi_j^{j'}(\ell)|\le CN^\tau\gamma^{-1}|\mathcal{R}_j^{j'}(\ell)|$. Moreover, differentiating (7.58) with respect to $\lambda=(\omega,\kappa)$, we get

$$\partial_\lambda^k\Psi_j^{j'}(\ell)=\sum\nolimits_{k_1+k_2=k}C(k_1,k_2)\big[\partial_\lambda^{k_1}(\omega\cdot\ell+\mu_j-\mu_{j'})^{-1}\big]\partial_\lambda^{k_2}\mathcal{R}_j^{j'}(\ell)\,,$$

and since, by (7.17), (7.18), (7.19), (6.83), (6.228),

$$\sup_{|k_1|\le k_0}|\partial_\lambda^{k_1}(\omega\cdot\ell+\mu_j-\mu_{j'})^{-1}|\le C(k_0)\langle\ell\rangle^{\tau(k_0+1)+k_0}\gamma^{-1-|k_1|}\,,$$

we deduce that, for all $0<|k|\le k_0$,

$$|\partial_\lambda^k\Psi_j^{j'}(\ell)|\le C(k_0)\langle\ell\rangle^{\tau(k_0+1)+k_0}\gamma^{-1-|k|}\sum\nolimits_{|k_2|\le|k|}\gamma^{|k_2|}|\partial_\lambda^{k_2}\mathcal{R}_j^{j'}(\ell)|\,. \tag{7.63}$$

Therefore for all $0\le|k|\le k_0$ we get

$$\begin{aligned}
|||\langle\partial_\varphi\rangle^{\mathtt{b}}\partial_\lambda^k\Psi|h|\|_s^2&\le\sum_{\ell,j}\langle\ell,j\rangle^{2s}\Big(\sum_{|\ell'-\ell|\le N,j'}|\langle\ell-\ell'\rangle^{\mathtt{b}}\partial_\lambda^k\Psi_j^{j'}(\ell-\ell')||h_{\ell',j'}|\Big)^2\\
&\overset{(7.63)}{\le_{k_0}} N^{2\tau_1}\gamma^{-2(1+|k|)}\sum_{|k_2|\le|k|}\gamma^{2|k_2|}\sum_{\ell,j}\langle\ell,j\rangle^{2s}\Big(\sum_{\ell',j'}|\langle\ell-\ell'\rangle^{\mathtt{b}}\partial_\lambda^{k_2}\mathcal{R}_j^{j'}(\ell-\ell')||h_{\ell',j'}|\Big)^2\\
&=N^{2\tau_1}\gamma^{-2(1+|k|)}\sum\nolimits_{|k_2|\le|k|}\gamma^{2|k_2|}|||\langle\partial_\varphi\rangle^{\mathtt{b}}\partial_\lambda^{k_2}\mathcal{R}|[\![h]\!]\|_s^2\\
&\overset{(7.21),(2.73)}{\le_{k_0}} N^{2\tau_1}\gamma^{-2(1+|k|)}\big(\mathfrak{M}^\sharp(s,\mathtt{b})^2|||h|||_{s_0}^2+\mathfrak{M}^\sharp(s_0,\mathtt{b})^2|||h|||_s^2\big)\\
&\overset{(2.4)}{=}C(k_0)N^{2\tau_1}\gamma^{-2(1+|k|)}\big(\mathfrak{M}^\sharp(s,\mathtt{b})^2\|h\|_{s_0}^2+\mathfrak{M}^\sharp(s_0,\mathtt{b})^2\|h\|_s^2\big)
\end{aligned} \tag{7.64}$$

and, recalling Definition 2.9, the second inequality in (7.60) follows. The proof of the first inequality is analogous.
Proof of (7.61)-(7.62). By (7.58), for all $(\omega,\kappa)\in\Lambda^{\gamma_1}_{\nu+1}(i_1)\cap\Lambda^{\gamma_2}_{\nu+1}(i_2)$, one has

$$\Delta_{12}\Psi_j^{j'}(\ell)=\frac{\Delta_{12}\mathcal{R}_j^{j'}(\ell)}{\delta_{\ell jj'}(i_1)}-\mathcal{R}_j^{j'}(\ell)(i_2)\frac{\Delta_{12}\delta_{\ell jj'}}{\delta_{\ell jj'}(i_1)\delta_{\ell jj'}(i_2)}\,,\quad \delta_{\ell jj'}:=\mathrm{i}(\omega\cdot\ell+\mu_j-\mu_{j'})\,.$$

By (7.17), (6.83), (6.228), (7.31) we get

$$|\Delta_{12}\delta_{\ell jj'}|=|\Delta_{12}(\mu_j-\mu_{j'})|\le C\varepsilon\gamma^{-1}\big||j|^{\frac32}-|j'|^{\frac32}\big|\|i_1-i_2\|_{2s_0+\sigma+\mu(\mathtt{b})}\,,$$

whence $\gamma_1^{-1}, \gamma_2^{-1} \le \gamma^{-1}$, $\varepsilon\gamma^{-2}$ small enough, imply

$$|\Delta_{12}\Psi_j^{j'}(\ell)| \le C N^{2\tau}\gamma^{-1}\big(|\mathcal{R}_j^{j'}(\ell)(i_2)|\|i_1 - i_2\|_{2s_0+\sigma+\mu(\mathtt{b})} + |\Delta_{12}\mathcal{R}_j^{j'}(\ell)|\big)$$

and (7.61), (7.62) follow arguing as in (7.64).

Finally, since $\mathbf{R}, \mathbf{Q}$ are even and reversible, (7.58), (7.59) imply that $\boldsymbol{\Psi}$ is even and reversibility preserving. □

By (7.54), (7.55) we have

$$\mathbf{L}_+ = \boldsymbol{\Phi}^{-1}\mathbf{L}\boldsymbol{\Phi} = \omega\cdot\partial_\varphi \mathbb{I}_2^\perp + \mathrm{i}\mathbf{D}_+ + \mathbf{R}_+ + \mathbf{Q}_+$$

which proves (7.24) and (7.15) at the step $\nu+1$, with

$$\mathrm{i}\mathbf{D}_+ := \mathrm{i}\mathbf{D} + [\mathbf{R}]\,, \quad \mathbf{R}_+ + \mathbf{Q}_+ = \boldsymbol{\Phi}^{-1}\big(\Pi_N^\perp \mathbf{R} + \Pi_N^\perp \mathbf{Q} + \mathbf{R}\boldsymbol{\Psi} - \boldsymbol{\Psi}[\mathbf{R}] + \mathbf{Q}\boldsymbol{\Psi}\big)\,. \tag{7.65}$$

The new operator $\mathbf{L}_+$ has the same form of $\mathbf{L}$ with $\mathbf{R}_+ + \mathbf{Q}_+$ which is the sum of a quadratic function of $\boldsymbol{\Psi}$ and $(\mathbf{R}, \mathbf{Q})$ and a remainder supported on high frequencies. The new normal form $\mathbf{D}_+$ is diagonal:

Lemma 7.8. (New diagonal part). *The new normal form is*

$$\mathrm{i}\mathbf{D}_+ = \mathrm{i}\mathbf{D} + [\mathbf{R}] = \mathrm{i}\begin{pmatrix} \mathcal{D}_+ & 0 \\ 0 & -\mathcal{D}_+ \end{pmatrix}, \quad \mathcal{D}_+ := \mathrm{diag}_{j\in\mathbb{S}_0^c}\mu_j^+\,, \quad \mu_j^+ := \mu_j + \mathtt{r}_j \in \mathbb{R}\,, \tag{7.66}$$

with $\mathtt{r}_j = \mathtt{r}_{-j}$, $\mu_j^+ = \mu_{-j}^+$, $\forall j \in \mathbb{S}_0^c$, *and* $|\mu_j^+ - \mu_j|^{k_0,\gamma} \lessdot \mathfrak{M}^\sharp(s_0)$.

Moreover, given tori $i_1(\omega,\kappa)$, $i_2(\omega,\kappa)$ *then, for all* $(\omega,\kappa) \in \Lambda_\nu^{\gamma_1}(i_1) \cap \Lambda_\nu^{\gamma_2}(i_2)$, *the difference*

$$|\mathtt{r}_j(i_1) - \mathtt{r}_j(i_2)| \le C\||\Delta_{12}\mathcal{R}|\|_{\mathcal{L}(H^{s_0})}\,. \tag{7.67}$$

Proof. By (7.58)-(7.59) the operator $[\mathbf{R}]$ in (7.56) satisfies

$$[\mathcal{R}]u = \sum\nolimits_{j\in\mathbb{S}_0^c}\big(\mathcal{R}_j^{-j}(0)u_{-j} + \mathcal{R}_j^j(0)u_j\big)e^{\mathrm{i}jx} = \sum\nolimits_{j\in\mathbb{S}_0^c}\big(\mathcal{R}_j^{-j}(0) + \mathcal{R}_j^j(0)\big)u_j e^{\mathrm{i}jx}$$

since $[\mathcal{R}]$ acts on the space of functions even in x, i.e. $u_j = u_{-j}$. Thus (7.66) holds with $\mathcal{R}_j^{-j}(0)+\mathcal{R}_j^j(0) =: \mathrm{i}\mathtt{r}_j$. Since $\mathcal{R}$ is even, by (2.15) we deduce $\mathtt{r}_{-j} = \mathtt{r}_j$. In addition, since $\mathcal{R} = A + \mathrm{i}B$ is reversible we have $\mathcal{R}(-\varphi) = -\overline{\mathcal{R}}(\varphi)$, and so the maps $\varphi \mapsto A_j^{j'}(\varphi)$ are odd and so the average $A_j^j(0) := \int_{\mathbb{T}^\nu} A_j^j(\varphi)\,d\varphi = 0$ as well as $A_j^{-j}(0) = 0$. Hence $\mathcal{R}_j^j(0) + \mathcal{R}_j^{-j}(0) = \mathrm{i}(B_j^j(0) + B_j^{-j}(0)) \in \mathrm{i}\mathbb{R}$ and each $\mathtt{r}_j \in \mathbb{R}$.

Recalling the definition of $\mathfrak{M}^\sharp(s_0)$ in (7.21) (with $s = s_0$) and Defintion 2.9, we have, for $\lambda = (\omega,\kappa)$, for all $0 \le |k| \le k_0$, $\||\partial_\lambda^k \mathcal{R}|h\|_{s_0} \le 2\gamma^{-|k|}\mathfrak{M}^\sharp(s_0)\|h\|_{s_0}$, which implies that (see (2.67))

$$|\partial_\lambda^k \mathcal{R}_j^j(0)| + |\partial_\lambda^k \mathcal{R}_j^{-j}(0)| \le C\gamma^{-|k|}\mathfrak{M}^\sharp(s_0)\,.$$

Hence

$$|\mu_j^+ - \mu_j|^{k_0,\gamma} \le |\mathcal{R}_j^j(0)|^{k_0,\gamma} + |\mathcal{R}_j^{-j}(0)|^{k_0,\gamma} \le C\mathfrak{M}^\sharp(s_0)\,.$$

The estimate (7.67) follows analogously by $|\Delta_{12}(\mathcal{R}_j^j(0) + \mathcal{R}_j^{-j}(0))| \le C\||\Delta_{12}\mathcal{R}|\|_{\mathcal{L}(H^{s_0})}$. □

7.1.2 The iteration

Let $\nu \ge 0$ and suppose that the statements $(\mathbf{S1})_\nu$-$(\mathbf{S4})_\nu$ are true. We prove $(\mathbf{S1})_{\nu+1}$-$(\mathbf{S4})_{\nu+1}$.

Proof of $(\mathbf{S1})_{\nu+1}$. Since the eigenvalues μ_j^ν are defined on $\mathcal{N}(\Lambda_\nu^\gamma, \gamma N_{\nu-1}^{-\tau-2})$, the set $\Lambda_{\nu+1}^\gamma$ is well-defined. Moreover μ_j^ν are well defined also on the set $\mathcal{N}(\Lambda_{\nu+1}^\gamma, \gamma N_\nu^{-\tau-2}) \subseteq \mathcal{N}(\Lambda_\nu^\gamma, \gamma N_{\nu-1}^{-\tau-2})$ because $\Lambda_{\nu+1}^\gamma \subseteq \Lambda_\nu^\gamma$. Let us prove (7.16) at the step $\nu+1$, namely that $\mathcal{N}(\Lambda^{\gamma}_{\nu+1}, \gamma N_\nu^{-\tau-2}) \subset \Lambda^{\gamma/2}_{\nu+1}$. Indeed, let $\lambda_0 = (\omega_0, \kappa_0) \in \Lambda^{\gamma}_{\nu+1}$ and $\lambda = (\omega, \kappa)$ with $|\lambda - \lambda_0| \leq \gamma N_\nu^{-\tau-2}$. Then, for all $|\ell| \leq N_\nu$, $j \neq j'$ (consider the case $\varsigma = 1$),

$$
\begin{aligned}
|\omega \cdot \ell + \mu_j^\nu(\lambda) - \mu_{j'}^\nu(\lambda)| &\geq |\omega_0 \cdot \ell + \mu_j^\nu(\lambda_0) - \mu_{j'}^\nu(\lambda_0)| - |\omega - \omega_0||\ell| - |(\mu_j^\nu - \mu_{j'}^\nu)(\lambda) - (\mu_j^\nu - \mu_{j'}^\nu)(\lambda_0)| \\
&\overset{(6.84),(6.228),(7.18),\varepsilon\gamma^{-2}\leq 1}{\geq} |\omega_0 \cdot \ell + \mu_j^\nu(\omega_0) - \mu_{j'}^\nu(\omega_0)| - \big(|\ell| + C(S)|j^{\frac32} - j'^{\frac32}|\big)|\lambda - \lambda_0| \\
&\geq \gamma |j^{\frac32} - j'^{\frac32}| \langle \ell \rangle^{-\tau} - \gamma N_\nu^{-\tau-1} - C(S)\gamma |j^{\frac32} - j'^{\frac32}| N_\nu^{-\tau-2} \geq \frac{\gamma}{2} |j^{\frac32} - j'^{\frac32}| \langle \ell \rangle^{-\tau}
\end{aligned}
$$

for $N_0 > 4C(S)$ large enough. Thus $\lambda = (\omega, \kappa) \in \Lambda^{\gamma/2}_{\nu+1}$ (defined in (7.19) with $\nu \rightsquigarrow \nu+1$ and $\gamma \rightsquigarrow \gamma/2$).

By (7.16) at the step $\nu+1$ and Lemma 7.7, for all $(\omega, \kappa) \in \mathcal{N}(\Lambda^{\gamma}_{\nu+1}, \gamma N_\nu^{-\tau-2})$ the solutions $\Psi_{\nu,m}$, $m = 1, 2$, of the homological equations (7.57), defined in (7.58), (7.59), are well defined and, by (7.60), (7.22), satisfy for all $0 \leq |k| \leq k_0$, the estimates (7.25) at $\nu+1$. In particular (7.25) at $\nu+1$ with $k=0$, $s = s_0$ imply

$$
\mathfrak{M}^\sharp_{\Psi_{\nu,m}}(s_0) \leq C(k_0) N_\nu^{\tau_1} N_{\nu-1}^{-\mathtt{a}} \gamma^{-1} \mathfrak{M}_0(s_0, \mathtt{b}), \quad m = 1, 2. \tag{7.68}
$$

Therefore, by (7.6), (7.14), the smallness condition (2.82) of Lemma 2.17 is verified for $N_0 := N_0(S, \mathtt{b})$ large enough and the map $\mathbf{\Phi}_\nu = \mathbb{I}_2^\perp + \mathbf{\Psi}_\nu$ is invertible. Its inverse has the form

$$
\mathbf{\Phi}_\nu^{-1} = \mathbb{I}_2^\perp + \check{\mathbf{\Psi}}_\nu, \qquad \check{\mathbf{\Psi}}_\nu := \begin{pmatrix} \check{\Psi}_{\nu,1} & \check{\Psi}_{\nu,2} \\ \overline{\check{\Psi}}_{\nu,2} & \overline{\check{\Psi}}_{\nu,1} \end{pmatrix} \tag{7.69}
$$

and, by Lemma 2.17, the $\check{\Psi}_{\nu,m}$ $m = 1, 2$, are $\mathcal{D}^{k_0}$-modulo-tame with the same modulo-tame constants of $\Psi_{\nu,m}$ (see (7.25) for $\nu+1$), i.e.

$$
\mathfrak{M}^\sharp_{\check{\Psi}_{\nu,m}}(s) \leq_{k_0} \gamma^{-1} N_\nu^{\tau_1} N_{\nu-1}^{-\mathtt{a}} \mathfrak{M}_0(s, \mathtt{b}), \quad \mathfrak{M}^\sharp_{\langle \partial_\varphi \rangle^{\mathtt{b}} \check{\Psi}_{\nu,m}}(s) \leq_{k_0, \mathtt{b}} \gamma^{-1} N_\nu^{\tau_1} N_{\nu-1} \mathfrak{M}_0(s, \mathtt{b}). \tag{7.70}
$$

Since $\mathbf{\Psi}_\nu$ is even and reversibility preserving, also $\check{\mathbf{\Psi}}_\nu$ is even and reversibility preserving.

By Lemma 7.8 the operator $\mathbf{D}_{\nu+1}$ is diagonal and its eigenvalues $\mu_j^{\nu+1} : \mathcal{N}(\Lambda^{\gamma}_{\nu+1}, \gamma N_\nu^{-\tau-2}) \to \mathbb{R}$ satisfy (7.18) at $\nu+1$.

Now we estimate the remainder (see (7.65))

$$
\mathbf{R}_{\nu+1} + \mathbf{Q}_{\nu+1} := \mathbf{\Phi}_\nu^{-1} \mathbf{H}_\nu, \qquad \mathbf{H}_\nu := \Pi^\perp_{N_\nu} \mathbf{R}_\nu + \Pi^\perp_{N_\nu} \mathbf{Q}_\nu + \mathbf{R}_\nu \mathbf{\Psi}_\nu - \mathbf{\Psi}_\nu [\mathbf{R}_\nu] + \mathbf{Q}_\nu \mathbf{\Psi}_\nu.
$$

By (7.69), (7.20), (7.53) we get

$$
\mathbf{R}_{\nu+1} = \begin{pmatrix} \mathcal{R}_{\nu+1} & 0 \\ 0 & \overline{\mathcal{R}}_{\nu+1} \end{pmatrix}, \qquad \mathbf{Q}_{\nu+1} := \begin{pmatrix} 0 & \mathcal{Q}_{\nu+1} \\ \overline{\mathcal{Q}}_{\nu+1} & 0 \end{pmatrix} \tag{7.71}
$$

where

$$
\begin{aligned}
\mathcal{R}_{\nu+1} &:= (\mathrm{Id} + \check{\Psi}_{\nu,1})(\Pi^\perp_{N_\nu} \mathcal{R}_\nu + \mathcal{R}_\nu \Psi_{\nu,1} - \Psi_{\nu,1}[\mathcal{R}_\nu] + \mathcal{Q}_\nu \overline{\Psi}_{\nu,2}) \\
&\quad + \check{\Psi}_{\nu,2}(\Pi^\perp_{N_\nu} \mathcal{Q}_\nu + \mathcal{R}_\nu \Psi_{\nu,2} - \Psi_{\nu,2}[\overline{\mathcal{R}}_\nu] + \mathcal{Q}_\nu \overline{\Psi}_{\nu,1}), \qquad (7.72) \\
\mathcal{Q}_{\nu+1} &:= (\mathrm{Id} + \check{\Psi}_{\nu,1})(\Pi^\perp_N \mathcal{Q}_\nu + \mathcal{R}_\nu \Psi_{\nu,2} - \Psi_{\nu,2}[\overline{\mathcal{R}}_\nu] + \mathcal{Q}_\nu \overline{\Psi}_{\nu,1}) \\
&\quad + \Pi^\perp_N \overline{\mathcal{R}}_\nu + \overline{\mathcal{R}}_\nu \overline{\Psi}_{\nu,1} - \overline{\Psi}_{\nu,1}[\overline{\mathcal{R}}_\nu] + \overline{\mathcal{Q}}_\nu \Psi_{\nu,2}. \qquad (7.73)
\end{aligned}
$$

Lemma 7.9. (Nash-Moser iterative scheme) *The operators $\mathcal{R}_{\nu+1}$, $\mathcal{Q}_{\nu+1}$ are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constants satisfying*

$$
\mathfrak{M}^\sharp_{\nu+1}(s) \leq_{k_0} N_\nu^{-\mathtt{b}} \mathfrak{M}^\sharp_\nu(s, \mathtt{b}) + N_\nu^{\tau_1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(s) \mathfrak{M}^\sharp_\nu(s_0). \tag{7.74}
$$

The operators $\langle \partial_\varphi \rangle^{\mathtt{b}} \mathcal{R}_{\nu+1}, \langle \partial_\varphi \rangle^{\mathtt{b}} \mathcal{Q}_{\nu+1}$ are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constants satisfying

$$
\mathfrak{M}^\sharp_{\nu+1}(s, \mathtt{b}) \leq_{k_0, \mathtt{b}} \mathfrak{M}^\sharp_\nu(s, \mathtt{b}) + N_\nu^{\tau_1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(s, \mathtt{b}) \mathfrak{M}_\nu(s_0) + N_\nu^{\tau_1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(s_0, \mathtt{b}) \mathfrak{M}^\sharp_\nu(s). \tag{7.75}
$$

Proof. We estimate each term in (7.72)-(7.73). The proof of (7.74) follows by Lemmata 2.18, 2.16, (7.60), (7.70). The proof of (7.75) follows by Lemma 2.16 (7.60), (7.70), (7.22) and Lemma 2.18. □

The estimates (7.74), (7.75), and (7.6), allow to prove that also (7.22) holds at the step $\nu+1$.

Lemma 7.10. *$\mathfrak{M}^\sharp_{\nu+1}(s) \leq N_\nu^{-\mathtt{a}}\mathfrak{M}_0(s,\mathtt{b})$ and $\mathfrak{M}^\sharp_{\nu+1}(s,\mathtt{b}) \leq N_\nu \mathfrak{M}_0(s,\mathtt{b})$.*

Proof. By (7.74) and (7.22) we get

$$\mathfrak{M}^\sharp_{\nu+1}(s) \leq_{k_0} N_\nu^{-\mathtt{b}} N_{\nu-1}\mathfrak{M}_0(s,\mathtt{b}) + N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}_0(s,\mathtt{b})\mathfrak{M}_0(s_0,\mathtt{b})N_{\nu-1}^{-2\mathtt{a}} \leq N_\nu^{-\mathtt{a}}\mathfrak{M}_0(s,\mathtt{b})$$

by (7.6), (7.14) and taking $N_0 := N_0(S,\mathtt{b}) > 0$ large enough. Then by (7.75), (7.22) we get that

$$\mathfrak{M}^\sharp_{\nu+1}(s,\mathtt{b}) \leq_{k_0,\mathtt{b}} N_{\nu-1}\mathfrak{M}_0(s,\mathtt{b}) + N_\nu^{\tau_1}N_{\nu-1}^{1-\mathtt{a}}\gamma^{-1}\mathfrak{M}_0(s,\mathtt{b})\mathfrak{M}_0(s_0,\mathtt{b}) \leq N_\nu\mathfrak{M}_0(s,\mathtt{b})$$

by (7.6), (7.14) and taking $N_0 := N_0(S,\mathtt{b}) > 0$ large enough. □

The proof of $(\mathbf{S1})_{\nu+1}$ is concluded by noting that the operators $\mathbf{R}_{\nu+1}$, $\mathbf{Q}_{\nu+1}$ are even and reversible because $\mathbf{\Phi}_\nu$ is even and reversibility preserving (Lemma 7.7).

Proof of $(\mathbf{S2})_{\nu+1}$. We now construct the smooth extension $\tilde\mu_j^{\nu+1}$ on all the parameter space $\mathtt{\Omega}\times[\kappa_1,\kappa_2]$. By the inductive hyphothesis there exists a k_0-times differentiable function $\tilde\mu_j^\nu : \mathtt{\Omega}\times[\kappa_1,\kappa_2] \mapsto \mathbb{R}$ such that $\mu_j^\nu = \tilde\mu_j^\nu$ on $\mathtt{\Lambda}_\nu^\gamma$ and $\tilde\mu_j^\nu = 0$ outside $\mathcal{N}(\mathtt{\Lambda}_\nu^\gamma, \gamma N_{\nu-1}^{-\tau-2})$. Note that all the sets $\mathtt{\Lambda}_\nu^\gamma$ in (7.19) are defined by only *finitely* many non-resonance conditions, namely (for brevity we omit to write the sets $\mathtt{DC}^\gamma_{K_n}\cap\mathtt{DC}^\gamma_{N_{\nu-1}}$)

$$\mathtt{\Lambda}_\nu^\gamma = \bigcap_{|\ell|\leq N_{\nu-1},|j|,|j'|\leq CN_{\nu-1}^2} \Big\{(\omega,\kappa)\in\mathtt{\Lambda}_{\nu-1}^\gamma : |\omega\cdot\ell+\mu_j^{\nu-1}-\varsigma\mu_{j'}^{\nu-1}| \geq \frac{\gamma|j^{\frac32}-\varsigma j'^{\frac32}|}{\langle\ell\rangle^\tau},\ j,j'\in\mathbb{S}_0^c, \varsigma\in\{+,-\}\Big\}.$$

Actually, provided $j^{\frac12}+j'^{\frac12}\geq CN_{\nu-1}$, $j\neq j'$, for all $(\omega,\kappa)\in\mathtt{\Lambda}_{\nu-1}^\gamma$ the functions

$$|\omega\cdot\ell+\mu_j^{\nu-1}-\mu_{j'}^{\nu-1}| \geq |\mu_j^{\nu-1}-\mu_{j'}^{\nu-1}| - |\omega||\ell| \geq \frac12|j^{\frac32}-j'^{\frac32}| - C|\ell| \geq C(j^{\frac12}+j'^{\frac12}) - CN_{\nu-1} \geq \frac12.$$

Since $\mu_j^{\nu+1} = \mu_j^\nu + \mathtt{r}_j^\nu$ (defined on $\mathcal{N}(\mathtt{\Lambda}_{\nu+1}^\gamma, \gamma N_\nu^{-\tau-2})$) we need only to extend the function $\mathtt{r}_j^\nu$.

Let $\psi_\nu \in \mathcal{C}^\infty : \mathbb{R}^{\nu+1}\to\mathbb{R}$ be a cut-off function satisfying: $0\leq\psi_\nu\leq 1$,

$$\psi_\nu(\lambda) = 1,\ \forall\lambda\in\mathtt{\Lambda}_{\nu+1}^\gamma,\ \mathrm{supp}(\psi_\nu)\subseteq\mathcal{N}(\mathtt{\Lambda}_{\nu+1}^\gamma,\gamma N_\nu^{-\tau-2}), |\partial_\lambda^k\psi_\nu(\lambda)| \leq C(k)\big(N_\nu^{\tau+2}\gamma^{-1}\big)^{|k|},\ \forall k\in\mathbb{N}^\nu,$$

and thus $|\psi_\nu|^{k_0,\gamma} \leq C(k_0)N_\nu^{(\tau+2)k_0}$. Hence, defining $\tilde{\mathtt{r}}_j^\nu := \psi_\nu\mathtt{r}_j^\nu$ and $\tilde\mu_j^{\nu+1} := \tilde\mu_j^\nu + \tilde{\mathtt{r}}_j^\nu$, we get the estimate

$$|\tilde\mu_j^{\nu+1}-\tilde\mu_j^\nu|^{k_0,\gamma} \leq |\psi_\nu|^{k_0,\gamma}|\mathtt{r}_j^\nu|^{k_0,\gamma} \leq C(k_0)N_\nu^{(\tau+2)k_0}\mathfrak{M}^\sharp_\nu(s_0) \leq \varepsilon\gamma^{-1}C(k_0,S,\mathtt{b})N_\nu^{(\tau+2)k_0}N_{\nu-1}^{-\mathtt{a}}$$

by Lemma 7.8, (7.22) and (7.12). This is (7.27) at $\nu+1$. Summing we also get (7.26) at the step $\nu+1$.

Proof of $(\mathbf{S3})_{\nu+1}$. At the ν-th step we have already constructed the operators

$$\mathcal{R}_\nu(i_m), \mathcal{Q}_\nu(i_m), \Psi_{\nu-1,1}(i_m), \Psi_{\nu-1,2}(i_m), \quad m=1,2,$$

which are defined on $\mathtt{\Lambda}_\nu^{\gamma_1}(i_1)\cap\mathtt{\Lambda}_\nu^{\gamma_2}(i_2)$ and they satisfies (7.22), (7.25). We now estimate the operator $\Delta_{12}\mathcal{R}_{\nu+1}$. The estimate for $\Delta_{12}\mathcal{Q}_{\nu+1}$ is analogous. By Lemma 7.7 we may construct the operators $\Psi_{\nu,1}(i_1)$, $\Psi_{\nu,2}(i_1)$, $\Psi_{\nu,1}(i_2)$, $\Psi_{\nu,2}(i_2)$, defined for all $\omega\in\mathtt{\Lambda}_{\nu+1}^{\gamma_1}(i_1)\cap\mathtt{\Lambda}_{\nu+1}^{\gamma_2}(i_2)$ and

$$\begin{aligned}|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\Psi_{\nu,1}|||_{\mathcal{L}(H^{s_0})} &\overset{(7.62)}{\leq_{\mathtt{b}}} CN_\nu^{2\tau}\gamma^{-1}\big(|||\langle\partial_\varphi\rangle^{\mathtt{b}}\mathcal{R}_\nu(i_2)|||_{\mathcal{L}(H^{s_0})}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})+\sigma} + |||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\mathcal{R}_\nu|||_{\mathcal{L}(H^{s_0})}\big) \\ &\overset{(2.66),(7.22),(7.12)}{\leq_{S,\mathtt{b}}} N_\nu^{2\tau}N_{\nu-1}\varepsilon\gamma^{-2}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})+\sigma} + N_\nu^{2\tau}\gamma^{-1}|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\mathcal{R}_\nu|||_{\mathcal{L}(H^{s_0})} \\ &\overset{(7.29)}{\leq_{S,\mathtt{b}}} N_\nu^{2\tau}N_{\nu-1}\varepsilon\gamma^{-2}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})+\sigma} \end{aligned} \tag{7.76}$$

and by (7.61), (2.66), (7.22), (7.28) we get

$$|||\Delta_{12}\Psi_{\nu,1}|||_{\mathcal{L}(H^{s_0})} \leq_{S,\mathtt{b}} N_\nu^{2\tau} N_{\nu-1}^{-\mathtt{a}} \varepsilon\gamma^{-2} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma} . \tag{7.77}$$

Similarly one can prove that $\Delta_{12}\Psi_{\nu,2}$ satisfies (7.76), (7.77). By (7.68), for $\varepsilon\gamma^{-2}$ small enough, the smallness condition (2.86) is verified. Therefore by (7.76), (7.77), Lemma 2.20 and (7.70), (2.66) we get

$$|||\Delta_{12}\check{\Psi}_{\nu,1}|||_{\mathcal{L}(H^{s_0})}, |||\Delta_{12}\check{\Psi}_{\nu,2}|||_{\mathcal{L}(H^{s_0})} \leq_{S,\mathtt{b}} N_\nu^{2\tau} N_{\nu-1}^{-\mathtt{a}} \varepsilon\gamma^{-2} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma} , \tag{7.78}$$

$$|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\check{\Psi}_{\nu,1}|||_{\mathcal{L}(H^{s_0})}, |||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\check{\Psi}_{\nu,2}|||_{\mathcal{L}(H^{s_0})} \leq_{S,\mathtt{b}} N_\nu^{2\tau} N_{\nu-1} \varepsilon\gamma^{-2} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma} . \tag{7.79}$$

We now estimate $\Delta_{12}\mathcal{R}_{\nu+1}$ where $\mathcal{R}_{\nu+1}$ is defined in (7.72). We consider the term $\mathcal{R}^\star_{\nu+1} := (\mathrm{Id} + \check{\Psi}_{\nu,1})(\Pi^\perp_{N_\nu}\mathcal{R}_\nu + \mathcal{R}_\nu\Psi_{\nu,1})$. The other terms in (7.72) satisfy the same estimate. One has

$$\begin{aligned}\Delta_{12}\mathcal{R}^\star_{\nu+1} = \; & \Delta_{12}\check{\Psi}_{\nu,1}\big(\Pi^\perp_{N_\nu}\mathcal{R}_\nu(i_1) + \mathcal{R}_\nu(i_1)\Psi_{\nu,1}(i_1)\big) \\ & + (\mathrm{Id} + \check{\Psi}_{\nu,1}(i_2))\big(\Pi^\perp_{N_\nu}\Delta_{12}\mathcal{R}_\nu + \Delta_{12}\mathcal{R}_\nu\Psi_{\nu,1}(i_1) + \mathcal{R}_\nu(i_2)\Delta_{12}\Psi_{\nu,1}\big) .\end{aligned} \tag{7.80}$$

Hence by Lemma 2.19, (7.78), (7.70), (7.61), (2.66), (7.60), taking $\varepsilon\gamma^{-2}$ small enough, we get

$$\begin{aligned}|||\Delta_{12}\mathcal{R}^\star_{\nu+1}|||_{\mathcal{L}(H^{s_0})} \leq_{\mathtt{b}} & \big(N_\nu^{-\mathtt{b}}\mathfrak{M}^\sharp_\nu(s_0,\mathtt{b}) + N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(s_0)^2\big)\|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma} + \\ & + N_\nu^{-\mathtt{b}}|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\mathcal{R}_\nu|||_{\mathcal{L}(H^{s_0})} + N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(s_0)|||\Delta_{12}\mathcal{R}_\nu|||_{\mathcal{L}(H^{s_0})} .\end{aligned} \tag{7.81}$$

Moreover, using also (7.79), (7.62) and since (7.22), (7.14) imply $N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(s_0) \leq 1$, we get

$$\begin{aligned}|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\mathcal{R}^\star_{\nu+1}|||_{\mathcal{L}(H^{s_0})} \leq_{S,\mathtt{b}} & \big(\varepsilon\gamma^{-1}N_{\nu-1} + \mathfrak{M}^\sharp_\nu(s_0,\mathtt{b})\big)\|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma} \\ & + |||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\mathcal{R}_\nu|||_{\mathcal{L}(H^{s_0})} + N_\nu^{\tau_1}\gamma^{-1}|||\Delta_{12}\mathcal{R}_\nu|||_{\mathcal{L}(H^{s_0})}\mathfrak{M}^\sharp_\nu(s_0,\mathtt{b}) .\end{aligned} \tag{7.82}$$

The other terms in (7.72) may be estimated in the same way, whence $\Delta_{12}\mathcal{R}_{\nu+1}$ satisfies (7.81), (7.82).

We now prove (7.28), (7.29) at the step $\nu+1$. By (7.81), (7.22), (7.7), (7.28), (7.29) we get

$$\begin{aligned}|||\Delta_{12}\mathcal{R}_{\nu+1}|||_{\mathcal{L}(H^{s_0})} & \leq_{S,\mathtt{b}} \big(\varepsilon\gamma^{-1}N_{\nu-1}N_\nu^{-\mathtt{b}} + N_\nu^{\tau_1}\varepsilon^2\gamma^{-3}N_{\nu-1}^{-2\mathtt{a}}\big)\|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma} \\ & \overset{(7.6)}{\leq_{S,\mathtt{b}}} \varepsilon\gamma^{-1}N_\nu^{-\mathtt{a}}\|i_1 - i_2\|_{s_0+\mu(\mathtt{b})+\sigma} .\end{aligned}$$

for $\varepsilon\gamma^{-2} \leq 1$ and $N_0(S,\mathtt{b}) > 0$ large. Hence (7.28) at the step $\nu+1$ is proved. Similarly, by (7.82), (7.22), (7.7), (7.28), (7.29), we get

$$|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\mathcal{R}_{\nu+1}|||_{\mathcal{L}(H^{s_0})} \leq_{S,\mathtt{b}} \varepsilon\gamma^{-1}N_{\nu-1}\big(1+\varepsilon\gamma^{-2}N_\nu^{\tau_1}N_{\nu-1}^{-\mathtt{a}}\big)\|i_1-i_2\|_{s_0+\mu(\mathtt{b})+\sigma} \leq_{S,\mathtt{b}} \varepsilon\gamma^{-1}N_\nu\|i_1-i_2\|_{s_0+\mu(\mathtt{b})+\sigma}$$

by (7.6), $\varepsilon\gamma^{-2} \leq 1$ and taking $N_0 := N_0(S,\mathtt{b}) > 0$ large. Thus (7.29) at the step $\nu+1$ is proved.

The proof of (7.30) at the step $\nu+1$ follows by Lemma 7.8. The estimate (7.31) follows by a telescopic argument using (7.30) and (7.28).

Proof of $(\mathbf{S4})_{\nu+1}$. The proof is the same as that of $(\mathbf{S4})_{\nu+1}$ of Theorem 4.2 in [7]. It uses $(\mathbf{S3})_\nu$. □

7.2 Almost-invertibility of $\mathcal{L}_\omega$

By (6.243) and Theorem 7.5 (applied to $\mathbf{L}_0 = \mathcal{L}^{(3)}_M$) we obtain

$$\mathcal{L}_\omega = \mathbf{W}_{2,n}\mathbf{L}_n\mathbf{W}_{1,n}^{-1} + \mathbf{R}^{(3),\perp}_M , \quad \mathbf{W}_{1,n} := \mathcal{W}_1^\perp\mathbf{U}_n , \quad \mathbf{W}_{2,n} := \mathcal{W}_2^\perp\mathbf{U}_n , \tag{7.83}$$

where the operator $\mathbf{L}_n$ is defined in (7.35) and $\mathbf{R}^{(3),\perp}_M$ (defined in (6.243)) satisfies the estimates (6.244), (6.245). Then (6.236), (7.37), (7.9), (7.10), imply that for all $s_0 \leq s \leq S$

$$\|\mathbf{W}_1^{\pm 1}h\|_s^{k_0,\gamma}, \|\mathbf{W}_2^{\pm 1}h\|_s^{k_0,\gamma} \leq_S \|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma} \tag{7.84}$$

for some $\sigma := \sigma(\tau,\nu,k_0) > 0$.

In order to verify the inversion assumption (5.41)-(5.42) required to construct an approximate inverse (and thus define the successive approximate solution of the Nash-Moser non-linear iteration) we decompose the operator $\mathbf{L}_n$ in (7.35) as

$$\mathbf{L}_n = \mathbf{D}_n^< + \mathbf{R}_n^\perp + \mathbf{R}_n + \mathbf{Q}_n \tag{7.85}$$

where

$$\mathbf{D}_n^< := \Pi_{K_n}\big(\omega\cdot\partial_\varphi \mathbb{I}_2^\perp + \mathrm{i}\mathbf{D}_n\big)\Pi_{K_n} + \Pi_{K_n}^\perp\,, \quad \mathbf{R}_n^\perp := \Pi_{K_n}^\perp\big(\omega\cdot\partial_\varphi \mathbb{I}_2^\perp + \mathrm{i}\mathbf{D}_n\big)\Pi_{K_n}^\perp - \Pi_{K_n}^\perp\,, \tag{7.86}$$

the diagonal operator $\mathbf{D}_n$ are defined in (7.15) (with $\nu = n$), and the constant K_n in (1.39).

Lemma 7.11. **(First order Melnikov non-resonance conditions)** *For all $\lambda = (\omega,\kappa)$ in*

$$\Lambda_{n+1}^{\gamma,I} := \Lambda_{n+1}^{\gamma,I}(i) := \big\{\lambda \in \Lambda_{n+1}^\gamma : |\omega\cdot\ell + \mu_j^n| \geq 2\gamma j^{\frac{3}{2}}\langle\ell\rangle^{-\tau}\,, \quad \forall|\ell|\leq K_n\,,\ j\in\mathbb{N}\setminus\mathbb{S}^+\big\} \tag{7.87}$$

(recall (7.34)), the operator $\mathbf{D}_n^<$ in (7.86) is invertible and

$$\|(\mathbf{D}_n^<)^{-1}g\|_s^{k_0,\gamma} \leq_{k_0} \gamma^{-1}\|g\|_{s+\tau_1}^{k_0,\gamma}\,, \quad \tau_1 := \tau + k_0(\tau+1)\,. \tag{7.88}$$

Proof. The estimate (7.88) follows by $\big|\partial_\lambda^k(\omega\cdot\ell + \mu_j^n(\lambda))^{-1}\big| \leq C(k)\langle\ell\rangle^{\tau(|k|+1)+|k|}\gamma^{-(|k|+1)}$, $\forall|k|\leq k_0$. □

Standard smoothing properties imply that the operator $\mathbf{R}_n^\perp$ defined in (7.86) satisfies, for all $b>0$,

$$\|\mathbf{R}_n^\perp h\|_{s_0}^{k_0,\gamma} \lessdot K_n^{-b}\|h\|_{s_0+b+\frac{3}{2}}^{k_0,\gamma}\,, \quad \|\mathbf{R}_n^\perp h\|_s^{k_0,\gamma} \lessdot \|h\|_{s+\frac{3}{2}}^{k_0,\gamma}\,. \tag{7.89}$$

By the decompositions (7.83), (7.85), Theorem 7.5, Proposition 6.31, the estimates (7.88), (7.89), (7.84) we deduce the following theorem:

Theorem 7.12. **(Almost invertibility of $\mathcal{L}_\omega$)** *Assume (5.9) and that, for all $S > s_0$, the smallness condition (7.33) holds. Let $\mathtt{a},\mathtt{b}$ as in (7.6). Then for all*

$$(\omega,\kappa) \in \mathbf{\Lambda}_{n+1}^\gamma := \mathbf{\Lambda}_{n+1}^\gamma(i) := \Lambda_{n+1}^\gamma \cap \Lambda_{n+1}^{\gamma,I} \tag{7.90}$$

(see (7.34), (7.87)) the operator $\mathcal{L}_\omega$ defined in (5.40) (see also (6.7)) can be decomposed as

$$\mathcal{L}_\omega = \mathbf{L}_\omega + \mathbf{R}_\omega + \mathbf{R}_\omega^\perp\,, \tag{7.91}$$

$$\mathbf{L}_\omega := \mathbf{W}_{2,n}\mathbf{D}_n^<\mathbf{W}_{1,n}^{-1}\,, \quad \mathbf{R}_\omega := \mathbf{W}_{2,n}(\mathbf{R}_n + \mathbf{Q}_n)\mathbf{W}_{1,n}^{-1}\,, \quad \mathbf{R}_\omega^\perp := \mathbf{W}_{2,n}\mathbf{R}_n^\perp\mathbf{W}_{1,n}^{-1} + \mathbf{R}_M^{(3),\perp}\,,$$

where $\mathbf{L}_\omega$ is invertible and, for some $\sigma := \sigma(\nu,\tau,k_0) > 0$, for all $s_0 \leq s \leq S$, $g \in H^{s+\sigma}$,

$$\|\mathbf{L}_\omega^{-1}g\|_s^{k_0,\gamma} \leq_S \gamma^{-1}\big(\|g\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma+\mu(\mathtt{b})}^{k_0,\gamma}\|g\|_{s_0+\sigma}^{k_0,\gamma}\big) \tag{7.92}$$

(with $\mu(\mathtt{b})$ defined in (7.10)) and

$$\|\mathbf{R}_\omega h\|_s^{k_0,\gamma} \leq_S \varepsilon\gamma^{-1}N_{n-1}^{-\mathtt{a}}\big(\|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma+\mu(\mathtt{b})}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,, \tag{7.93}$$

$$\|\mathbf{R}_\omega^\perp h\|_{s_0}^{k_0,\gamma} \leq_S K_n^{-b}\big(\|h\|_{s_0+b+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma+\mu(\mathtt{b})+b}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,, \qquad \forall b>0\,, \tag{7.94}$$

$$\|\mathbf{R}_\omega^\perp h\|_s^{k_0,\gamma} \leq_S \|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma+\mu(\mathtt{b})}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\,. \tag{7.95}$$

We finally remark that the operators

$$\mathbf{W}_{1,\infty} := \mathcal{W}_1^\perp\mathbf{U}_\infty\,, \qquad \mathbf{W}_{2,\infty} := \mathcal{W}_2^\perp\mathbf{U}_\infty \quad \text{where} \quad \mathbf{U}_\infty := \lim_{n\to+\infty}\mathbf{U}_n \tag{7.96}$$

see (7.32), and $\mathcal{W}_1^\perp$, $\mathcal{W}_2^\perp$ are defined in (6.235), (6.234) completely diagonalize the linearized operator $\mathcal{L}_\omega$ defined in (5.40). We deduce that $\mathbf{W}_{1,\infty}(\varphi)$, $\mathbf{W}_{2,\infty}(\varphi)$ satisfy the tame estimates (1.26)-(1.27) by small modifications of the arguments of sections 6-7.

8 The Nash-Moser iteration

In this section we prove Theorem 4.1. It will be a consequence of Theorem 8.2 below where we construct iteratively a sequence of better and better approximate solutions of the operator $\mathcal{F}(i,\alpha)$ defined in (4.17). We consider the finite-dimensional subspaces

$$E_n := \Big\{ \mathfrak{I}(\varphi) = (\Theta, I, z)(\varphi), \quad \Theta = \Pi_n \Theta, \ I = \Pi_n I, \ z = \Pi_n z \Big\}$$

where Π_n is the projector

$$\Pi_n := \Pi_{K_n} : z(\varphi, x) = \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{S}_0^c} z_{\ell j} e^{\mathrm{i}(\ell\cdot\varphi + jx)} \mapsto \Pi_n z(\varphi, x) := \sum_{|(\ell,j)| \le K_n} z_{\ell j} e^{\mathrm{i}(\ell\cdot\varphi + jx)} \tag{8.1}$$

with $K_n = K_0^{\chi^n}$ (see (1.39) and (5.43)) and we denote with the same symbol also $\Pi_n p(\varphi) := \sum_{|\ell| \le K_n} p_\ell e^{\mathrm{i}\ell\cdot\varphi}$.

We also define $\Pi_n^\perp := \mathrm{Id} - \Pi_n$. The projectors Π_n, $\Pi_n^\perp$ satisfy the smoothing properties (2.8) for the weighted Sobolev norm defined in (2.5).

In view of the Nash-Moser Theorem 8.2 we introduce the constants

$$\mathtt{a}_1 := \max\{6\sigma_1 + 13, \chi(pk_0(\tau+2) + p\tau + \mu(\mathtt{b}) + 2\sigma_1) + 1\}, \mathtt{a}_2 := \chi^{-1}\mathtt{a}_1 - pk_0(\tau+2) - \mu(\mathtt{b}) - 2\sigma_1 \tag{8.2}$$

$$\mathtt{b}_1 := \mathtt{a}_1 + \mu(\mathtt{b}) + 3\sigma_1 + 3 + \chi^{-1}\mu_1\,, \qquad \mu_1 := 3(\mu(\mathtt{b}) + 2\sigma_1) + 1\,, \qquad \chi = 3/2\,, \tag{8.3}$$

$$\sigma_1 := \max\{\bar\sigma\,, \sigma\,, s_0 + 2k_0 + 5\}\,, \tag{8.4}$$

where $\bar\sigma := \bar\sigma(\tau,\nu,k_0) > 0$ is defined in Theorem 5.10, $\sigma = \sigma(\tau,\nu,k_0) > 0$ is the constant which appears in Theorem 7.3-$(\mathbf{S3})_\nu$-$(\mathbf{S4})_\nu$, $s_0 + 2k_0 + 5$ is the largest loss of regularity in the estimates of the Hamiltonian vector field X_P in Lemma 5.1, $\mu(\mathtt{b})$ in (7.10), the constant $\mathtt{b} := [\mathtt{a}] + 2 \in \mathbb{N}$ where $\mathtt{a}$ is defined in (7.6), and the exponent p in (5.43) satisfies

$$p\mathtt{a} > (\chi - 1)\mathtt{a}_1 + \chi\sigma_1 = \frac{1}{2}\mathtt{a}_1 + \frac{3}{2}\sigma_1\,. \tag{8.5}$$

By remark 7.1 the constant $\mathtt{a} \ge \chi k_0(\tau+2) + 1$. Hence, by the definition of $\mathtt{a}_1$ in (8.2), there exists $p := p(\tau,\nu,k_0)$ such that (8.5) holds. For example we fix

$$p := \max\Big\{ \frac{5\sigma_1 + 7}{\chi k_0(\tau+2) + 1}, \frac{\chi(\mu(\mathtt{b}) + 2\sigma_1) + 1}{\chi k_0 + 1} \Big\}\,. \tag{8.6}$$

Remark 8.1. The constant $\mathtt{a}_1$ is the exponent in (8.11). The constant $\mathtt{a}_2$ is the exponent in (8.9). The constant μ_1 is the exponent in $(\mathcal{P}3)_n$. The conditions $\mathtt{a}_1 > (2\sigma_1 + 4)\chi/(2-\chi) = 6\sigma_1 + 12$, $\mathtt{b}_1 > \mathtt{a}_1 + \mu(\mathtt{b}) + 3\sigma_1 + 2 + \chi^{-1}\mu_1$, as well as $p\mathtt{a} > (\chi-1)\mathtt{a}_1 + \chi\sigma_1$, $\mu_1 > (\mu(\mathtt{b}) + 2\sigma_1)\chi/(\chi-1) = 3(\mu(\mathtt{b}) + 2\sigma_1)$ arise for the convergence of the iterative scheme (8.23)-(8.24), see Lemma 8.4. In addition we require $\mathtt{a}_1 \ge \chi(pk_0(\tau+2) + \mu(\mathtt{b}) + 2\sigma_1) + \chi p\tau + 1$ so that $\mathtt{a}_2 > p\tau$, more precisely $\mathtt{a}_2 \ge p\tau + \chi^{-1}$. This condition is used in the proof of Lemma 8.6. □

Theorem 8.2. (Nash-Moser) *There exist δ_0, $C_* > 0$, such that, if*

$$K_0^{\tau_2}\varepsilon\gamma^{-2} < \delta_0, \quad \tau_2 := \max\{p\tau_0, 2\sigma_1 + \mathtt{a}_1 + 4\}\,, \quad K_0 := \gamma^{-1}, \quad \gamma := \varepsilon^a\,, \quad 0 < a < \frac{1}{2+\tau_2}\,, \tag{8.7}$$

where $\tau_0 := \tau_0(\tau,\nu)$ is defined in Theorem 7.3, then, for all $n \ge 0$:

$(\mathcal{P}1)_n$ *there exists a k_0-times differentiable function $\tilde W_n : \mathbb{R}^\nu \times [\kappa_1,\kappa_2] \to E_{n-1} \times \mathbb{R}^\nu$, $\lambda = (\omega,\kappa) \mapsto \tilde W_n(\lambda) := (\tilde{\mathfrak{I}}_n, \tilde\alpha_n - \omega)$, for $n \ge 1$, and $\tilde W_0 := 0$, satisfying*

$$\|\tilde W_n\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \le C_* K_0^{pk_0(\tau+2)}\varepsilon\gamma^{-1}\,. \tag{8.8}$$

Let $\tilde U_n := U_0 + \tilde W_n$ where $U_0 := (\varphi,0,0,\omega)$. The difference $\tilde H_n := \tilde U_n - \tilde U_{n-1}$, $n \ge 1$, satisfies

$$\|\tilde H_1\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \le C_*\varepsilon\gamma^{-1}K_0^{pk_0(\tau+2)}\,, \quad \|\tilde H_n\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \le C_*\varepsilon\gamma^{-1}K_{n-1}^{-\mathtt{a}_2}\,, \quad \forall n > 1\,. \tag{8.9}$$

$(\mathcal{P}2)_n$ *Setting* $\tilde{\imath}_n := (\varphi, 0, 0) + \tilde{\mathfrak{I}}_n$ *we define*

$$\mathcal{G}_0 := \Omega \times [\kappa_1, \kappa_2] \,, \quad \mathcal{G}_{n+1} := \mathcal{G}_n \bigcap \mathbf{\Lambda}^{\gamma}_{n+1}(\tilde{\imath}_n) \,, \quad n \geq 0 \,, \tag{8.10}$$

where $\mathbf{\Lambda}^{\gamma}_{n+1}(\tilde{\imath}_n)$ *is defined in* (7.90). *Then, for all* λ *in* $\mathcal{N}(\mathcal{G}_n, \gamma K_{n-1}^{-p(\tau+2)})$*, setting* $\gamma_{-1} = \gamma$ *and* $K_{-1} := 1$*, we have*

$$\|\mathcal{F}(\tilde{U}_n)\|_{s_0}^{k_0,\gamma} \leq C_* \varepsilon K_{n-1}^{-\mathtt{a}_1} \,. \tag{8.11}$$

$(\mathcal{P}3)_n$ (High norms). $\|\tilde{W}_n\|_{s_0+\mathtt{b}_1}^{k_0,\gamma} \leq C_* \varepsilon \gamma^{-1} K_{n-1}^{\mu_1}$ *for all* $\omega \in \mathcal{N}(\mathcal{G}_n, \gamma K_{n-1}^{-p(\tau+2)})$.

Proof. To simplify notations, in this proof we denote $\| \, \|^{k_0,\gamma}$ by $\| \, \|$.

Step 1: *Proof of* $(\mathcal{P}1,2,3)_0$. They follow by $\|\mathcal{F}(U_0)\|_s = O(\varepsilon)$ and taking C_* large enough.

Step 2: *Assume that* $(\mathcal{P}1,2,3)_n$ *hold for some* $n \geq 0$*, and prove* $(\mathcal{P}1,2,3)_{n+1}$. We are going to define the successive approximation $\tilde{U}_{n+1}$ by a modified Nash-Moser scheme. For that we prove the almost-approximate invertibility of the linearized operator

$$L_n := L_n(\lambda) := d_{i,\alpha} \mathcal{F}(\tilde{\imath}_n(\lambda))$$

applying Theorem 5.10 to $L_n(\lambda)$. The verification of the inversion assumption (5.41)-(5.42) is the purpose of Theorem 7.12 that we apply with $i = \tilde{\imath}_n$. By (8.7) the smallness condition (7.33) holds for ε small enough. Therefore Theorem 7.12 applies, and we deduce that the inversion assumption (5.41)-(5.42) holds for all $\lambda \in \mathbf{\Lambda}^{\gamma/2}_{n+1}(\tilde{\imath}_n)$, see (7.90). Actually the inversion assumption holds for all $\lambda \in \mathcal{N}(\mathbf{\Lambda}^{\gamma}_{n+1}(\tilde{\imath}_n), 2\gamma K_n^{-p(\tau+2)})$ because

$$\mathcal{N}\big(\mathbf{\Lambda}^{\gamma}_{n+1}(\tilde{\imath}_n), 2\gamma K_n^{-p(\tau+2)}\big) \subseteq \mathbf{\Lambda}^{\gamma/2}_{n+1}(\tilde{\imath}_n) \,, \quad \forall n \geq 0 \,,$$

which is a consequence of (7.16) and the similar inclusion $\mathcal{N}(\mathtt{\Lambda}^{\gamma,I}_{n+1}(\tilde{\imath}_n), 2\gamma K_n^{-p(\tau+2)}) \subseteq \mathtt{\Lambda}^{\gamma/2,I}_{n+1}(\tilde{\imath}_n)$.

Now we apply Theorem 5.10 to the linearized operator $L_n(\lambda)$ with $\mathtt{\Lambda}_o = \mathcal{N}(\mathbf{\Lambda}^{\gamma}_{n+1}(\tilde{\imath}_n), 2\gamma K_n^{-p(\tau+2)})$ and

$$S := s_0 + \mathtt{b}_1 \quad \text{where } \mathtt{b}_1 \text{ is defined in (8.3)} \,. \tag{8.12}$$

It implies the existence of an almost-approximate inverse $\mathbf{T}_n := \mathbf{T}_n(\lambda, \tilde{\imath}_n(\lambda))$ which satisfies

$$\|\mathbf{T}_n g\|_s \leq_{s_0+\mathtt{b}_1} \gamma^{-1} \big(\|g\|_{s+\sigma_1} + \|\tilde{\mathfrak{I}}_n\|_{s+\sigma_1+\mu(\mathtt{b})} \|g\|_{s_0+\sigma_1} \big) \,, \quad \forall s_0 < s \leq s_0 + \mathtt{b}_1 \tag{8.13}$$

$$\|\mathbf{T}_n g\|_{s_0} \leq_{s_0+\mathtt{b}_1} \gamma^{-1} \|g\|_{s_0+\sigma_1} \,. \tag{8.14}$$

For all

$$\lambda \in \mathcal{N}(\mathcal{G}_{n+1}, 2\gamma K_n^{-p(\tau+2)}) \subset \mathcal{N}(\mathcal{G}_n, \gamma K_{n-1}^{-p(\tau+2)}) \,, \ n \geq 0 \,, \tag{8.15}$$

we define the successive approximation

$$U_{n+1} := \tilde{U}_n + H_{n+1} \,, \quad H_{n+1} := (\widehat{\mathfrak{I}}_{n+1}, \widehat{\alpha}_{n+1}) := -\mathbf{\Pi}_n \mathbf{T}_n \Pi_n \mathcal{F}(\tilde{U}_n) \in E_n \times \mathbb{R}^{\nu} \tag{8.16}$$

where $\mathbf{\Pi}_n$ is defined by (see (8.1))

$$\mathbf{\Pi}_n(\mathfrak{I}, \alpha) := (\Pi_n \mathfrak{I}, \alpha) \,, \quad \mathbf{\Pi}_n^{\perp}(\mathfrak{I}, \alpha) := (\Pi_n^{\perp} \mathfrak{I}, 0) \,, \quad \forall (\mathfrak{I}, \alpha) \,. \tag{8.17}$$

We now show that the iterative scheme in (8.16) is rapidly converging. We write

$$\mathcal{F}(U_{n+1}) = \mathcal{F}(\tilde{U}_n) + L_n H_{n+1} + Q_n$$

where $L_n := d_{i,\alpha} \mathcal{F}(\tilde{\imath}_n)$ and

$$Q_n := Q(\tilde{U}_n, H_{n+1}) \,, \quad Q(\tilde{U}_n, H) := \mathcal{F}(\tilde{U}_n + H) - \mathcal{F}(\tilde{U}_n) - L_n H \,, \quad H \in E_n \times \mathbb{R}^{\nu} \,. \tag{8.18}$$

Then, by the definition of H_{n+1} in (8.16), we have (recall also (8.17))

$$\begin{aligned}\mathcal{F}(U_{n+1}) &= \mathcal{F}(\tilde U_n) - L_n \mathbf{\Pi}_n \mathbf{T}_n \Pi_n \mathcal{F}(\tilde U_n) + Q_n \\ &= \mathcal{F}(\tilde U_n) - L_n \mathbf{T}_n \Pi_n \mathcal{F}(\tilde U_n) + L_n \mathbf{\Pi}_n^\perp \mathbf{T}_n \Pi_n \mathcal{F}(\tilde U_n) + Q_n \\ &= \mathcal{F}(\tilde U_n) - \Pi_n L_n \mathbf{T}_n \Pi_n \mathcal{F}(\tilde U_n) + (L_n \mathbf{\Pi}_n^\perp - \Pi_n^\perp L_n) \mathbf{T}_n \Pi_n \mathcal{F}(\tilde U_n) + Q_n \\ &= \Pi_n^\perp \mathcal{F}(\tilde U_n) + R_n + Q_n + P_n \end{aligned} \tag{8.19}$$

where

$$R_n := (L_n \mathbf{\Pi}_n^\perp - \Pi_n^\perp L_n) \mathbf{T}_n \Pi_n \mathcal{F}(\tilde U_n) \, , \qquad P_n := -\Pi_n (L_n \mathbf{T}_n - \mathrm{Id}) \Pi_n \mathcal{F}(\tilde U_n) \, . \tag{8.20}$$

We first note that, for all $\lambda \in \Omega \times [\kappa_1, \kappa_2]$, $s \geq s_0$,

$$\| \mathcal{F}(\tilde U_n) \|_s \leq_s \| \mathcal{F}(U_0) \|_s + \| \mathcal{F}(\tilde U_n) - \mathcal{F}(U_0) \|_s \overset{(4.17),(5.3),(8.4),(8.8)}{\leq_s} \varepsilon + \| \tilde W_n \|_{s+\sigma_1} \tag{8.21}$$

and, by (8.8), (8.7),

$$\gamma^{-1} \| \mathcal{F}(\tilde U_n) \|_{s_0} \leq 1 \, . \tag{8.22}$$

Lemma 8.3. *For all $\lambda \in \mathcal{N}(\mathcal{G}_{n+1}, 2\gamma K_n^{-p(\tau+2)})$ we have, setting $\mu_2 := \mu(\mathtt{b}) + 3\sigma_1 + 2$,*

$$\| \mathcal{F}(U_{n+1}) \|_{s_0} \leq_{s_0+\mathtt{b}_1} \frac{1}{\gamma} K_n^{\mu_2 - \mathtt{b}_1} (\varepsilon + \| \tilde W_n \|_{s_0+\mathtt{b}_1}) + \frac{K_n^{2\sigma_1+4}}{\gamma} \| \mathcal{F}(\tilde U_n) \|_{s_0}^2 + K_{n-1}^{-p\mathtt{a}} K_n^{\sigma_1} \frac{\varepsilon}{\gamma^2} \| \mathcal{F}(\tilde U_n) \|_{s_0} \tag{8.23}$$

$$\| W_1 \|_{s_0+\mathtt{b}_1} \leq_{s_0+\mathtt{b}_1} \varepsilon \gamma^{-1} \, , \quad \| W_{n+1} \|_{s_0+\mathtt{b}_1} \leq_{s_0+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+2\sigma_1} \gamma^{-1} (\varepsilon + \| \tilde W_n \|_{s_0+\mathtt{b}_1}) \, , \ n \geq 1 \, . \tag{8.24}$$

Proof. We first estimate H_{n+1} defined in (8.16).

Estimates of H_{n+1}. By (8.16) and (2.8), (8.13), (8.14), (8.8), we get

$$\begin{aligned} \| H_{n+1} \|_{s_0+\mathtt{b}_1} &\leq_{s_0+\mathtt{b}_1} \gamma^{-1} \big(K_n^{\sigma_1} \| \mathcal{F}(\tilde U_n) \|_{s_0+\mathtt{b}_1} + K_n^{\mu(\mathtt{b})+2\sigma_1} \| \tilde{\mathfrak{I}}_n \|_{s_0+\mathtt{b}_1} \| \mathcal{F}(\tilde U_n) \|_{s_0} \big) \\ &\overset{(8.21),(8.22)}{\leq_{s_0+\mathtt{b}_1}} K_n^{\mu(\mathtt{b})+2\sigma_1} \gamma^{-1} \big(\varepsilon + \| \tilde W_n \|_{s_0+\mathtt{b}_1} \big) \, , \end{aligned} \tag{8.25}$$

$$\| H_{n+1} \|_{s_0} \leq_{s_0+\mathtt{b}_1} \gamma^{-1} K_n^{\sigma_1} \| \mathcal{F}(\tilde U_n) \|_{s_0} \, . \tag{8.26}$$

Now we estimate the terms Q_n in (8.18) and P_n, R_n in (8.20) in $\| \ \|_{s_0}$ norm.

Estimate of Q_n. By (8.18), (4.17), (5.4) and (8.8), (2.8), we have the quadratic estimate

$$\| Q(\tilde U_n, H) \|_{s_0} \leq_{s_0} \varepsilon K_n^4 \| \widehat{\mathfrak{I}} \|_{s_0}^2 \, , \ \forall \widehat{\mathfrak{I}} \in E_n \, . \tag{8.27}$$

Then the term Q_n in (8.18) satisfies, by (8.27), (8.26), $\varepsilon \gamma^{-1} \leq 1$,

$$\| Q_n \|_{s_0} \leq_{s_0+\mathtt{b}_1} K_n^{2\sigma_1+4} \gamma^{-1} \| \mathcal{F}(\tilde U_n) \|_{s_0}^2 \, . \tag{8.28}$$

Estimate of P_n. According to (5.62), we write the term P_n in (8.20) as

$$P_n = -\Pi_n (L_n \mathbf{T}_n - \mathrm{Id}) \Pi_n \mathcal{F}(\tilde U_n) = -P_n^{(1)} - P_{n,\omega} - P_{n,\omega}^\perp$$

$$P_n^{(1)} := \Pi_n \mathcal{P}(\tilde\imath_n) \Pi_n \mathcal{F}(\tilde U_n) \, , \quad P_{n,\omega} := \Pi_n \mathcal{P}_\omega(\tilde\imath_n) \Pi_n \mathcal{F}(\tilde U_n) \, , \quad P_{n,\omega}^\perp := \Pi_n \mathcal{P}_\omega^\perp(\tilde\imath_n) \Pi_n \mathcal{F}(\tilde U_n) \, .$$

By (8.8), (8.7), (8.22), using that, by (2.8),

$$\| \mathcal{F}(\tilde U_n) \|_{s_0+\sigma_1} \leq \| \Pi_n \mathcal{F}(\tilde U_n) \|_{s_0+\sigma_1} + \| \Pi_n^\perp \mathcal{F}(\tilde U_n) \|_{s_0+\sigma_1} \leq K_n^{\sigma_1} \big(\| \mathcal{F}(\tilde U_n) \|_{s_0} + K_n^{-\mathtt{b}_1} \| \mathcal{F}(\tilde U_n) \|_{s_0+\mathtt{b}_1} \big)$$

the bounds (5.63)-(5.66) imply the following estimates:

$$\begin{aligned} \| P_n^{(1)} \|_{s_0} &\leq_{s_0+\mathtt{b}_1} \gamma^{-1} K_n^{2\sigma_1} \big(\| \mathcal{F}(\tilde U_n) \|_{s_0} + K_n^{-\mathtt{b}_1} \| \mathcal{F}(\tilde U_n) \|_{s_0+\mathtt{b}_1} \big) \| \mathcal{F}(\tilde U_n) \|_{s_0} \, , \\ &\overset{(8.21),(2.8)}{\leq_{s_0+\mathtt{b}_1}} \gamma^{-1} K_n^{2\sigma_1} \big(\| \mathcal{F}(\tilde U_n) \|_{s_0} + K_n^{\sigma_1 - \mathtt{b}_1} (\varepsilon + \| \tilde W_n \|_{s_0+\mathtt{b}_1}) \big) \| \mathcal{F}(\tilde U_n) \|_{s_0} \, , \end{aligned} \tag{8.29}$$

$$\| P_{n,\omega} \|_{s_0} \leq_{s_0+\mathtt{b}_1} \varepsilon \gamma^{-2} N_{n-1}^{-\mathtt{a}} K_n^{\sigma_1} \| \mathcal{F}(\tilde U_n) \|_{s_0} \, , \tag{8.30}$$

$$\begin{aligned} \| P_{n,\omega}^\perp \|_{s_0} &\leq_{s_0+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+2\sigma_1-\mathtt{b}_1} \gamma^{-1} (\| \mathcal{F}(\tilde U_n) \|_{s_0+\mathtt{b}_1} + \varepsilon \| \tilde{\mathfrak{I}}_n \|_{s_0+\mathtt{b}_1}) \\ &\overset{(8.21),(2.8)}{\leq_{s_0+\mathtt{b}_1}} K_n^{\mu(\mathtt{b})+3\sigma_1-\mathtt{b}_1} \gamma^{-1} (\varepsilon + \| \tilde W_n \|_{s_0+\mathtt{b}_1}) \, . \end{aligned} \tag{8.31}$$

Estimate of R_n. For $H := (\widehat{\mathfrak{I}}, \widehat{\alpha})$ we have $(L_n \mathbf{\Pi}_n^\perp - \Pi_n^\perp L_n)H = \varepsilon[d_i X_P(\tilde{\imath}_n), \Pi_n^\perp]\widehat{\mathfrak{I}} = [\Pi_n, d_i X_P(\tilde{\imath}_n)]\widehat{\mathfrak{I}}$ where X_P is the Hamiltonian vector field of the perturbation P in (4.14), see (4.17). Thus, applying the estimate (5.3), using (2.8) and recalling (8.4), the following estimate holds:

$$\|(L_n \mathbf{\Pi}_n^\perp - \Pi_n^\perp L_n)H\|_{s_0} \leq_{s_0+\mathtt{b}_1} \varepsilon K_n^{-\mathtt{b}_1+\sigma_1+2}\big(\|\widehat{\mathfrak{I}}\|_{s_0+\mathtt{b}_1} + \|\tilde{\mathfrak{I}}_n\|_{s_0+\mathtt{b}_1}\|\widehat{\mathfrak{I}}\|_{s_0+2}\big)\,. \tag{8.32}$$

Hence, applying (8.13), (8.32), (8.7), (8.8), (2.8), (8.22) the term R_n defined in (8.20) satisfies

$$\begin{aligned} \|R_n\|_{s_0} &\leq_{s_0+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+2\sigma_1+2-\mathtt{b}_1}(\varepsilon\gamma^{-1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0+\mathtt{b}_1} + \varepsilon\|\tilde{\mathfrak{I}}_n\|_{s_0+\mathtt{b}_1}) \\ &\stackrel{(8.21)}{\leq_{s_0+\mathtt{b}_1}} K_n^{\mu(\mathtt{b})+3\sigma_1+2-\mathtt{b}_1}(\varepsilon + \|\tilde{W}_n\|_{s_0+\mathtt{b}_1})\,. \end{aligned} \tag{8.33}$$

We can finally estimate $\mathcal{F}(U_{n+1})$ in $\| \ \|_{s_0}$. By (8.19) and (8.28), (8.29)-(8.31), (8.33), (8.7), (8.8), we get (8.23). Moreover by (8.16) and (8.13) we have the bound (8.24) for

$$\|W_1\|_{s_0+\mathtt{b}_1} = \|H_1\|_{s_0+\mathtt{b}_1} \leq_{s_0+\mathtt{b}_1} \gamma^{-1}\|\mathcal{F}(U_0)\|_{s_0+\mathtt{b}_1+\sigma_1} \leq_{s_0+\mathtt{b}_1} \varepsilon\gamma^{-1}\,.$$

The estimate (8.24) for $W_{n+1} := \tilde{W}_n + H_{n+1}$, $n \geq 1$, follows by (8.25). □

As a corollary we get

Lemma 8.4. *For all $\lambda \in \mathcal{N}(\mathcal{G}_{n+1}, 2\gamma K_n^{-p(\tau+2)})$ we have*

$$\|\mathcal{F}(U_{n+1})\|_{s_0} \leq C_*\varepsilon K_n^{-\mathtt{a}_1}\,, \qquad \|W_{n+1}\|_{s_0+\mathtt{b}_1} \leq C_*\varepsilon\gamma^{-1}K_n^{\mu_1}\,, \tag{8.34}$$

$$\|H_1\|_{s_0+\mu(\mathtt{b})+\sigma_1} \leq C\varepsilon\gamma^{-1}\,, \quad \|H_{n+1}\|_{s_0+\mu(\mathtt{b})+\sigma_1} \leq_{s_0} \varepsilon\gamma^{-1}K_n^{\mu(\mathtt{b})+2\sigma_1}K_{n-1}^{-\mathtt{a}_1}\,, \quad n \geq 1\,. \tag{8.35}$$

Proof. First note that, by (8.15), if $\lambda \in \mathcal{N}(\mathcal{G}_{n+1}, 2\gamma K_n^{-p(\tau+2)})$ then $\lambda \in \mathcal{N}(\mathcal{G}_n, \gamma K_{n-1}^{-p(\tau+2)})$ and so (8.11) and $(\mathcal{P}3)_n$ hold. Then the first inequality in (8.34) follows by (8.23), $(\mathcal{P}2)_n$, $(\mathcal{P}3)_n$, $\gamma^{-1} = K_0 \leq K_n$, $\varepsilon\gamma^{-2} \leq c$ small, and by (8.2), (8.3), (8.5)-(8.6) (see also remark 8.1). For $n = 0$ we use also (8.7). The second inequality in (8.34) follows similarly by (8.24), $(\mathcal{P}3)_n$, the choice of μ_1 in (8.3) and K_0 large enough. Since $H_1 = W_1$ the first inequality in (8.35) follows by the first inequality in (8.24). For $n \geq 1$, the estimate (8.35) follows by (2.8), (8.26) and (8.11). □

We now define a k_0-times differentiable extension of $(H_{n+1})_{|\mathcal{N}(\mathcal{G}_{n+1}, \gamma K_n^{-p(\tau+2)})}$ to the whole $\mathbb{R}^\nu \times [\kappa_1, \kappa_2]$.

Lemma 8.5. (Extension) *There is a k_0-times differentiable function $\tilde{H}_{n+1}$ defined on the whole $\mathbb{R}^\nu \times [\kappa_1, \kappa_2]$ such that*

$$\tilde{H}_{n+1} = H_{n+1}\,, \quad \forall \lambda \in \mathcal{N}(\mathcal{G}_{n+1}, \gamma K_n^{-p(\tau+2)})\,, \tag{8.36}$$

and (8.9) *holds also at the step $n+1$.*

Proof. The function $H_{n+1}(\lambda)$ is defined for all $\lambda \in \mathcal{N}\big(\mathcal{G}_{n+1}, 2\gamma K_n^{-p(\tau+2)}\big)$. Then we define

$$\tilde{H}_{n+1}(\lambda) := \begin{cases} \psi_{n+1}(\lambda)H_{n+1}(\lambda) & \forall \lambda \in \mathcal{N}\big(\mathcal{G}_{n+1}, 2\gamma K_n^{-p(\tau+2)}\big) \\ 0 & \forall \lambda \notin \mathcal{N}\big(\mathcal{G}_{n+1}, 2\gamma K_n^{-p(\tau+2)}\big) \end{cases}$$

where ψ_{n+1} is a $\mathcal{C}^\infty$ cut-off function satisfying $0 \leq \psi_{n+1} \leq 1$,

$$\begin{gathered} \psi_{n+1}(\lambda) = 1\,,\ \forall \lambda \in \mathcal{N}(\mathcal{G}_{n+1}, \gamma K_n^{-p(\tau+2)}), \quad \mathrm{supp}(\psi_{n+1}) \subseteq \mathcal{N}\big(\mathcal{G}_{n+1}, 2\gamma K_n^{-p(\tau+2)}\big)\,, \\ |\partial_\lambda^k \psi_{n+1}(\lambda)| \leq C(k)\big(K_n^{p(\tau+2)}\gamma^{-1}\big)^{|k|},\ \forall k \in \mathbb{N}^{\nu+1}\,. \end{gathered}$$

Then (8.36) holds and we have the estimate $\|\tilde{H}_{n+1}\|_{s_0+\mu(\mathtt{b})+\sigma_1} \leq K_n^{p(\tau+2)k_0}\|H_{n+1}\|_{s_0+\mu(\mathtt{b})+\sigma_1}$. For $n = 0$ and (8.35) we get the first inequality in (8.9). For $n \geq 1$ we deduce using (8.35) and the definition of $\mathtt{a}_2$ in (8.2), the estimate (8.9) also at the step $n+1$. □

We now define

$$\tilde W_{n+1} = \tilde W_n + \tilde H_{n+1}\,, \quad \tilde U_{n+1} := \tilde U_n + \tilde H_{n+1} = U_0 + \tilde W_n + \tilde H_{n+1} = U_0 + \tilde W_{n+1}\,,$$

which are defined for all $\lambda \in \mathbb{R}^\nu \times [\kappa_1, \kappa_2]$ and satisfy

$$\tilde W_{n+1} = W_{n+1}\,, \tilde U_{n+1} = U_{n+1}\,, \quad \forall \lambda \in \mathcal{N}(\mathcal{G}_{n+1}, \gamma K_n^{-p(\tau+2)})\,.$$

Therefore $(\mathcal{P}2)_{n+1}$, $(\mathcal{P}3)_{n+1}$ are proved by Lemma 8.4. Moreover by (8.9), which has been proved up to the step $n+1$ in Lemma 8.5, we have

$$\|\tilde W_{n+1}\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \leq \sum\nolimits_{k=1}^{n+1} \|\tilde H_k\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \leq C_* K_0^{pk_0(\tau+2)} \varepsilon \gamma^{-1}$$

and thus (8.8) holds also at the step $n+1$. This completes the proof of Theorem 8.2. □

8.1 Proof of Theorem 4.1

Let $\gamma = \varepsilon^a$ with $a \in (0, a_0)$ and $a_0 := 1/(2+\tau_2)$. Then the smallness condition (8.7) holds for $0 < \varepsilon < \varepsilon_0$ small enough and Theorem 8.2 holds. By (8.9) the sequence of functions $\tilde U_n := (\tilde\imath_n, \tilde\alpha_n)$ is a Cauchy sequence in $\| \,\|^{k_0,\gamma}_{s_0}$ and we define its limit function

$$U_\infty := (i_\infty, \alpha_\infty) = (\varphi, 0, 0, \omega) + W_\infty\,, \quad W_\infty : \Omega \times [\kappa_1, \kappa_2] \to H^{s_0}_\varphi \times H^{s_0}_\varphi \times H^{s_0}_{\varphi,x} \times \mathbb{R}^\nu\,, \quad W_\infty := \lim_{n\to+\infty} \tilde W_n\,.$$

By (8.8) and (8.9) we also deduce

$$\|U_\infty - U_0\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \leq C_* \varepsilon \gamma^{-1} K_0^{pk_0(\tau+2)}\,, \quad \|U_\infty - \tilde U_n\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \leq C \varepsilon \gamma^{-1} K_n^{-\mathtt{a}_2}\,, \quad \forall n \geq 1\,. \tag{8.37}$$

Moreover by Theorem 8.2-$(\mathcal{P}2)_n$, we deduce that $\mathcal{F}(\lambda, U_\infty(\lambda)) = 0$ for all λ belonging to

$$\bigcap_{n\geq 0} \mathcal{G}_n = \mathtt{\Lambda} \cap \bigcap_{n\geq 1} \mathbf{\Lambda}^\gamma_n(\tilde\imath_{n-1}) \overset{(7.90),(7.34),(7.87)}{=} \mathtt{\Lambda} \cap \Big[\bigcap_{n\geq 1} \mathtt{\Lambda}^\gamma_n(\tilde\imath_{n-1})\Big] \bigcap \Big[\bigcap_{n\geq 1} \mathtt{\Lambda}^{\gamma,I}_n(\tilde\imath_{n-1})\Big]\,, \tag{8.38}$$

where $\mathtt{\Lambda} := \Omega \times [\kappa_1, \kappa_2]$. By (8.37) for $n = 0$ and since $K_0 = \gamma^{-1}$ (see (8.7)) we deduce the estimates (4.21) and (4.22) with $k_1 := pk_0(\tau+2)$.

In order to conclude the proof of Theorem 4.1 we have to provide the caractherization of $\mathcal{C}^\gamma_\infty$ in (4.25). We first consider the set

$$\mathcal{G}_\infty := \mathtt{\Lambda} \cap \Big[\bigcap_{n\geq 1} \mathtt{\Lambda}^{2\gamma}_n(i_\infty)\Big] \bigcap \Big[\bigcap_{n\geq 1} \mathtt{\Lambda}^{2\gamma,I}_n(i_\infty)\Big]\,. \tag{8.39}$$

Lemma 8.6. $\mathcal{G}_\infty \subseteq \bigcap_{n\geq 0} \mathcal{G}_n$ *defined in* (8.10).

Proof. By (8.37), (8.7), we have

$$\varepsilon\gamma^{-1} C(S) N_0^\tau \|i_\infty - i_0\|_{s_0+\mu(\mathtt{b})+\sigma_1} \leq \varepsilon\gamma^{-1} C(S) K_0^{p\tau} C_* \varepsilon \gamma^{-1} K_0^{pk_0(\tau+2)} \leq \gamma$$
$$\varepsilon\gamma^{-1} C(S) N_{n-1}^\tau \|i_\infty - \tilde\imath_{n-1}\|_{s_0+\mu(\mathtt{b})+\sigma_1} \leq \varepsilon\gamma^{-1} C(S) K_{n-1}^{p\tau} C \varepsilon \gamma^{-1} K_n^{-\mathtt{a}_2} \leq \gamma\,,\ \forall n \geq 2\,,$$

noting that the exponent τ_2 in (8.7) satisfies $\tau_2 > \mathtt{a}_1 > 3(pk_0(\tau+2)+p\tau)/2$ by (8.2) and that $\mathtt{a}_2 \geq p\tau + \chi^{-1}$ (see (8.2) and remark 8.1). Recall also that S has been fixed in (8.12) and that $\sigma_1 \geq \sigma$, see (8.4). Therefore Theorem 7.3-$(\mathbf{S4})_\nu$ implies

$$\mathtt{\Lambda}^{2\gamma}_n(i_\infty) \subset \mathtt{\Lambda}^\gamma_n(\tilde\imath_{n-1})\,, \quad \forall n \geq 1\,.$$

By similar arguments we deduce that $\mathtt{\Lambda}^{2\gamma,I}_n(i_\infty) \subset \mathtt{\Lambda}^{\gamma,I}_n(\tilde\imath_{n-1})$ and the lemma is proved. □

Then we define the "final eigenvalues"

$$\mu_j^\infty := \mathtt{m}_3^\infty j^{\frac12}(1+\kappa j^2)^{\frac12} + \mathtt{m}_1^\infty j^{\frac12} + r_j^\infty\,, \quad j \in \mathbb{N}^+ \setminus \mathbb{S}^+\,, \tag{8.40}$$

where

$$\mathtt{m}_3^\infty := \mathtt{m}_3(i_\infty)\,, \quad \mathtt{m}_1^\infty := \mathtt{m}_1(i_\infty)\,, \quad r_j^\infty := \lim_{n\to+\infty} \tilde r_j^n(i_\infty)\,, \quad j \in \mathbb{N}^+ \setminus \mathbb{S}^+\,, \tag{8.41}$$

where $\mathtt{m}_3$, $\mathtt{m}_1$ are defined in (6.72), (6.226) and $\tilde r_j^n$ are given in Theorem 7.3-$(\mathbf{S2})_\nu$. Note that the sequence $(\tilde r_j^n(i_\infty))_{n\in\mathbb{N}}$ is a Cauchy sequence in $|\ |^{k_0,\gamma}$ by (7.27). As a consequence its limit function $r_j^\infty(\omega,\kappa)$ is well defined, it is k_0-times differentiable and satisfies

$$|r_j^\infty - \tilde r_j^n(i_\infty)|^{k_0,\gamma} \le C\varepsilon\gamma^{-1} N_n^{k_0(\tau+2)} N_{n-1}^{-\mathtt{a}}\,,\ n \ge 0\,. \tag{8.42}$$

In particular, since $\tilde r_j^0(i_\infty) = 0$ and $K_0 = \gamma^{-1}$ we get $|r_j^\infty|^{k_0,\gamma} \le C\varepsilon\gamma^{-1} K_0^{pk_0(\tau+2)+1}$ and (4.24) holds with $k_1 = pk_0(\tau+2)+1$ (recall that the constant $C := C(S,k_0)$ with S fixed in (8.12)).

Consider the *final Cantor set* $\mathcal{C}_\infty^\gamma$ in (4.25).

Lemma 8.7. $\mathcal{C}_\infty^\gamma \subseteq \mathcal{G}_\infty$ *defined in* (8.39).

Proof. By (8.39), we have to prove that $\mathcal{C}_\infty^\gamma \subseteq \Lambda_n^{2\gamma}(i_\infty)$, $\forall n \in \mathbb{N}$. We argue by induction. For $n=0$ the inclusion is trivial, since $\Lambda_0^{2\gamma}(i_\infty) = \Omega \times [\kappa_1,\kappa_2] = \Lambda$. Now assume that $\mathcal{C}_\infty^\gamma \subseteq \Lambda_n^{2\gamma}(i_\infty)$. Theorem 7.3-$(\mathbf{S2})_\nu$ implies $\tilde\mu_j^n(i_\infty)(\lambda) = \mu_j^n(i_\infty)(\lambda)$, $\forall \lambda \in \Lambda_n^{2\gamma}(i_\infty)$. Hence $\forall \lambda \in \mathcal{C}_\infty^\gamma \subseteq \Lambda_n^{2\gamma}(i_\infty)$, by (7.17), (8.40), (8.42), we get

$$|(\mu_j^n - \mu_{j'}^n)(i_\infty) - (\mu_j^\infty - \mu_{j'}^\infty)| \le C\varepsilon\gamma^{-1} N_n^{k_0(\tau+2)} N_{n-1}^{-\mathtt{a}}\,,$$

and therefore (consider in (4.25) the case $\varsigma = 1$ and $j \ne j'$)

$$\begin{aligned}|\omega\cdot\ell + \mu_j^n(i_\infty) - \mu_{j'}^n(i_\infty)| &\ge |\omega\cdot\ell + \mu_j^\infty - \mu_{j'}^\infty| - C\varepsilon\gamma^{-1} N_n^{k_0(\tau+2)} N_{n-1}^{-\mathtt{a}} \\ &\ge 4\gamma|j^{\frac32} - j'^{\frac32}|\langle\ell\rangle^{-\tau} - C\varepsilon\gamma^{-1}|j^{\frac32} - j'^{\frac32}| N_n^{k_0(\tau+2)} N_{n-1}^{-\mathtt{a}} \\ &\ge 2\gamma|j^{\frac32} - j'^{\frac32}|\langle\ell\rangle^{-\tau}\,, \quad \forall |\ell| \le N_n\,,\end{aligned}$$

provided $\varepsilon\gamma^{-2} \le C N_{n-1}^{\mathtt{a}} N_n^{-k_0(\tau+2)-\tau}$, $\forall n \ge 0$, which holds true by (7.6), (8.7), see also remark 7.1. We have proved that $\mathcal{C}_\infty^\gamma \subseteq \Lambda_{n+1}^{2\gamma}(i_\infty)$. Similarly we prove that $\mathcal{C}_\infty^\gamma \subseteq \Lambda_n^{2\gamma,I}(i_\infty)$, $\forall n \in \mathbb{N}$. □

Lemmata 8.6, 8.7 imply that

Corollary 8.8. $\mathcal{C}_\infty^\gamma \subseteq \bigcap_{n\ge0} \mathcal{G}_n$ *defined in* (8.10).

9 Appendix: tame estimates for the flow of pseudo-PDEs

In this Appendix we prove tame estimates for the flow Φ^t of the pseudo-PDE

$$\begin{cases}\partial_t u = \mathrm{i}a(\varphi,x)|D|^{\frac12}u \\ u(0,x) = u_0(x)\,,\end{cases} \qquad \varphi \in \mathbb{T}^\nu\,, \quad x \in \mathbb{T}\,, \tag{9.1}$$

where $a(\varphi,x) = a(\lambda,\varphi,x)$ is a real valued function which is $\mathcal{C}^\infty$ with respect to the variables (φ,x) and k_0-times differentiable with respect to the parameters $\lambda = (\omega,\kappa)$. The function $a := a(i)$ may depend also on the "approximate" torus $i(\varphi)$. We look for the solution of (9.1) by a Galerkin approximation, as limit of the solutions of the truncated equations

$$\begin{cases}\partial_t u = \mathrm{i}\Pi_N\big(a(\varphi,x)|D|^{\frac12}\Pi_N u\big) \\ u(0,x) = \Pi_N u_0(x)\,,\end{cases} \qquad \varphi \in \mathbb{T}^\nu\,, \quad x \in \mathbb{T}\,, \tag{9.2}$$

where, for any $N \in \mathbb{N}$, we denote by Π_N the L^2-orthogonal projector on the finite dimensional subspace

$$E_N := \big\{u \in L^2(\mathbb{T}) : u(x) = \sum\nolimits_{|j|\le N} u_j e^{\mathrm{i}jx}\big\}\,.$$

We denote by $\Phi_N(t) = \Phi_N(\lambda, t, \varphi) : E_N \to E_N$ the flow of (9.2). It solves

$$\begin{cases} \partial_t \Phi_N(t) = \mathrm{i}\Pi_N a(\varphi, x)|D|^{\frac{1}{2}}\Phi_N(t) \\ \Phi_N(0) = \Pi_N , \end{cases} \qquad \varphi \in \mathbb{T}^\nu . \tag{9.3}$$

We introduce the "paraproduct" decomposition for the product of two functions $a, u : \mathbb{T} \to \mathbb{C}$,

$$au = T_a u + R_u a \tag{9.4}$$

$$T_a u := \sum_{k,\xi \in \mathbb{Z}, |k-\xi| \le |\xi|} \widehat{a}(k-\xi)\widehat{u}(\xi)e^{\mathrm{i}kx} , \quad R_u a := \sum_{k,\xi \in \mathbb{Z}, |k-\xi| < |\xi|} \widehat{u}(k-\xi)\widehat{a}(\xi)e^{\mathrm{i}kx} . \tag{9.5}$$

Note that

$$T_a = \mathrm{Op}(a_0(x,\xi)) \qquad \text{with} \qquad a_0(x,\xi) := \sum\nolimits_{|k| \le |\xi|} \widehat{a}(k)e^{\mathrm{i}kx} . \tag{9.6}$$

For all $s \ge 0$, we have the following estimates

$$\|T_a u\|_{H^s_x} \le C(s)\|a\|_{H^1_x}\|u\|_{H^s_x} , \qquad \|R_u(a)\|_{H^s_x} \le C(s)\|a\|_{H^{s+(1/2)}_x}\|u\|_{H^{1/2}_x} \tag{9.7}$$

(the operator $u \mapsto R_u(a)$ is smoothing) which follow arguing as in Lemma 2.13.

Lemma 9.1. $\big\||D|^{\frac{1}{2}}(T_a)^* - T_a|D|^{\frac{1}{2}}\big\|_{\mathcal{L}(L^2_x)} \le C\|a\|_{H^2_x}$ *and* $\big\|[\langle D\rangle^s, T_a|D|^{\frac{1}{2}}]u\big\|_{L^2_x} \le_s \|a\|_{H^2_x}\|u\|_{H^s_x}$, $\forall s \ge 0$.

Proof. By (2.31) the adjoint of $T_a = \mathrm{Op}(a_0)$ is the pseudo-differential operator $(T_a)^* = \mathrm{Op}(a_0^*)$ with symbol

$$a_0^*(x,\xi) = \overline{\sum\nolimits_{k\in\mathbb{Z}} \widehat{a}_0(k, \xi-k)e^{\mathrm{i}kx}} \overset{(9.6)}{=} \overline{\sum\nolimits_{|k|\le|\xi-k|} \widehat{a}(k)e^{\mathrm{i}kx}} = \sum\nolimits_{|k|\le|\xi+k|} \widehat{a}(k)e^{\mathrm{i}kx}$$

since $\overline{\widehat{a}(k)} = \widehat{a}(-k)$ because $a(x)$ is real valued. Thus

$$|D|^{\frac{1}{2}}(T_a)^* u = \sum\nolimits_\xi \sum\nolimits_{|k|\le|\xi+k|} |\xi+k|^{\frac{1}{2}}\widehat{a}(k)\widehat{u}(\xi)e^{\mathrm{i}(k+\xi)x} = R_1 + R_2 \tag{9.8}$$

where, writing

$$\vartheta(\xi,k) := |\xi+k|^{\frac{1}{2}} - |\xi|^{\frac{1}{2}} = \frac{|\xi+k| - |\xi|}{|\xi+k|^{\frac{1}{2}} + |\xi|^{\frac{1}{2}}} , \quad \text{for} \quad (\xi,k) \ne (0,0) , \tag{9.9}$$

we split

$$R_1 := \sum_\xi \sum_{|k|\le|\xi+k|} |\xi|^{\frac{1}{2}}\widehat{a}(k)\widehat{u}(\xi)e^{\mathrm{i}(k+\xi)x} , \quad R_2 := \sum_\xi \sum_{|k|\le|\xi+k|} \vartheta(\xi,k)\widehat{a}(k)\widehat{u}(\xi)e^{\mathrm{i}(k+\xi)x} . \tag{9.10}$$

In addition, by (9.5),

$$T_a|D|^{\frac{1}{2}}u(x) = \sum\nolimits_\xi \sum\nolimits_{|k|\le|\xi|} |\xi|^{\frac{1}{2}}\widehat{a}(k)\widehat{u}(\xi)e^{\mathrm{i}(k+\xi)x} . \tag{9.11}$$

We estimate

$$\big(|D|^{\frac{1}{2}}(T_a)^* - T_a|D|^{\frac{1}{2}}\big)u = (R_1 - T_a|D|^{\frac{1}{2}}u) + R_2 . \tag{9.12}$$

Estimate of R_2. By (9.9) the triangular inequality implies $|\vartheta(\xi,k)| \le |k|$, for any $k, \xi \in \mathbb{Z}$. Then by the Cauchy-Schwartz inequality we get

$$\begin{aligned} \|R_2\|^2_{L^2_x} &\le \sum_j \Big(\sum_{|j-\xi|\le|j|} |\vartheta(\xi, j-\xi)||\widehat{a}(j-\xi)||\widehat{u}(\xi)| \Big)^2 \\ &\le \sum_j \Big(\sum_{|j-\xi|\le|j|} |j-\xi||\widehat{a}(j-\xi)||\widehat{u}(\xi)| \frac{\langle j-\xi\rangle}{\langle j-\xi\rangle} \Big)^2 \\ &\le C \sum_j \sum_{|j-\xi|\le|j|} \langle j-\xi\rangle^4 |\widehat{a}(j-\xi)|^2 |\widehat{u}(\xi)|^2 \\ &\le C \sum_\xi |\widehat{u}(\xi)|^2 \sum_j \langle j-\xi\rangle^4 |\widehat{a}(j-\xi)|^2 \le C\|a\|^2_{H^2_x}\|u\|^2_{L^2_x} . \end{aligned} \tag{9.13}$$

Estimate of $R_1 - T_a|D|^{\frac{1}{2}}u$. By (9.10) and (9.11) we write

$$R_1 - T_a|D|^{\frac{1}{2}}u = T_1 - T_2 \tag{9.14}$$

$$T_1 := \sum_{\xi} \sum_{|\xi|<|k|\le|\xi+k|} |\xi|^{\frac{1}{2}} \widehat{a}(k)\widehat{u}(\xi)e^{\mathrm{i}(\xi+k)x}, \quad T_2 := \sum_{\xi} \sum_{|\xi+k|<|k|\le|\xi|} |\xi|^{\frac{1}{2}} \widehat{a}(k)\widehat{u}(\xi)e^{\mathrm{i}(\xi+k)x}.$$

We estimate the L^2_x norm of T_2. The estimate for T_1 is analogous. We have

$$\|T_2\|^2_{L^2_x} \le \sum_j \Big(\sum_{|j|\le|j-\xi|\le|\xi|} |\xi|^{\frac{1}{2}} |\widehat{a}(j-\xi)||\widehat{u}(\xi)| \Big)^2$$

and, since in the sum $|\xi| \le |j| + |\xi - j| \le 2|j-\xi|$, the Cauchy-Schwartz inequality implies

$$\begin{aligned}
\|T_2\|^2_{L^2_x} &\le 4\sum_j \Big(\sum_{|j|\le|j-\xi|\le|\xi|} |j-\xi|^{\frac{1}{2}} |\widehat{a}(j-\xi)||\widehat{u}(\xi)| \frac{\langle j-\xi\rangle}{\langle j-\xi\rangle} \Big)^2 \\
&\le C \sum_j \sum_{|j|\le|j-\xi|\le|\xi|} \langle j-\xi\rangle^3 |\widehat{a}(j-\xi)|^2 |\widehat{u}(\xi)|^2 \\
&\le C \sum_{\xi} |\widehat{u}(\xi)|^2 \sum_j \langle j-\xi\rangle^3 |\widehat{a}(j-\xi)|^2 \le C\|a\|^2_{H^{\frac{3}{2}}_x} \|u\|^2_{L^2_x}. \qquad (9.15)
\end{aligned}$$

The first estimate of Lemma 9.1 follows by (9.12), (9.13), (9.14), (9.15) (and the similar bound for T_1).

Let us prove the second estimate of Lemma 9.1. By (9.11) the commutator

$$[\langle D\rangle^s, T_a|D|^{\frac{1}{2}}]u = \sum\nolimits_{\xi} \sum\nolimits_{|j-\xi|\le|\xi|} \psi(\xi,j)\widehat{a}(j-\xi)\widehat{u}(\xi)e^{\mathrm{i}jx}$$

where $\psi(\xi,j) := (\langle j\rangle^s - \langle\xi\rangle^s)|\xi|^{\frac{1}{2}}$. Since $|j-\xi| \le |\xi|$ we have $|\psi(\xi,j)| \le_s \langle\xi\rangle^s|j-\xi|$. Hence using as before the Cauchy-Schwartz inequality we get

$$\begin{aligned}
\|[\langle D\rangle^s, T_a|D|^{\frac{1}{2}}]u\|^2_{L^2_x} &\le_s \sum_j \Big(\sum_{|j-\xi|\le|\xi|} |\psi(\xi,j)||\widehat{a}(j-\xi)||\widehat{u}(\xi)| \Big)^2 \\
&\le_s \Big(\sum_{|j-\xi|\le|\xi|} \langle\xi\rangle^s |j-\xi||\widehat{a}(j-\xi)||\widehat{u}(\xi)| \frac{\langle j-\xi\rangle}{\langle j-\xi\rangle} \Big)^2 \\
&\le_s \sum_{\xi} \langle\xi\rangle^{2s} |\widehat{u}(\xi)|^2 \sum_j \langle j-\xi\rangle^4 |\widehat{a}(j-\xi)|^2 \le_s \|a\|^2_{H^2_x} \|u\|^2_{H^s_x}.
\end{aligned}$$

The lemma is proved. ∎

Proposition 9.2. *Assume $\|a\|_{s_0+\frac{5}{2}} \le 1$. Then, $\forall \varphi \in \mathbb{T}^\nu$, for all $s \ge 0$ the flow $\Phi^t_N(\varphi)$ of (9.2) satisfies*

$$\sup\nolimits_{t\in[0,1]} \|\Phi^t_N(\varphi)(u_0)\|_{H^s_x} \le C\|u_0\|_{H^s_x}, \qquad \forall 0 \le s \le 1 \tag{9.16}$$

$$\sup\nolimits_{t\in[0,1]} \|\Phi^t_N(\varphi)(u_0)\|_{H^s_x} \le C(s)\big(\|u_0\|_{H^s_x} + \|a\|_{H^{s+\frac{1}{2}}_x} \|u_0\|_{H^1_x}\big), \quad \forall s \ge 1, \tag{9.17}$$

uniformly for all $N \in \mathbb{N}$. The flow of (9.1) is a linear bounded operator $\Phi^t(\varphi) : H^s_x(\mathbb{T}) \to H^s_x(\mathbb{T})$ satisfying

$$\sup\nolimits_{t\in[0,1]} \|\Phi^t(\varphi)(u_0)\|_{H^s_x} \le C\|u_0\|_{H^s_x}, \qquad \forall 0 \le s \le 1 \tag{9.18}$$

$$\sup\nolimits_{t\in[0,1]} \|\Phi^t(\varphi)(u_0)\|_{H^s_x} \le C(s)\big(\|u_0\|_{H^s_x} + \|a\|_{H^{s+\frac{1}{2}}_x} \|u_0\|_{H^1_x}\big), \quad \forall s \ge 1. \tag{9.19}$$

Proof. Proof of (9.16), (9.17).

Step 1. $s = 0$. For any $N \in \mathbb{N}$, the equation (9.2) is an ODE on the finite dimensional space E_N which admits a unique solution $u_N(t) = u_N(\lambda, t, \varphi, \cdot) = \Phi_N^t(u_0) \in E_N$. The L^2_x-norm of the solution $u_N(t)$ satisfies (using that Π_N is L^2 self-adjoint)

$$\begin{aligned}\partial_t \|u_N(t)\|^2_{L^2_x} &= (\mathrm{i}\Pi_N a|D|^{\frac{1}{2}} u_N, u_N)_{L^2_x} + (u_N, \mathrm{i}\Pi_N a|D|^{\frac{1}{2}} u_N)_{L^2_x} \\ &= (\mathrm{i}a|D|^{\frac{1}{2}} u_N, u_N)_{L^2_x} + (u_N, \mathrm{i}a|D|^{\frac{1}{2}} u_N)_{L^2_x} = (\mathrm{i}[a, |D|^{\frac{1}{2}}]u_N, u_N)_{L^2_x} \end{aligned} \tag{9.20}$$

because a is real. Lemma 2.8, (2.13), (2.39), (2.40), and $\|a\|_{s_0+\frac{5}{2}} \leq 1$, imply the commutator estimate $\|[a, |D|^{\frac{1}{2}}]\|_{\mathcal{L}(L^2_x)} \leq C$. Hence $\partial_t \|u_N(t)\|^2_{L^2_x} \leq C\|u_N(t)\|^2_{L^2_x}$ and Gronwall inequality implies (9.16) for $s = 0$.

Step 2. $s \geq 1$. The Sobolev norm $\|u_N\|^2_{H^s_x} = \|\langle D\rangle^s u_N\|^2_{L^2_x}$ satisfies

$$\begin{aligned}\partial_t \|\langle D\rangle^s u_N\|^2_{L^2_x} &= \big(\langle D\rangle^s \Pi_N \mathrm{i}a|D|^{\frac{1}{2}} u_N, \langle D\rangle^s u_N\big)_{L^2_x} + \big(\langle D\rangle^s u_N, \langle D\rangle^s \Pi_N \mathrm{i}a|D|^{\frac{1}{2}} u_N\big)_{L^2_x} \\ &= \big(\langle D\rangle^s \mathrm{i}a|D|^{\frac{1}{2}} u_N, \langle D\rangle^s u_N\big)_{L^2_x} + \big(\langle D\rangle^s u_N, \langle D\rangle^s \mathrm{i}a|D|^{\frac{1}{2}} u_N\big)_{L^2_x} \\ &= \big(\langle D\rangle^s \mathrm{i}T_a(|D|^{\frac{1}{2}} u_N), \langle D\rangle^s u_N\big)_{L^2_x} + \big(\langle D\rangle^s u_N, \langle D\rangle^s \mathrm{i}T_a(|D|^{\frac{1}{2}} u_N)\big)_{L^2_x} \qquad (9.21) \\ &\quad + \big(\langle D\rangle^s \mathrm{i}R_{|D|^{\frac{1}{2}} u_N} a, \langle D\rangle^s u_N\big)_{L^2_x} + \big(\langle D\rangle^s u_N, \langle D\rangle^s \mathrm{i}R_{|D|^{\frac{1}{2}} u_N} a\big)_{L^2_x} \qquad (9.22)\end{aligned}$$

by the paraproduct decomposition (9.4) of $a|D|^{\frac{1}{2}} u_N = T_a|D|^{\frac{1}{2}} u_N + R_{|D|^{\frac{1}{2}} u_N} a$.

Estimate of (9.21). We write

$$\begin{aligned}(9.21) &= \big(\mathrm{i}T_a|D|^{\frac{1}{2}}\langle D\rangle^s u_N, \langle D\rangle^s u_N\big)_{L^2_x} + \big(\mathrm{i}[\langle D\rangle^s, T_a|D|^{\frac{1}{2}}]u_N, \langle D\rangle^s u_N\big)_{L^2_x} \\ &\quad + \big(\langle D\rangle^s u_N, \mathrm{i}T_a|D|^{\frac{1}{2}}\langle D\rangle^s u_N\big)_{L^2_x} + \big(\langle D\rangle^s u_N, \mathrm{i}[\langle D\rangle^s, T_a|D|^{\frac{1}{2}}]u_N\big)_{L^2_x} \\ &= \big(\mathrm{i}[\langle D\rangle^s, T_a|D|^{\frac{1}{2}}]u_N, \langle D\rangle^s u_N\big)_{L^2_x} + \big(\langle D\rangle^s u_N, \mathrm{i}[\langle D\rangle^s, T_a|D|^{\frac{1}{2}}]u_N\big)_{L^2_x} \\ &\quad + \big(\mathrm{i}(T_a|D|^{\frac{1}{2}} - |D|^{\frac{1}{2}}(T_a)^*)\langle D\rangle^s u_N, \langle D\rangle^s u_N\big)_{L^2_x}. \end{aligned} \tag{9.23}$$

Thus (9.23) and Lemma 9.1 imply that the term in (9.21) satisfies

$$\big|\big(\langle D\rangle^s \mathrm{i}T_a|D|^{\frac{1}{2}} u_N, \langle D\rangle^s u_N\big)_{L^2_x} + \big(\langle D\rangle^s u_N, \langle D\rangle^s \mathrm{i}T_a|D|^{\frac{1}{2}} u_N\big)_{L^2_x}\big| \leq_s \|a\|_{H^2_x}\|u_N\|^2_{H^s_x}. \tag{9.24}$$

Estimate of (9.22). Cauchy-Schwartz inequality and (9.7) imply

$$\big|\big(\langle D\rangle^s \mathrm{i}R_{|D|^{\frac{1}{2}} u_N} a, \langle D\rangle^s u_N\big)_{L^2_x} + \big(\langle D\rangle^s u_N, \langle D\rangle^s \mathrm{i}R_{|D|^{\frac{1}{2}} u_N} a\big)_{L^2_x}\big| \leq_s \|\langle D\rangle^s u_N\|_{L^2_x}\|a\|_{H^{s+\frac{1}{2}}_x}\|u_N\|_{H^1_x}. \tag{9.25}$$

By (9.21)-(9.22), (9.24), (9.25), $\|a\|_{H^2_x} \leq 1$, we deduce the differential inequality: $\forall s \geq 1$

$$\partial_t \|u_N\|^2_{H^s_x} \leq_s \|a\|_{H^{s+(1/2)}_x}\|u_N\|_{H^s_x}\|u_N\|_{H^1_x} + \|a\|_{H^2_x}\|u_N\|^2_{H^s_x} \leq_s \|a\|^2_{H^{s+(1/2)}_x}\|u_N\|^2_{H^1_x} + \|u_N\|^2_{H^s_x}. \tag{9.26}$$

For $s = 1$ and since $\|a\|_{H^2_x} \leq 1$, (9.26) reduces to $\partial_t \|u_N\|^2_{H^1_x} \leq C\|u_N\|^2_{H^1_x}$, which implies $\|\Phi_N^t(u_0)\|_{H^1_x} \leq C'\|u_0\|_{H^1_x}$, $\forall t \in [0, 1]$. For $s > 1$, (9.26) reduces to $\partial_t \|u_N\|^2_{H^s_x} \leq C(s)\big(\|a\|^2_{H^{s+(1/2)}_x}\|u_0\|^2_{H^1_x} + \|u_N\|^2_{H^s_x}\big)$ and the estimate (9.17) follows by the Gronwall inequality in differential form.

Since $\Phi_N^t : H^0_x(\mathbb{T}) \to H^0_x(\mathbb{T})$ and $\Phi_N^t : H^1_x(\mathbb{T}) \to H^1_x(\mathbb{T})$ are linear bounded operators, a classical interpolation result implies that $\Phi_N^t : H^s_x(\mathbb{T}) \to H^s_x(\mathbb{T})$ is also bounded $\forall s \in [0, 1]$ and (9.16) holds.

Proof of (9.18), (9.19). Now we pass to the limit $N \to +\infty$. By (9.16) the sequence of functions $u_N(t, \cdot)$ is bounded in $L^\infty_t H^s_x$ and, up to subsequences,

$$u_N \overset{w^*}{\rightharpoonup} u \ \text{ in } \ L^\infty_t H^s_x, \qquad \|u\|_{L^\infty_t H^s_x} \leq \liminf_{N \to +\infty} \|u_N\|_{L^\infty_t H^s_x}. \tag{9.27}$$

Claim: *the sequence $u_N \to u$ in $\mathcal{C}^0_t H^s_x \cap \mathcal{C}^1_t H^{s-\frac{1}{2}}_x$, and $u(t, x)$ solves the equation* (9.1).

We first prove that u_N is a Cauchy sequence in $\mathcal{C}^0_t L^2_x$. Indeed, by (9.2), the difference $h_N := u_{N+1} - u_N$ solves

$$\partial_t h_N = \mathrm{i}\Pi_{N+1}(a|D|^{\frac12} h_N) + \mathrm{i}(\Pi_{N+1} - \Pi_N) a|D|^{\frac12} u_N\,, \quad h_N(0) = (\Pi_{N+1} - \Pi_N) u_0\,,$$

and therefore

$$\begin{aligned}\partial_t \|h_N(t)\|^2_{L^2_x} &= (\partial_t h_N\,,\, h_N)_{L^2_x} + (h_N\,,\, \partial_t h_N)_{L^2_x} \\ &= (\mathrm{i}\Pi_{N+1}(a|D|^{\frac12} h_N), h_N)_{L^2_x} + (h_N, \mathrm{i}\Pi_{N+1}(a|D|^{\frac12} h_N))_{L^2_x} \\ &\quad + (\mathrm{i}(\Pi_{N+1} - \Pi_N) a|D|^{\frac12} u_N, h_N)_{L^2_x} + (h_N, \mathrm{i}(\Pi_{N+1} - \Pi_N) a|D|^{\frac12} u_N)_{L^2_x}\,. \end{aligned} \tag{9.28}$$

Since Π_{N+1} is self-adjoint with respect to the L^2 scalar product

$$\begin{aligned}(\mathrm{i}\Pi_{N+1}(a|D|^{\frac12} h_N), h_N)_{L^2_x} + (h_N, \mathrm{i}\Pi_{N+1}(a|D|^{\frac12} h_N))_{L^2_x} &= (\mathrm{i}a|D|^{\frac12} h_N), h_N)_{L^2_x} + (h_N, \mathrm{i}a|D|^{\frac12} h_N)_{L^2_x} \\ &= (\mathrm{i}[a, |D|^{\frac12}] h_N)\,,\, h_N)_{L^2_x} \le C \|h_N(t)\|^2_{L^2_x}\,. \end{aligned} \tag{9.29}$$

Moreover

$$\begin{aligned}&(\mathrm{i}(\Pi_{N+1} - \Pi_N) a|D|^{\frac12} u_N, h_N)_{L^2_x} + (h_N,\, \mathrm{i}(\Pi_{N+1} - \Pi_N) a|D|^{\frac12} u_N)_{L^2_x} \\ &\le 2\|(\Pi_{N+1} - \Pi_N) a|D|^{\frac12} u_N\|_{L^2_x} \|h_N\|_{L^2_x} \le \|h_N\|^2_{L^2_x} + \|(\Pi_{N+1} - \Pi_N) a|D|^{\frac12} u_N\|^2_{L^2_x} \\ &\lessdot \|h_N\|^2_{L^2_x} + \big(N^{-2} \|a|D|^{\frac12} u_N\|_{H^2_x}\big)^2 \lessdot \|h_N\|^2_{L^2_x} + \big(N^{-2} \|u_0\|_{H^{5/2}_x}\big)^2 \end{aligned} \tag{9.30}$$

using that $\|a\|_{H^2_x} \le 1$. Hence (9.28)-(9.30) imply that $\partial_t \|h_N(t)\|^2_{L^2_x} \lessdot \|h_N(t)\|^2_{L^2_x} + N^{-4} \|u_0\|^2_{H^{5/2}_x}$ and, since $\|h_N(0)\|_{L^2_x} \le N^{-2} \|u_0\|_{H^2_x}$, by Gronwall lemma we deduce that

$$\|u_{N+1} - u_N\|_{\mathcal{C}^0_t L^2_x} = \sup_{t \in [0,1]} \|u_{N+1}(t,\cdot) - u_N(t,\cdot)\|_{L^2_x} \lessdot N^{-2} \|u_0\|_{H^{5/2}_x}\,.$$

The above inequality implies that u_N is a Cauchy sequence in $\mathcal{C}^0_t L^2_x$. Hence $u_N \to \tilde u \in \mathcal{C}^0_t L^2_x$. By (9.27) we have $u = \tilde u \in \mathcal{C}^0_t L^2_x \cap L^\infty_t H^s_x$. Next, for any $\bar s \in [0, s)$ we use the interpolation inequality

$$\|u_N - u\|_{L^\infty_t H^{\bar s}_x} \le \|u_N - u\|^{1-\lambda}_{L^\infty_t L^2_x} \|u_N - u\|^{\lambda}_{L^\infty_t H^{\bar s}_x}\,,$$

and, since u_N is bounded in $L^\infty_t H^s_x$ (see (9.16), (9.17)), $u \in L^\infty_t H^s_x$, and $u_N \to u \in \mathcal{C}^0_t L^2_x$, we deduce that $u_N \to u$ in each $L^\infty_t H^{\bar s}_x$. Since $u_N \in \mathcal{C}^0_t H^{\bar s}_x$ are continuous in t, the limit function $u \in \mathcal{C}^0_t H^{\bar s}_x$ is continuous as well. Moreover we also deduce that

$$\partial_t u_N = \mathrm{i}\Pi_N(a|D|^{\frac12} u_N) \to \mathrm{i}a|D|^{\frac12} u \quad \text{in} \quad \mathcal{C}^0_t H^{\bar s - 1/2}_x\,, \quad \forall \bar s \in [0, s)\,.$$

As a consequence $u \in \mathcal{C}^1_t H^{\bar s - \frac12}_x$ and $\partial_t u = \mathrm{i}a|D|^{\frac12} u$ solves (9.1).

Finally, arguing as in [43], Proposition 5.1.D, it follows that the function $t \to \|u(t)\|^2_{H^s_x}$ is Lipschitz. Furthermore, if $t_n \to t$ then $u(t_n) \rightharpoonup u(t)$ weakly in H^s_x, because $u(t_n) \to u(t)$ in $H^{\bar s}_x$ for any $\bar s \in [0, s)$. As a consequence the sequence $u(t_n) \to u(t)$ strongly in H^s_x. This proves that $u \in \mathcal{C}^0_t H^s_x$ and therefore $\partial_t u = \mathrm{i}a|D|^{\frac12} u \in \mathcal{C}^0_t H^{s-\frac12}_x$.

Uniqueness. If $u_1, u_2 \in \mathcal{C}^0_t H^s_x \cap \mathcal{C}^1_t H^{s-\frac12}_x$, $s \ge 1/2$, are solutions of (9.1), then $h := u_1 - u_2$ solves

$$\partial_t h = \mathrm{i}a|D|^{\frac12} h\,, \qquad h(0) = 0\,.$$

Arguing as in the proof of (9.26) we deduce the energy inequality $\partial_t \|h(t)\|^2_{L^2_x} \le C \|h(t)\|^2_{L^2_x}$. Since $h(0) = 0$, Gronwall lemma implies that $\|h(t)\|^2_{L^2_x} = 0$, for any $t \in [0,1]$, i.e. $h(t) = 0$. Therefore the problem (9.1) has a unique solution $u(t)$ that we denote by $\Phi^t(u_0)$. The estimate (9.18), (9.19) then follows by (9.27), (9.16), (9.17), since $u_N(t) = \Phi^t_N(u_0)$. □

In the next lemma we prove the smooth dependence of the flow with respect to parameters.

Lemma 9.3. *Let $a(z,\cdot) \in \mathcal{C}^\infty(\mathbb{T})$ and p_0-times differentiable, resp. $\mathcal{C}^{p_0}$, with respect to $z \in \mathcal{B}_X$, where $\mathcal{B}_X$ is an open subset of a Banach space X. Then, for any $p \le p_0$, the flow map $\Phi(z,t)$, $t \in [0,1]$, is smooth in z, more precisely, the map*

$$\mathcal{B}_X \ni z \mapsto \Phi(z,t) \in \mathcal{L}(H^s_x, H_x^{s-\frac{p}{2}-\frac12}), \quad \forall s \ge (p/2) + (1/2),$$

is p-times differentiable, resp. $\mathcal{C}^p$. Moreover, for any $z \in \mathcal{B}_X$, the derivative $\partial_z^p \Phi(z,t)$ is a multilinear form from X^p in $\mathcal{L}(H^s_x, H_x^{s-\frac{p}{2}})$.

Proof. We denote for simplicity $\| \ \|_{\mathcal{L}(H^s_x)} := \| \ \|_{\mathcal{L}(H^s_x, H^s_x)}$. We argue by induction on p. We first prove the statement for $p = 0$. Let $s \ge 1/2$. By (6.131), we have that $\Delta_z\Phi(z,t) := \Phi(z+z_1,t) - \Phi(z,t)$ solves

$$\partial_t \Delta_z \Phi(t) = \mathrm{i} a(z+z_1,x)|D|^{\frac12}\Delta_z\Phi(t) + \mathrm{i}\Delta_z a |D|^{\frac12}\Phi(z,t), \quad \Delta_z\Phi(0) = 0,$$

where $\Delta_z a := a(z+z_1,x) - a(z,x)$. By Duhamel principle $\Delta_z\Phi(z,t) = \int_0^t \Phi(z+z_1,t-\tau)\mathrm{i}\Delta_z a|D|^{\frac12}\Phi(z,\tau)d\tau$. Hence

$$\sup_{t\in[0,1]} \|\Delta_z\Phi(z,t)\|_{\mathcal{L}(H^s_x, H_x^{s-\frac12})} \le \sup_{t\in[0,1]}\|\Phi(z+z_1,t)\|_{\mathcal{L}(H_x^{s-\frac12})}\|\Delta_z a\|_{\mathcal{C}_x^{s-\frac12}} \sup_{t\in[0,1]}\|\Phi(z,t)\|_{\mathcal{L}(H^s_x)} \to 0 \quad (9.31)$$

as $z_1 \to 0$, because $\|a(z+z_1) - a(z)\|_{\mathcal{C}_x^{s-\frac12}} \to 0$ by continuity.

Now we assume that for all $0 \le q \le p < p_0$, the flow $z \mapsto \Phi(z,t) \in \mathcal{L}(H^s_x, H_x^{s-\frac{q}{2}-\frac12})$, $s \ge q/2 + 1/2$, is q-times differentiable, with $\partial_z^q\Phi : X^q \to \mathcal{L}(H^s_x, H_x^{s-\frac{q}{2}})$ and we prove that $z \mapsto \Phi(z,t) \in \mathcal{L}(H^s_x, H_x^{s-\frac{p+1}{2}-\frac12})$, $s \ge (p+1)/2 + 1/2$, is $(p+1)$-times differentiable with $\partial_z^{p+1}\Phi(z,t) : X^{p+1} \to \mathcal{L}(H^s_x, H_x^{s-\frac{p+1}{2}})$.

The derivate $\partial_z^p\Phi(z,t)$ solves the equation, for any $z_1,\dots,z_p \in X$,

$$\partial_t\big(\partial_z^p\Phi(z,t)[z_1,\dots,z_p]\big) = \mathrm{i}a(z,x)|D|^{\frac12}\partial_\lambda^p\Phi(z,t)[z_1,\dots,z_p] + F_p(z,t)[z_1,\dots,z_p], \ \partial_z^p\Phi(z,0) = 0 \quad (9.32)$$

where $F_0 := 0$ and, for any $1 \le q \le p+1$,

$$F_q(z,t)[z_1,\dots,z_q] := \sum_{0\le q_1\le q-1, \sigma\in\mathcal{P}_q} \mathrm{i}\partial_z^{q-q_1}a(z)[z_{\sigma(1)},\dots,z_{\sigma(q-q_1)}]|D|^{\frac12}\partial_z^{q_1}\Phi(z,t)[z_{\sigma(q-q_1+1)},\dots,z_{\sigma(q)}] \quad (9.33)$$

denoting by $\mathcal{P}_q$ the set of permutations of the indices $\{1,\dots,q\}$. For $0 \le q \le p$ we have

$$F_{q+1}(z,t) = \partial_z F_q(z,t) + \mathrm{i}\partial_z a(z,x)[\cdot]|D|^{\frac12}\partial_z^q\Phi(z,t). \quad (9.34)$$

The candidate $(p+1)$-derivative of the operator $\Phi(z,t)$ is the multilinear $(p+1)$-form

$$\mathcal{A}_p(z,t)[z_1,\dots,z_{p+1}] := \int_0^t \Phi(z,t-\tau)F_{p+1}(z,\tau)[z_1,\dots,z_{p+1}]\,d\tau \quad (9.35)$$

obtained by differentiating formally the equation (9.32) and using the Duhamel principle. We now estimate $\partial_z^p\Phi(z+z_{p+1},t) - \partial_z^p\Phi(z,t) - \mathcal{A}_p(z,t)[z_{p+1}]$. Note that, since $\mathcal{A}_p(z,t)$ is a multilinear $(p+1)$-form, then $\mathcal{A}_p(z,t)[z_{p+1}]$ is a multilinear p-form. Taking the difference of (9.32) evaluated at $z+z_{p+1}$ and z, and using the Duhamel principle we get that

$$\Delta_z\partial_z^p\Phi(z,t) := \partial_z^p\Phi(z+z_{p+1},t) - \partial_z^p\Phi(z,t) = \int_0^t \Phi(z+z_{p+1},t-\tau)\big(\mathrm{i}\Delta_z a|D|^{\frac12}\partial_z^p\Phi(z,t) + \Delta_z F_p\big)d\tau$$

where $\Delta_z a := a(z+z_{p+1},x) - a(z,x)$ and $\Delta_z F_p := F_p(z+z_{p+1},t) - F_p(z,t)$. Hence, by (9.35) and (9.34) with $q = p$, we get

$$\begin{aligned}
&\Delta_z\partial_z^p\Phi(z,t) - \mathcal{A}_p(z,t)[z_{p+1}] \\
&= \int_0^t \Phi(z+z_{p+1},t-\tau)\mathrm{i}\Delta_z a|D|^{\frac12}\partial_z^p\Phi(z,\tau)d\tau - \int_0^t \Phi(z,t-\tau)\mathrm{i}\partial_z a(z)[z_{p+1}]|D|^{\frac12}\partial_z^p\Phi(z,\tau)d\tau \quad (9.36)\\
&\quad + \int_0^t \Phi(z+z_{p+1},t-\tau)\Delta_z F_p d\tau - \int_0^t \Phi(z,t-\tau)\partial_z F_p(z,\tau)[z_{p+1}]d\tau \quad (9.37)
\end{aligned}$$

Estimate of (9.36). Set $\Delta_z \Phi(t) := \Phi(z + z_{p+1}, t) - \Phi(z,t)$. For all $0 \le \tau < t$, we have

$$
\begin{aligned}
& \Big\| \big\{ \Phi(z+z_{p+1}, t-\tau) \mathrm{i} \Delta_z a - \Phi(z, t-\tau) \mathrm{i} \partial_z a(z)[z_{p+1}] \big\} |D|^{\frac12} \partial_z^p \Phi(z,\tau)[z_1, \ldots, z_p] \Big\|_{\mathcal{L}(H^s_x, H_x^{s - \frac{p+1}{2} - \frac12})} \\
& \le \Big\| \Phi(z+z_{p+1}, t-\tau) \mathrm{i} (\Delta_z a - \partial_z a[z_{p+1}]) |D|^{\frac12} \partial_z^p \Phi(z,\tau)[z_1, \ldots, z_p] \Big\|_{\mathcal{L}(H^s_x, H_x^{s - \frac{p+1}{2} - \frac12})} \\
& + \Big\| \Delta_z \Phi(t-\tau) \mathrm{i} \partial_z a(z)[z_{p+1}] |D|^{\frac12} \partial_z^p \Phi(z,\tau)[z_1, \ldots, z_p] \Big\|_{\mathcal{L}(H^s_x, H_x^{s - \frac{p+1}{2} - \frac12})} \\
& \le \sup_{t \in [0,1]} \| \Phi(z+z_{p+1}, t) \|_{\mathcal{L}(H_x^{s - \frac{p+1}{2} - \frac12})} \| \Delta_z a - \partial_z a[z_{p+1}] \|_{\mathcal{C}_x^{s - \frac{p+1}{2} - \frac12}} \sup_{t \in [0,1]} \| \partial_z^p \Phi(z,\tau)[z_1, \ldots, z_p] \|_{\mathcal{L}(H^s_x, H_x^{s - \frac{p}{2} - \frac12})} \\
& + \sup_{t \in [0,1]} \| \Delta_z \Phi(t) \|_{\mathcal{L}(H_x^{s - \frac{p+1}{2}}, H_x^{s - \frac{p+1}{2} - \frac12})} \| \partial_z a(z)[z_{p+1}] \|_{\mathcal{C}_x^{s - \frac{p+1}{2}}} \sup_{t \in [0,1]} \| \partial_z^p \Phi(z,\tau)[z_1, \ldots, z_p] \|_{\mathcal{L}(H^s_x, H_x^{s - \frac{p}{2}})} \\
& \le_{s,p} \Big(\| \Delta_z a - \partial_z a[z_{p+1}] \|_{\mathcal{C}_x^{s - \frac{p+1}{2}}} + \sup_{t \in [0,1]} \| \Delta_z \Phi(t) \|_{\mathcal{L}(H_x^{s - \frac{p+1}{2}}, H_x^{s - \frac{p+1}{2} - \frac12})} \| z_{p+1} \| \Big) \| z_1 \| \ldots \| z_p \| \qquad (9.38)
\end{aligned}
$$

using the inductive assumption on $\partial_z^p \Phi(z,\tau)$.

Estimate of (9.37). By the expression in (9.33) (with $q = p$), the fact that $z \mapsto a(z)$ is $(p+1)$-times differentiable, the inductive differentiability properties of the flow, the map $z \mapsto F_p(z,t)[z_1, \ldots, z_p] \in \mathcal{L}(H^s_x, H_x^{s - \frac{p}{2} - \frac12})$ is differentiable. Arguing as above, we have, for all $0 \le t \le \tau$,

$$
\begin{aligned}
& \Big\| \Big\{ \Phi(z+z_{p+1}, t-\tau) \Delta_z F_p(z,\tau) - \Phi(z, t-\tau) \partial_z F_p(z,\tau)[z_{p+1}] \Big\} [z_1, \ldots, z_p] \Big\|_{\mathcal{L}(H^s_x, H_x^{s - \frac{p+1}{2} - \frac12})} \\
& \le_{s,p} \sup_{t \in [0,1]} \| \big(\Delta_z F_p(z,\tau) - \partial_z F_p(z,\tau)[z_{p+1}] \big) [z_1, \ldots, z_p] \|_{\mathcal{L}(H^s_x, H_x^{s - \frac{p+1}{2} - \frac12})} \\
& + \sup_{t \in [0,1]} \| \Delta_z \Phi(z,t) \|_{\mathcal{L}(H_x^{s - \frac{p+1}{2}}, H_x^{s - \frac{p+1}{2} - \frac12})} \| \partial_z F_p(z)[z_{p+1}][z_1, \ldots, z_p] \|_{\mathcal{L}(H^s_x, H_x^{s - \frac{p+1}{2}})} . \qquad (9.39)
\end{aligned}
$$

In conclusion, by (9.36), (9.37), (9.38), (9.39), the differentiability of $a(z)$ and (9.31), we deduce that

$$
\sup_{t \in [0,1]} \sup_{\|z_1\|, \ldots, \|z_p\| \le 1} \| \big(\Delta_z \partial_z^p \Phi(z,t) - \mathcal{A}_p(z,t)[z_{p+1}] \big) [z_1, \ldots, z_p] \|_{\mathcal{L}(H^s_x, H_x^{s - \frac{p+1}{2} - \frac12})} \| z_{p+1} \|^{-1} \to 0 ,
$$

for $z_{p+1} \to 0$, namely $\partial_z^p \Phi(z,t)$ is differentiable and $\partial_z^{p+1} \Phi(z,t) = \mathcal{A}_p(z,t)$. Moreover, by (9.35), (9.33) for $q = p+1$, the continuity of $z \mapsto \partial_z^p a(z)$ and the inductive differentiability properties of the flow, we have that $z \mapsto \partial_z^{p+1} \Phi(z,t)$ is continuous and $\partial_z^{p+1} \Phi(z,t)[z_1, \ldots, z_{p+1}] \in \mathcal{L}(H^s_x, H_x^{s - \frac{p+1}{2}})$. $\square$

We now want to prove tame estimates for the flow operator $\Phi^t := \Phi(t) := \Phi(\lambda, \varphi, t)$ acting in the Sobolev spaces H^s of functions $u(\varphi, x)$. Recall that the Sobolev norm $\| \ \|_s$ in (1.19) is equivalent to $\| \ \|_s \simeq \| \ \|_{H^s_\varphi L^2_x} + \| \ \|_{L^2_\varphi H^s_x}$, see (2.2). Note also the continuous embeddings

$$
H^{s+s_0}(\mathbb{T}^{\nu+1}) \hookrightarrow H^{s_0}(\mathbb{T}^\nu, H^s_x) \hookrightarrow L^\infty(\mathbb{T}^\nu, H^s_x) . \qquad (9.40)
$$

Lemma 9.4. *For any $|\beta| \le \beta_0, |k| \le k_0, t \in [0,1], h \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$, the function $\partial_\lambda^k \partial_\varphi^\beta \Phi^t(\varphi) h$ is $\mathcal{C}^\infty(\mathbb{T}^{\nu+1})$.*

Proof. Since $h(\varphi, x) \in \mathcal{C}^\infty(\mathbb{T}^\nu \times \mathbb{T})$ then $\mathbb{T}^\nu \ni \varphi \mapsto h(\varphi, \cdot) \in H^s_x$ is a $\mathcal{C}^\infty$ map for any $s > 0$. By Lemma 9.3, the map $\mathbb{T}^\nu \ni \varphi \mapsto \partial_\lambda^k \partial_\varphi^\beta \Phi^t(\varphi)[h(\varphi)] \in H^s_x$ is $\mathcal{C}^\infty$ and, for any $\alpha \in \mathbb{N}^\nu$, $\partial_\varphi^\alpha \big\{ \partial_\lambda^k \partial_\varphi^\beta \Phi^t(\varphi) h \big\} = \sum_{\alpha_1 + \alpha_2 = \alpha} C_{\alpha_1, \alpha_2} \partial_\lambda^k \partial_\varphi^{\beta + \alpha_1} \Phi^t(\varphi)[\partial_\varphi^{\alpha_2} h]$. By Lemma 9.3 each function $\partial_\lambda^k \partial_\varphi^{\beta + \alpha_1} \Phi^t(\varphi)[\partial_\varphi^{\alpha_2} h] \in \mathcal{C}^\infty_x$. $\square$

Proposition 9.5. *Assume that*

$$
\| a \|_{2s_0 + \frac32} \le 1 , \quad \| a \|_{2s_0+1} \le \delta(s) \qquad (9.41)
$$

for some $\delta(s) > 0$ small. Then the following tame estimates hold:

$$
\sup_{t \in [0,1]} \| \Phi(t) u_0 \|_s \le C(s) \| u_0 \|_s , \qquad \forall s \in [0, s_0 + 1] , \qquad (9.42)
$$

$$
\sup_{t \in [0,1]} \| \Phi(t) u_0 \|_s \le C(s) \big(\| u_0 \|_s + \| a \|_{s + s_0 + \frac12} \| u_0 \|_{s_0} \big) , \quad \forall s \ge s_0 . \qquad (9.43)
$$

Proof. We take $u_0 \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$, so that Φu_0 is $\mathcal{C}^\infty(\mathbb{T}^{\nu+1})$.
PROOF OF (9.42). For $s = 0$, integrating (9.18) in φ, we have

$$\|\Phi(t)u_0\|_0^2 = \|\Phi(t)u_0\|_{L^2_\varphi L^2_x}^2 = \int_{\mathbb{T}^\nu} \|\Phi(\varphi,t)u_0\|_{L^2_x}^2 \, d\varphi \le C \int_{\mathbb{T}^\nu} \|u_0\|_{L^2_x}^2 \, d\varphi = C\|u_0\|_{L^2_\varphi L^2_x}^2 \,. \tag{9.44}$$

Now we suppose that (9.42) holds for $s \in \mathbb{N}$, $s \le s_0$, and we prove it for $s+1$. By (2.2)

$$\|\Phi(t)u_0\|_{s+1} \simeq \|\Phi(t)u_0\|_{L^2_\varphi H^{s+1}_x} + \|\Phi(t)u_0\|_{H^{s+1}_\varphi L^2_x} \,. \tag{9.45}$$

The first term in (9.45) is estimated, using (9.19), (9.40), (9.41), by

$$\begin{aligned} \|\Phi(t)u_0\|_{L^2_\varphi H^{s+1}_x} &\le_s \|u_0\|_{L^2_\varphi H^{s+1}_x} + \|a\|_{L^\infty_\varphi H^{s+\frac{3}{2}}_x} \|u_0\|_{L^2_\varphi H^1_x} \le_s \|u_0\|_{s+1} + \|a\|_{s+s_0+\frac{3}{2}} \|u_0\|_1 \\ &\le_s \|u_0\|_{s+1} \,. \end{aligned} \tag{9.46}$$

The second term in (9.45) is estimated, using (9.44) and (9.42), by

$$\begin{aligned} \|\Phi(t)u_0\|_{H^{s+1}_\varphi L^2_x} &\simeq \|\Phi(t)u_0\|_{L^2_\varphi L^2_x} + \sup_{m=1,\ldots,\nu} \|\partial_{\varphi_m}(\Phi(t)u_0)\|_{H^s_\varphi L^2_x} \\ &\le_s \|u_0\|_{L^2_\varphi L^2_x} + \sup_{m=1,\ldots,\nu} \big(\|\Phi(t)[\partial_{\varphi_m} u_0]\|_s + \|\partial_{\varphi_m}\Phi(t)u_0\|_s \big) \qquad (9.47) \\ &\le_s \|u_0\|_{s+1} + \|\partial_{\varphi_m}\Phi(t)u_0\|_s \,. \qquad (9.48) \end{aligned}$$

For estimating the last term in (9.48) note that, differentiating (6.131), the operator $\partial_{\varphi_m}\Phi(t)$ solves

$$\partial_t(\partial_{\varphi_m}\Phi(t)) = \mathrm{i}a|D|^{\frac{1}{2}}(\partial_{\varphi_m}\Phi(t)) + \mathrm{i}(\partial_{\varphi_m}a)|D|^{\frac{1}{2}}\Phi(t) \,, \quad \partial_{\varphi_m}\Phi(0) = 0 \,,$$

and then, by Duhamel principle (variation of constants method), it has the representation

$$\partial_{\varphi_m}\Phi(t) = \mathrm{i}\int_0^t \Phi(t-\tau)(\partial_{\varphi_m}a)|D|^{\frac{1}{2}}\Phi(\tau)\, d\tau \,. \tag{9.49}$$

By the inductive assumption (9.42) up to $s \le s_0$, and (9.40), we get

$$\begin{aligned} \|\Phi(t-\tau)(\partial_{\varphi_m}a)|D|^{\frac{1}{2}}\Phi(\tau)[u_0]\|_s &\le_s \|(\partial_{\varphi_m}a)|D|^{\frac{1}{2}}\Phi(\tau)[u_0]\|_s \le \|a\|_{\mathcal{C}^{s+1}} \|\Phi(\tau)u_0\|_{s+\frac{1}{2}} \\ &\le_s \|a\|_{2s_0+1} \sup\nolimits_{t\in[0,1]} \|\Phi(t)u_0\|_{s+1} \,. \end{aligned} \tag{9.50}$$

Therefore (9.45)-(9.50) imply

$$\|\Phi(t)u_0\|_{s+1} \le C(s)\big(\|u_0\|_{s+1} + \|a\|_{2s_0+1} \sup\nolimits_{t\in[0,1]} \|\Phi(t)u_0\|_{s+1}\big)$$

and, for $C(s)\|a\|_{2s_0+1} \le 1/2$, we deduce (9.42) for $s+1$. After s_0-steps we prove (9.42) at s_0+1. Then a classical interpolation result implies that $\Phi(t)$ satisfies the estimate (9.42) also for all $s \in (0, s_0+1)$.
PROOF OF (9.43). We argue again by induction on s. For $s \in [s_0, s_0+1]$ the tame estimate (9.43) is trivially implied by (9.42). Then we suppose that (9.43) holds up to $s \ge s_0$ and we prove it at $s+1$.

We estimate $\|\Phi(t)u_0\|_{s+1}$ as in (9.45)-(9.47). Then we estimate the last terms in (9.47) in a tame way. The inductive hyphothesis (9.43) and Lemma 2.1 (with $a_0 = 2s_0 + \frac{1}{2}$, $b_0 = s_0$, $p = s - s_0$, $q = 1$) imply

$$\begin{aligned} \|\Phi(t)[\partial_{\varphi_m} u_0]\|_s &\le_s \|u_0\|_{s+1} + \|a\|_{s+s_0+\frac{1}{2}} \|u_0\|_{s_0+1} \le_s \|u_0\|_{s+1} + \|a\|_{s+s_0+\frac{3}{2}} \|u_0\|_{s_0} + \|a\|_{2s_0+\frac{1}{2}} \|u_0\|_{s+1} \\ &\le_s \|u_0\|_{s+1} + \|a\|_{s+s_0+\frac{3}{2}} \|u_0\|_{s_0} \end{aligned} \tag{9.51}$$

since $\|a\|_{2s_0+\frac{1}{2}} \le 1$. Then we estimate $\|\partial_{\varphi_m}\Phi(t)u_0\|_s$. By the inductive assumption (9.43), the tame estimates for the product of functions, (9.41) and (9.42), we get, for all $t, \tau \in [0,1]$,

$$\begin{aligned} \|\Phi(t-\tau)(\partial_{\varphi_m}a)|D|^{\frac{1}{2}}\Phi(\tau)[u_0]\|_s &\le_s \|(\partial_{\varphi_m}a)|D|^{\frac{1}{2}}\Phi(\tau)[u_0]\|_s + \|a\|_{s+s_0+\frac{1}{2}} \|(\partial_{\varphi_m}a)|D|^{\frac{1}{2}}\Phi(\tau)[u_0]\|_{s_0} \\ &\le_s \|a\|_{s+s_0+\frac{1}{2}} \|u_0\|_{s_0+\frac{1}{2}} + \|a\|_{s_0+1} \|\Phi(\tau)u_0\|_{s+\frac{1}{2}} \,. \end{aligned} \tag{9.52}$$

Then (9.45), (9.46), (9.47), (9.49), (9.51), (9.52) imply

$$\sup_{t\in[0,1]}\|\Phi(t)u_0\|_{s+1} \leq_s \|u_0\|_{s+1} + \|a\|_{s+s_0+\frac{3}{2}}\|u_0\|_{s_0} + \|a\|_{s_0+1}\sup_{\tau\in[0,1]}\|\Phi(\tau)u_0\|_{s+1} \\ + \|a\|_{s+s_0+\frac{1}{2}}\|u_0\|_{s_0+1}\,.$$

Then, using (9.41) and Lemma 2.1 (with $a_0 = 2s_0 + \frac{1}{2}$, $b_0 = s_0$, $p = s - s_0$, $q = 1$), we get

$$\begin{aligned}\sup_{t\in[0,1]}\|\Phi(t)u_0\|_{s+1} &\leq_s \|u_0\|_{s+1} + \|a\|_{s+s_0+\frac{3}{2}}\|u_0\|_{s_0} + \|a\|_{s+s_0+\frac{1}{2}}\|u_0\|_{s_0+1} \\ &\leq_s \|u_0\|_{s+1} + \|a\|_{s+s_0+\frac{3}{2}}\|u_0\|_{s_0}\end{aligned}$$

which is (9.43) for $s + 1$.

We have proved (9.42), (9.43), for $u_0 \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$. The estimates for $u_0 \in H^s$ follow by density. □

We also prove the following tame estimates.

Lemma 9.6. *For all $n \geq 1$, if $\|a\|_{s_0+\frac{n}{2}+2} \leq \delta(s)$ small, then the following tame estimates hold: $\forall s \geq s_0$*

$$\|\langle D\rangle^{-\frac{n}{2}}\Phi(t)\langle D\rangle^{\frac{n}{2}}h\|_s\,, \|\langle D\rangle^{\frac{n}{2}}\Phi(t)\langle D\rangle^{-\frac{n}{2}}h\|_s \leq_s \|h\|_s + \|a\|_{s+s_0+\frac{n}{2}+2}\|h\|_{s_0}\,. \tag{9.53}$$

Proof. Let $\Phi_n(t) := \langle D\rangle^{-\frac{n}{2}}\Phi(t)\langle D\rangle^{\frac{n}{2}}$. We consider $h \in \mathcal{C}^\infty$ so that $\Phi_n(t)h \in \mathcal{C}^\infty$.
We have $\Phi_n(0) = \mathrm{Id}$ and

$$\partial_t\Phi_n(t) = \langle D\rangle^{-\frac{n}{2}}\mathrm{i}a|D|^{\frac{1}{2}}\Phi(t)\langle D\rangle^{\frac{n}{2}} = \mathrm{i}a|D|^{\frac{1}{2}}\Phi_n(t) + \mathrm{i}[\langle D\rangle^{-\frac{n}{2}}, a|D|^{\frac{1}{2}}]\langle D\rangle^{\frac{n}{2}}\Phi_n(t)\,.$$

Therefore by Duhamel principle we get

$$\Phi_n(t) = \Phi(t) + \Psi_n(t)\,, \quad \Psi_n(t) := \int_0^t \Phi(t-\tau)A_n\Phi_n(\tau)\,d\tau \quad \text{where} \quad A_n := \mathrm{i}[\langle D\rangle^{-\frac{n}{2}}, a|D|^{\frac{1}{2}}]\langle D\rangle^{\frac{n}{2}}\,. \tag{9.54}$$

By Lemmata 2.7, 2.8, and (2.40), (2.39), we get the estimate

$$|A_n|_{0,s,0} \leq_s \|a\|_{s+\frac{n}{2}+2}\,. \tag{9.55}$$

Then by (9.54), using (9.42) (for $s = s_0$) and Lemma 2.13, we get

$$\sup_{t\in[0,1]}\|\Phi_n(t)h\|_{s_0} \leq C\|h\|_{s_0} + C\|a\|_{s_0+\frac{n}{2}+2}\sup_{t\in[0,1]}\|\Phi_n(t)h\|_{s_0}\,.$$

For $C\|a\|_{s_0+\frac{n}{2}+2} \leq 1/2$, we deduce $\sup_{t\in[0,1]}\|\Phi_n(t)h\|_{s_0} \leq C\|h\|_{s_0}$. Then (9.43), (9.55) and Lemma 2.13, imply, for all $s > s_0$,

$$\begin{aligned}\|\Psi_n(t)h\|_s &\leq_s \sup_{t\in[0,1]}\big(\|A_n\Phi_n(t)h\|_s + \|a\|_{s+s_0+\frac{1}{2}}\|h\|_{s_0}\big) \\ &\leq_s \|a\|_{s+s_0+\frac{n}{2}+2}\|h\|_{s_0} + \|a\|_{s_0+\frac{n}{2}+2}\|h\|_s + \|a\|_{s_0+\frac{n}{2}+2}\sup_{t\in[0,1]}\|\Psi_n(t)h\|_s\,.\end{aligned} \tag{9.56}$$

Hence, for $\|a\|_{s_0+\frac{n}{2}+2} \leq \delta(s)$ small, we deduce the estimate (9.53) by (9.54), (9.43), (9.56).
If $h \in H^s$, the estimate (9.53) follows by density. □

Now we prove similar tame estimates for $\partial_\lambda^k\partial_\varphi^\beta\Phi$ when the vector field $\mathrm{i}a(\lambda,\varphi,x)|D|^{1/2}$ depends also on λ. The operator $\partial_\lambda^k\partial_\varphi^\beta\Phi$ loses $|D_x|^{\frac{|\beta|+|k|}{2}}$ derivatives which are compensated by applying $\langle D\rangle^{-\frac{|\beta|+|k|}{2}}$.

Proposition 9.7. *Assume that*

$$\|a\|_{2s_0+\beta_0+1} \leq \delta(s)\,, \quad \|a\|^{k_0,\gamma}_{2s_0+\frac{5}{2}+\beta_0+k_0} \leq 1 \tag{9.57}$$

with $\delta(s) > 0$ small enough. Then, for all $|k| \leq k_0$, $|\beta| \leq \beta_0$, the following tame estimates hold:

$$\|\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s \leq_s \gamma^{-|k|}\|h\|_s\,, \qquad \forall s \in [0, s_0+1]\,, \tag{9.58}$$

$$\|\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s \leq_s \gamma^{-|k|}\big(\|h\|_s + \|a\|^{k_0,\gamma}_{s+s_0+|\beta|+|k|+1}\|h\|_{s_0}\big)\,, \quad \forall s \geq s_0\,, \tag{9.59}$$

and

$$\|\langle D\rangle \partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}-1} h\|_s \le_s \gamma^{-|k|}\|h\|_s\,, \qquad \forall s\in[0,s_0+1]\,, \tag{9.60}$$

$$\|\langle D\rangle \partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}-1} h\|_s \le_s \gamma^{-|k|}\big(\|h\|_s + \|a\|^{k_0,\gamma}_{s+s_0+|\beta|+|k|+2}\|h\|_{s_0}\big)\,, \quad \forall s\ge s_0\,. \tag{9.61}$$

We prove Proposition 9.7 by induction. We introduce the following notation

- **Notation** : Given $k_1, k \in \mathbb{N}^{\nu+1}$, we say that $k_1 \prec k$ if each component $k_{1,m} \le k_m$, $\forall m = 1,\dots,\nu+1$, and there exists $\bar m \in \{1,\dots,\nu+1\}$ such that $k_{1,\bar m} \ne k_{\bar m}$. Given $(k_1,\beta_1),(k,\beta)\in\mathbb{N}^{\nu+1}\times\mathbb{N}^\nu$ we say that $(k_1,\beta_1)\prec(k,\beta)$ if each component $k_{1,m}\le k_m$, $\beta_{1,n}\le\beta_n$, $\forall m=1,\dots,\nu+1$, $\forall n=1,\dots,\nu$ and $(k_1,\beta_1)\ne(k,\beta)$.

We first estimate $\|\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{L^2_\varphi H^s_x}$.

Lemma 9.8. *Assume* (9.57)*. Then, for all* $\varphi\in\mathbb{T}^\nu$*,* $|k|\le k_0$*,* $|\beta|\le\beta_0$*,*

$$\|\partial_\lambda^k\partial_\varphi^\beta\Phi(\varphi)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{H^s_x} \le_s \gamma^{-|k|}\|h\|_{H^s_x}\,, \qquad \forall s\in[0,1]\,, \tag{9.62}$$

$$\|\partial_\lambda^k\partial_\varphi^\beta\Phi(\varphi)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{H^s_x} \le_s \gamma^{-|k|}\big(\|h\|_{H^s_x} + \|a\|^{k_0,\gamma}_{s+s_0+|\beta|+\frac{|k|}{2}+\frac12}\|h\|_{H^1_x}\big)\,, \quad \forall s\ge1\,. \tag{9.63}$$

Proof. We take $h\in\mathcal{C}^\infty$, so that $\|\partial_\lambda^k\partial_\varphi^\beta\Phi(\varphi)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h$ is $\mathcal{C}^\infty$.
We argue by induction on (k,β). For $k=\beta=0$ the estimates (9.62)-(9.63) are proved by (9.18)-(9.19). Then supposing that (9.62)-(9.63) hold for all $(k_1,\beta_1)\prec(k,\beta)$, $|k|\le k_0$, $|\beta|\le\beta_0$, we prove them for $\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}$. Differentiating (6.131) and using the Duhamel principle we get

$$\partial_\lambda^k\partial_\varphi^\beta\Phi(t) = \int_0^t \Phi(t-\tau)F_{\beta,k}(\tau)\,d\tau \tag{9.64}$$

where

$$F_{\beta,k}(\tau) := \sum_{k_1+k_2=k,\beta_1+\beta_2=\beta,(k_1,\beta_1)\prec(k,\beta)} C(k_1,k_2,\beta_1,\beta_2)(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\,. \tag{9.65}$$

We now prove (9.63). For all $(k_1,\beta_1)\prec(k,\beta)$, $k_1+k_2=k$, $\beta_1+\beta_2=\beta$, for all $t,\tau\in[0,1]$, using (9.19), tame estimates for the product, (9.57), we deduce

$$\begin{aligned}
&\|\Phi(t-\tau)(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{H^s_x}\\
&\le_s \|(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{H^s_x} + \|a\|_{s+s_0+\frac12}\|(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{H^1_x}\\
&\le_s \gamma^{-|k_2|}\|a\|^{k_0,\gamma}_{s+s_0+|\beta|+1}\|\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{H^{\frac32}_x} + \gamma^{-|k_2|}\|\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{H^{s+\frac12}_x}\,.
\end{aligned} \tag{9.66}$$

Now, since $(k_1,\beta_1)\prec(k,\beta)$,

$$\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} = \partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta_1|+|k_1|}{2}}\langle D\rangle^{-\frac m2}\,, \quad m:=|\beta|-|\beta_1|+|k|-|k_1|\ge1\,,$$

and, applying the inductive estimates (9.63) for $\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta_1|+|k_1|}{2}}$, (9.57), we get

$$(9.66) \le_s \gamma^{-|k|}\big(\|h\|_{H^s_x} + \|a\|^{k_0,\gamma}_{s+s_0+\frac12+|\beta|+\frac{|k|}{2}}\|h\|_{H^1_x}\big)$$

which, by (9.64), (9.65), proves (9.63) for h which is $\mathcal{C}^\infty$. The estimate (9.63) for $h\in H^s$ follows by density. The estimates (9.62) follow in the same way using (9.18). □

Then, integrating in φ we get the following corollary

Lemma 9.9. *Assume* (9.57)*. Then, for all* $\varphi \in \mathbb{T}^\nu$, $|k| \le k_0$, $|\beta| \le \beta_0$, *we have*

$$\|\partial_\lambda^k \partial_\varphi^\beta \Phi(\varphi)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{L^2_\varphi H^s_x} \le_s \gamma^{-|k|} \|h\|_{L^2_\varphi H^s_x}\,, \qquad \forall s \in [0,1]\,, \tag{9.67}$$

$$\|\partial_\lambda^k \partial_\varphi^\beta \Phi(\varphi)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{L^2_\varphi H^s_x} \le_s \gamma^{-|k|} \big(\|h\|_{L^2_\varphi H^s_x} + \|a\|^{k_0,\gamma}_{s+s_0+\frac12+|\beta|+\frac{|k|}{2}} \|h\|_{L^2_\varphi H^1_x}\big)\,, \quad \forall s \ge 1\,, \tag{9.68}$$

and

$$\|\langle D\rangle \partial_\lambda^k \partial_\varphi^\beta \Phi(\varphi)\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1} h\|_{L^2_\varphi H^s_x} \le_s \gamma^{-|k|} \|h\|_{L^2_\varphi H^s_x}\,, \qquad \forall s \in [0,1]\,, \tag{9.69}$$

$$\|\langle D\rangle \partial_\lambda^k \partial_\varphi^\beta \Phi(\varphi)\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1} h\|_{L^2_\varphi H^s_x} \le_s \gamma^{-|k|} \big(\|h\|_{L^2_\varphi H^s_x} + \|a\|^{k_0,\gamma}_{s+s_0+|\beta|+\frac{|k|}{2}+\frac32} \|h\|_{L^2_\varphi H^1_x}\big)\,,\ \forall s \ge 1\,. \tag{9.70}$$

Proof of Proposition 9.7. Let $h \in \mathcal{C}^\infty$. We argue by induction. For $k = 0, \beta = 0$ the estimates (9.58)-(9.59) follow by (9.42)-(9.43). We first argue by induction on k assuming that we have already proved (9.58)-(9.59) for all $k_1 \prec k$, $|\beta| \le \beta_0$. Then we prove the tame estimates (9.58)-(9.59) for the operator $\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}}$, for all $|\beta| \le \beta_0$. To do this we argue by induction on $|\beta|$, assuming (9.58)-(9.59) for all $|\beta| < n$ and we prove them for $|\beta| = n$ (also $n = 0$). To estimate $\|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_s$ we argue by induction on s.

Proof of (9.58) for $|\beta| = n$. For $s = 0$, by (9.67), we have

$$\|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_0 = \|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{L^2_\varphi L^2_x} \le C\gamma^{-|k|} \|h\|_{L^2_\varphi L^2_x} = C\gamma^{-|k|}\|h\|_0\,. \tag{9.71}$$

Now we suppose to have proved (9.58) with $|\beta| = n$, up to the Sobolev index $s < s_0 + 1$ and we prove it for $s + 1 \le s_0 + 1$. We have

$$\|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{s+1} \simeq \|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{L^2_\varphi H^{s+1}_x} + \|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{H^{s+1}_\varphi L^2_x}\,. \tag{9.72}$$

The first term in (9.72) is estimated, using (9.68), $s \le s_0$, (9.57), by

$$\|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{L^2_\varphi H^{s+1}_x} \le_s \gamma^{-|k|} \big(\|h\|_{s+1} + \|a\|^{k_0,\gamma}_{s+1+s_0+\frac12+|\beta|+\frac{|k|}{2}} \|h\|_1\big) \le_s \gamma^{-|k|}\|h\|_{s+1}\,.$$

Now we estimate the second term in (9.72). By the inductive hyphothesis

$$\begin{aligned}
\|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{H^{s+1}_\varphi L^2_x} &\simeq \|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{L^2_\varphi L^2_x} + \\
&\quad + \sup_{\alpha \in \mathbb{N}^\nu, |\alpha|=1} \|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} [\partial_\varphi^\alpha h]\|_{H^s_\varphi L^2_x} + \sup_{\alpha \in \mathbb{N}^\nu, |\alpha|=1} \|\partial_\lambda^k \partial_\varphi^{\beta+\alpha} \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{H^s_\varphi L^2_x} \\
&\overset{(9.71)}{\lessdot} \gamma^{-|k|}\|h\|_0 + \sup_{\alpha \in \mathbb{N}^\nu, |\alpha|=1} \|\partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} [\partial_\varphi^\alpha h]\|_s + \sup_{\alpha \in \mathbb{N}^\nu, |\alpha|=1} \|\partial_\lambda^k \partial_\varphi^{\beta+\alpha} \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_s \qquad (9.73) \\
&\le_s \gamma^{-|k|}\|h\|_{s+1} + \sup_{\alpha \in \mathbb{N}^\nu, |\alpha|=1} \|\partial_\lambda^k \partial_\varphi^{\beta+\alpha} \Phi \langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_s\,. \qquad (9.74)
\end{aligned}$$

Now, differentiating (6.131) and using Duhamel principle, we get

$$\partial_\lambda^k \partial_\varphi^{\beta+\alpha} \Phi(t) = \int_0^t \Phi(t-\tau) F_{\beta,k}(\tau)\, d\tau\,, \quad F_{\beta,k}(\tau) := F^{(1)}_{\beta,k}(\tau) + F^{(2)}_{\beta,k}(\tau) + F^{(3)}_{\beta,k}(\tau)\,, \tag{9.75}$$

where

$$\begin{aligned}
F^{(1)}_{\beta,k}(\tau) &:= \sum_{\beta_1+\beta_2=\beta+\alpha, k_1+k_2=k, k_1 \prec k} C(k_1,k_2,\beta_1,\beta_2) \partial_\lambda^{k_2} \partial_\varphi^{\beta_2} a |D|^{1/2} \partial_\lambda^{k_1} \partial_\varphi^{\beta_1} \Phi(\tau) \\
F^{(2)}_{\beta,k}(\tau) &:= \sum_{\beta_1+\beta_2=\beta+\alpha, |\beta_1| \le n-1} C(\beta_1,\beta_2) \partial_\varphi^{\beta_2} a |D|^{1/2} \partial_\lambda^{k} \partial_\varphi^{\beta_1} \Phi(\tau) \\
F^{(3)}_{\beta,k}(\tau) &:= \sum_{\beta_1+\beta_2=\beta+\alpha, |\beta_1| = n} C(\beta_1,\beta_2) \partial_\varphi^{\beta_2} a |D|^{1/2} \partial_\lambda^{k} \partial_\varphi^{\beta_1} \Phi(\tau)\,. \qquad (9.76)
\end{aligned}$$

Note that if $n=0$ the same formula applies, just without the second line. Therefore

$$\begin{aligned}
\|\partial_\lambda^k\partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s \lessdot & \sup_{\substack{k_1\prec k, k_1+k_2=k,\\ \beta_1+\beta_2=\beta+\alpha}}\sup_{t,\tau\in[0,1]}\|\Phi(t-\tau)(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s \\
&+\sup_{\substack{\beta_1+\beta_2=\beta+\alpha\\ |\beta_1|\le n-1}}\sup_{t,\tau\in[0,1]}\|\Phi(t-\tau)(\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s \\
&+\sup_{\substack{\beta_1+\beta_2=\beta+\alpha\\ |\beta_1|=n}}\sup_{t,\tau\in[0,1]}\|\Phi(t-\tau)(\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s. \qquad (9.77)
\end{aligned}$$

We estimate separately the three terms in the above inequality. By the estimate (9.42) for Φ, the inductive hyphothesis for $k_1+k_2=k$, $k_1\prec k$, $\beta_1+\beta_2=\beta+\alpha$, $t,\tau\in[0,1]$, and using (9.57), we get

$$\begin{aligned}
\|\Phi(t-\tau)(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s &\le_s \|(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s \\
&\le_s \gamma^{-|k_2|}\|a\|^{k_0,\gamma}_{2s_0+|\beta|+1}\|\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{s+1} \\
&\le_s \gamma^{-|k|}\|h\|_{s+1}\,. \qquad (9.78)
\end{aligned}$$

The second term in (9.77) is estimated as in (9.78). Then we consider the last term in (9.77). For $\beta_1+\beta_2=\beta+\alpha$, $|\beta_1|=n$, $s\le s_0$,

$$\|\Phi(t-\tau)(\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s \le_s \|a\|_{2s_0+|\beta|+1}\|\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{s+1}\,. \qquad (9.79)$$

By (9.72)-(9.79) we get

$$\sup_{|\beta|=n}\sup_{t\in[0,1]}\|\partial_\lambda^k\partial_\varphi^{\beta}\Phi(t)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{s+1} \le_s \gamma^{-|k|}\|h\|_{s+1}+\|a\|_{2s_0+|\beta|+1}\sup_{|\beta|=n}\sup_{t\in[0,1]}\|\partial_\lambda^k\partial_\varphi^{\beta}\Phi(t)\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{s+1}$$

which implies (9.58) for $|\beta|=n$ at $s+1$, because $\|a\|_{2s_0+|\beta|+1}\le\delta(s)$ is small enough (see (9.57)).

Proof of (9.59) for $|\beta|=n$. The estimate (9.59) for $s=s_0$ follows by (9.58). Then we assume to have proven (9.59) with $|\beta|=n$, up to the Sobolev index s and we prove it $\|\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{s+1}$. The first term in (9.72) is estimated, using (9.68), by

$$\|\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{L^2_\varphi H^{s+1}_x} \le_s \gamma^{-|k|}\big(\|h\|_{s+1}+\|a\|^{k_0,\gamma}_{s+s_0+1+|\beta|+|k|+1}\|h\|_1\big)\,. \qquad (9.80)$$

Now we estimate the second term in (9.72). We have as in (9.73) that

$$\begin{aligned}
&\|\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_{H^{s+1}_\varphi L^2_x} \\
&\simeq \gamma^{-|k|}\|h\|_0+\sup_{\alpha\in\mathbb{N}^\nu,|\alpha|=1}\|\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}[\partial_\varphi^\alpha h]\|_s+\sup_{\alpha\in\mathbb{N}^\nu,|\alpha|=1}\|\partial_\lambda^k\partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}h\|_s\,. \qquad (9.81)
\end{aligned}$$

The first term in (9.81) is estimated by the inductive hyphothesis (on s)

$$\begin{aligned}
\|\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}}[\partial_\varphi^\alpha h]\|_s &\le_s \gamma^{-|k|}\big(\|h\|_{s+1}+\|a\|^{k_0,\gamma}_{s+s_0+1+|\beta|+|k|}\|h\|_{s_0+1}\big) \\
&\le_s \gamma^{-|k|}\big(\|h\|_{s+1}+\|a\|^{k_0,\gamma}_{s+s_0+1+|\beta|+|k|+1}\|h\|_{s_0}\big) \qquad (9.82)
\end{aligned}$$

using (9.57) and the interpolation inequality (2.10) with $a_0=2s_0+|\beta|+|k|+1, b_0=s_0, p=s-s_0, q=1$, $\epsilon=1$.

Now we estimate the second term in (9.81). By (9.75)-(9.76) one has

$$
\begin{aligned}
\|\partial_\lambda^k \partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_s \lessdot & \sup_{\substack{k_1\prec k, k_1+k_2=k,\\ \beta_1+\beta_2=\beta+\alpha}} \sup_{t,\tau\in[0,1]} \|\Phi(t-\tau)(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2} a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_s \\
& + \sup_{\substack{\beta_1+\beta_2=\beta+\alpha\\ |\beta_1|\le n-1}} \sup_{t,\tau\in[0,1]} \|\Phi(t-\tau)(\partial_\varphi^{\beta_2} a)|D|^{\frac12}\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_s \\
& + \sup_{\substack{\beta_1+\beta_2=\beta+\alpha\\ |\beta_1|= n}} \sup_{t,\tau\in[0,1]} \|\Phi(t-\tau)(\partial_\varphi^{\beta_2} a)|D|^{\frac12}\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_s . \qquad (9.83)
\end{aligned}
$$

Note that if $n = 0$ the same formula applies, just without the second line. We estimate separately the terms in (9.83). By the estimate (9.43) on Φ, (9.58), and the inductive hyphothesis for $k_1 + k_2 = k$, $k_1 \prec k$, $\beta_1 + \beta_2 = \beta + \alpha$, $t, \tau \in [0, 1]$, we get

$$
\begin{aligned}
& \|\Phi(t-\tau)(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2} a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_s \\
& \le_s \gamma^{-|k_2|}\|\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{s+\frac12} + \gamma^{-|k_2|}\|a\|^{k_0,\gamma}_{s+s_0+|\beta|+1}\|\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{s_0+\frac12} \\
& \le_s \gamma^{-|k|}\big(\|h\|_{s+1} + \|a\|^{k_0,\gamma}_{s+s_0+1+|\beta|+|k|+1}\|h\|_{s_0}\big) \qquad (9.84)
\end{aligned}
$$

using (9.57) and since (2.10) with $a_0 = 2s_0 + |\beta| + 1$, $b_0 = s_0$, $p = s - s_0$, $q = 1$, $\epsilon = 1$, implies

$$
\|a\|^{k_0,\gamma}_{s+s_0+|\beta|+1}\|h\|_{s_0+1} \le \|a\|^{k_0,\gamma}_{2s_0+|\beta|+1}\|h\|_{s+1} + \|a\|^{k_0,\gamma}_{s+s_0+|\beta|+2}\|h\|_{s_0} . \qquad (9.85)
$$

The second term in (9.83) is estimated similarly by (9.84). Then we consider the third term in (9.83). For $\beta_1 + \beta_2 = \beta + \alpha$, $|\beta_1| = n$, by (9.43), (9.58)

$$
\begin{aligned}
& \|\Phi(t-\tau)(\partial_\varphi^{\beta_2} a)|D|^{\frac12}\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_s \\
& \le_s \|a\|_{s+s_0+|\beta|+1}\|\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{s_0+1} + \|a\|_{2s_0+|\beta|+1}\|\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{s+1} \\
& \le_s \gamma^{-|k|}\|a\|_{s+s_0+|\beta|+1}\|h\|_{s_0+1} + \|a\|_{2s_0+|\beta|+1}\|\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{s+1} \\
& \overset{(9.85)}{\le_s} \gamma^{-|k|}(\|h\|_{s+1} + \|a\|^{k_0,\gamma}_{s+s_0+|\beta|+2}\|h\|_{s_0}) + \|a\|_{2s_0+|\beta|+1}\|\partial_\lambda^{k}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{s+1} . \qquad (9.86)
\end{aligned}
$$

By (9.80), (9.81), (9.82), (9.83), (9.84), (9.86) we get

$$
\begin{aligned}
\sup_{|\beta|=n}\sup_{t\in[0,1]} \|\partial_\lambda^{k}\partial_\varphi^{\beta}\Phi(t)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{s+1} \le_s & \gamma^{-|k|}\big(\|h\|_{s+1} + \|a\|^{k_0,\gamma}_{s+s_0+|\beta|+|k|+2}\|h\|_{s_0}\big) \\
& + \|a\|_{2s_0+|\beta|+1}\sup_{|\beta|=n}\sup_{t\in[0,1]} \|\partial_\lambda^{k}\partial_\varphi^{\beta}\Phi(t)\langle D\rangle^{-\frac{|\beta|+|k|}{2}} h\|_{s+1}
\end{aligned}
$$

which implies (9.59) at $s + 1$ for $|\beta| = n$, because $\|a\|_{2s_0+|\beta|+1} \le \delta(s)$ is small enough (see (9.57)).

Proof of (9.60)-(9.61). We argue by induction on s. The estimate (9.60) for $s = 0$ is proved by (9.69) for $s = 0$. Now let us suppose to have estimated the operator $\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}$ up to the Sobolev index s and let us prove it for $s + 1$. We have to estimate

$$
\|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1} h\|_{s+1} \simeq \|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1} h\|_{L^2_\varphi H^{s+1}_x} + \|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1} h\|_{H^{s+1}_\varphi L^2_x} .
$$

The first term is estimated by (9.70) as

$$
\|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1} h\|_{L^2_\varphi H^{s+1}_x} \le_s \gamma^{-|k|}\big(\|h\|_{s+1} + \|a\|^{k_0,\gamma}_{s+1+s_0+|\beta|+\frac{|k|}{2}+\frac32}\|h\|_1\big) , \qquad (9.87)
$$

and the second term, using (9.69), as

$$
\begin{aligned}
&\|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}h\|_{H^{s+1}_\varphi L^2_x}\simeq\|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}h\|_{L^2_\varphi L^2_x} \qquad (9.88)\\
&+\sup_{\alpha\in\mathbb{N}^\nu,|\alpha|=1}\|\langle D\rangle\partial_\lambda^k\partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}h\|_{H^s_\varphi L^2_x}+\sup_{\alpha\in\mathbb{N}^\nu,|\alpha|=1}\|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\partial_\varphi^\alpha h\|_s\\
&\lessdot\gamma^{-|k|}\|h\|_0+\sup_{\alpha\in\mathbb{N}^\nu,|\alpha|=1}\|\langle D\rangle\partial_\lambda^k\partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}h\|_s+\sup_{\alpha\in\mathbb{N}^\nu,|\alpha|=1}\|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\partial_\varphi^\alpha h\|_s\,.
\end{aligned}
$$

By the inductive hyphothesis, for all $\alpha\in\mathbb{N}^\nu$, $|\alpha|=1$,

$$
\|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\partial_\varphi^\alpha h\|_s\le_s\gamma^{-|k|}\|h\|_{s+1}\,,\quad\forall s\le s_0 \qquad (9.89)
$$

$$
\begin{aligned}
\|\langle D\rangle\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\partial_\varphi^\alpha h\|_s&\le_s\gamma^{-|k|}\big(\|h\|_{s+1}+\|a\|^{k_0,\gamma}_{s+s_0+|\beta|+|k|+2}\|h\|_{s_0+1}\big)\,,\quad\forall s>s_0\\
&\le_s\gamma^{-|k|}\big(\|h\|_{s+1}+\|a\|^{k_0,\gamma}_{s+1+s_0+|\beta|+1+|k|+1}\|h\|_{s_0}\big)\,, \qquad (9.90)
\end{aligned}
$$

since (2.10) with $a_0=2s_0+|\beta|+|k|+2$, $b_0=s_0$, $p=s-s_0$, $q=1$, $\epsilon=1$, and (9.57) imply

$$
\|a\|^{k_0,\gamma}_{s+s_0+|\beta|+|k|+2}\|h\|_{s_0+1}\le\|h\|_{s+1}+\|a\|^{k_0,\gamma}_{s+1+s_0+|\beta|+|k|+2}\|h\|_{s_0}\,.
$$

Finally

$$
\|\langle D\rangle\partial_\lambda^k\partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}h\|_s\le\|\partial_\lambda^k\partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}h\|_{s+1}=\|\partial_\lambda^k\partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|+1}{2}}[\langle D\rangle^{-\frac12}h]\|_{s+1}
$$

and (9.58)-(9.59) imply

$$
\begin{aligned}
&\|\partial_\lambda^k\partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|+1}{2}}[\langle D\rangle^{-\frac12}h]\|_{s+1}\le_s\gamma^{-|k|}\|h\|_{s+1}\,,\quad\forall s\le s_0\,,\\
&\|\partial_\lambda^k\partial_\varphi^{\beta+\alpha}\Phi\langle D\rangle^{-\frac{|\beta|+|k|+1}{2}}[\langle D\rangle^{-\frac12}h]\|_{s+1}\le_s\gamma^{-|k|}\big(\|h\|_{s+1}+\|a\|^{k_0,\gamma}_{s+1+s_0+|\beta|+1+|k|+1}\|h\|_{s_0}\big)\,,\quad\forall s\ge s_0\,.
\end{aligned}
$$

Collecting all the above estimates we have proved (9.60)-(9.61) with Sobolev index $s+1$.

We have then proved the estimates (9.58)-(9.61) for $h\in\mathcal{C}^\infty$. If $h\in H^s$ they follows by density. The proof of Proposition 9.7 is completed. □

Proposition 9.10. *For $\beta_0\in\mathbb{N}$ assume that*

$$
\|a\|_{2s_0+\frac{\beta_0+k_0}{2}+3}\le\delta(s)\,,\quad\|a\|^{k_0,\gamma}_{2s_0+3+\frac32\beta_0+\frac{k_0}{2}}\le1\,, \qquad (9.91)
$$

for $\delta(s)>0$ small. Then, for all $\beta\in\mathbb{N}^\nu,k\in\mathbb{N}^{\nu+1}$ with $|\beta|\le\beta_0$, $|k|\le k_0$, $s\ge s_0$, we have

$$
\sup\nolimits_{t\in[0,1]}\|\langle D\rangle^{-\frac{|\beta|+|k|}{2}}\partial_\lambda^k\partial_\varphi^\beta\Phi(\varphi,t)h\|_s\le_s\gamma^{-|k|}\big(\|h\|_s+\|a\|^{k_0,\gamma}_{s+s_0+2+\frac32|\beta|+\frac12|k|}\|h\|_{s_0}\big)\,. \qquad (9.92)
$$

$$
\sup\nolimits_{t\in[0,1]}\|\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\partial_\lambda^k\partial_\varphi^\beta\Phi(\varphi,t)\langle D\rangle h\|_s\le_s\gamma^{-|k|}\big(\|h\|_s+\|a\|^{k_0,\gamma}_{s+s_0+3+\frac32|\beta|+\frac12|k|}\|h\|_{s_0}\big)\,. \qquad (9.93)
$$

Proof. We prove only (9.93). The proof of (9.92) is the same (easier). We take $h\in\mathcal{C}^\infty$ and we argue by induction on (k,β). For $k=0\,,\beta=0$ the estimate (9.93) is proved by (9.53) with $n=2$. Then supposing that (9.93) holds for all $(k_1,\beta_1)\prec(k,\beta)$, $|k|\le k_0$, $|\beta|\le\beta_0$, we prove it for $\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\partial_\lambda^k\partial_\varphi^\beta\Phi\langle D\rangle$ for which we use the integral representation (9.64)-(9.65). For all $\beta_1+\beta_2=\beta$, $k_1+k_2=k$, $(k_1,\beta_1)\prec(k,\beta)$, $t,\tau\in[0,1]$, one has

$$
\begin{aligned}
\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\Phi(t-\tau)(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle=\ &\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\Phi(t-\tau)\langle D\rangle^{\frac{|\beta|+|k|}{2}+1} \qquad (9.94)\\
&\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)\langle D\rangle^{\frac{|\beta|+|k|}{2}+1}\\
&|D|^{\frac12}\langle D\rangle^{-\frac{m}{2}}\langle D\rangle^{-\frac{|\beta_1|+|k_1|}{2}-1}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle
\end{aligned}
$$

where $m := |\beta| - |\beta_1| + |k| - |k_1| \geq 1$. These three terms satisfy tame estimates. By (9.53) (which can be applied because of (9.91)) we have

$$\|\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\Phi(t-\tau)\langle D\rangle^{\frac{|\beta|+|k|}{2}+1}h\|_s \leq_s \|h\|_s + \|a\|_{s+s_0+2+\frac{|\beta|+|k|}{2}}\|h\|_{s_0}\,. \tag{9.95}$$

Lemma 2.6, 2.7, and (2.39), (2.40), imply

$$|\!\langle D\rangle^{-\frac{|\beta|+|k|}{2}-1}\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a\langle D\rangle^{\frac{|\beta|+|k|}{2}+1}|\!|_{0,s,0} \leq_s \|\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a\|_{s+\frac{|\beta|+|k|}{2}} \leq_s \gamma^{-|k_2|}\|a\|^{k_0,\gamma}_{s+\frac{3}{2}|\beta|+\frac{|k|}{2}}\,. \tag{9.96}$$

Since $(k_1,\beta_1) \prec (k,\beta)$, using the inductive estimates (9.92) for $\langle D\rangle^{-\frac{|\beta_1|+|k_1|}{2}-1}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle$, we get

$$\begin{aligned}\||D|^{\frac{1}{2}}\langle D\rangle^{-\frac{m}{2}}\langle D\rangle^{-\frac{|\beta_1|+|k_1|}{2}-1}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle h\|_s &\leq_s \|\langle D\rangle^{-\frac{|\beta_1|+|k_1|}{2}-1}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle h\|_s \\ &\leq_s \gamma^{-|k_1|}\big(\|h\|_s + \|a\|^{k_0,\gamma}_{s+s_0+2+\frac{3}{2}|\beta|+\frac{|k|}{2}}\|h\|_{s_0}\big)\,.\end{aligned} \tag{9.97}$$

In conclusion, (9.94)-(9.97) imply (9.93). If $h \in H^s$ the estimate (9.93) follows by density. □

As a corollary we get

Proposition 9.11. *Assume* (9.57). *Then the flow $\Phi(t,\lambda)$ of* (9.1) *is $\mathcal{D}^{k_0}$-$\frac{k_0}{2}$-tame (Definition 2.8), more precisely, for all $k \in \mathbb{N}^{\nu+1}$, $|k| \leq k_0$, $s \geq s_0$,*

$$\|\partial_\lambda^k\Phi(\varphi,t)h\|_s \leq_s \gamma^{-|k|}\big(\|h\|_{s+\frac{|k|}{2}} + \|a\|^{k_0,\gamma}_{s+s_0+|k|+1}\|h\|_{s_0+\frac{|k|}{2}}\big)\,, \tag{9.98}$$

$$\sup_{t\in[0,1]}\|\partial_\lambda^k(\Phi(t)-\mathrm{Id})h\|_s \leq_s \gamma^{-|k|}\big(\|a\|^{k_0,\gamma}_{s_0}\|h\|_{s+\frac{|k|+1}{2}} + \|a\|^{k_0,\gamma}_{s+s_0+k_0+\frac{3}{2}}\|h\|_{s_0+\frac{|k|+1}{2}}\big)\,. \tag{9.99}$$

Proof. By (9.59) (with $\beta = 0$) we have

$$\begin{aligned}\|\partial_\lambda^k\Phi(\varphi,t)h\|_s = \|\partial_\lambda^k\Phi(\varphi,t)\langle D\rangle^{-\frac{|k|}{2}}\langle D\rangle^{\frac{|k|}{2}}h\|_s &\leq_s \gamma^{-|k|}\big(\|\langle D\rangle^{\frac{|k|}{2}}h\|_s + \|a\|^{k_0,\gamma}_{s+s_0+|k|+1}\|\langle D\rangle^{\frac{|k|}{2}}h\|_{s_0}\big) \\ &\leq_s \gamma^{-|k|}\big(\|h\|_{s+\frac{|k|}{2}} + \|a\|^{k_0,\gamma}_{s+s_0+|k|+1}\|h\|_{s_0+\frac{|k|}{2}}\big)\end{aligned}$$

which proves (9.98).
Proof of (9.99). By (9.1), i.e. (6.131), we write $\Phi(t) - \mathrm{Id} = \int_0^t \mathrm{i}a|D|^{\frac{1}{2}}\Phi(\tau)\,d\tau$. Then (9.99) for $k = 0$ follows by (2.72) and (9.43). For $|k| > 0$, (9.99) follows by interpolation and using (9.98). □

Finally we consider also the dependence of the flow Φ with respect to the torus $i := i(\varphi) := (\varphi,0,0) + \mathfrak{I}(\varphi)$ (recall the notation (4.19)). Assuming that there exists $\sigma > 0$ such that for any $s \geq 0$, the map

$$\mathfrak{I}(\lambda) \in \mathcal{Y}^{s+\sigma} \mapsto a(\lambda, i(\lambda)) \in H^s\,, \quad \mathcal{Y}^s := H^s(\mathbb{T}^\nu,\mathbb{R}^\nu) \times H^s(\mathbb{T}^\nu,\mathbb{R}^\nu) \times \big(H^s(\mathbb{T}^{\nu+1},\mathbb{R}^2) \cap H^\perp_{\mathbb{S}^+}\big)$$

is differentiable, then, by Lemma 9.3, the flow $\Phi(t)$ is differentiable with respect to i. Note that in the lemma below we do not estimate the derivatives of $\partial_i\Phi(t)$ with respect to λ since it is not required, see remark 7.4. We state an analogous version of Lemma 9.4 (the proof is similar) which takes into account the dependence with respect to the torus i.

Lemma 9.12. *For any $|\beta| \leq \beta_0$, h, i, $\widehat{\imath}$ which are $\mathcal{C}^\infty(\mathbb{T}^{\nu+1})$, the function $\partial_\varphi^\beta\partial_i\Phi^t(i)[\widehat{\imath}]h \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$.*

Proposition 9.13. *Let $s_1 > s_0$ and assume the condition*

$$\|a\|_{2s_0+\frac{\beta_0+1}{2}+3} \leq \delta(s_1)\,, \qquad \|a\|_{s_1+s_0+3+\frac{3}{2}\beta_0} \leq 1 \tag{9.100}$$

for $\delta(s_1) > 0$ small enough. Then, for all $\beta \in \mathbb{N}^\nu$ with $|\beta| \leq \beta_0$, for all $s \in [s_0,s_1]$

$$\|\langle D\rangle^{-\frac{|\beta|+1}{2}}\partial_\varphi^\beta\big(\partial_i\Phi(t)[\hat{\imath}]\big)h\|_s \leq_s \|\partial_i a[\hat{\imath}]\|_{s+\frac{3}{2}|\beta|+\frac{1}{2}}\|h\|_s \tag{9.101}$$

$$\|\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\partial_\varphi^\beta\big(\partial_i\Phi(t)[\hat{\imath}]\big)\langle D\rangle h\|_s \leq_s \|\partial_i a[\hat{\imath}]\|_{s+\frac{3}{2}|\beta|+\frac{3}{2}}\|h\|_s\,. \tag{9.102}$$

Proof. We prove (9.102). The proof of (9.101) is similar. We take $h, \widehat{\imath}$ in $\mathcal{C}^\infty$ with respect to φ and x, so that $\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\partial_\varphi^\beta\big(\partial_i\Phi(t)[\hat{\imath}]\big)\langle D\rangle h$ is $\mathcal{C}^\infty$. Differentiating (6.131) and using Duhamel principle we get

$$\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}] = \int_0^t \Phi(t-\tau)F_\beta(\tau)\,d\tau\,, \quad F_\beta := F_\beta^{(1)} + F_\beta^{(2)} \tag{9.103}$$

where

$$F_\beta^{(1)}(\tau) := \sum_{\beta_1+\beta_2=\beta, |\beta_1|<|\beta|} C(\beta_1,\beta_2)(\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau) \tag{9.104}$$

$$F_\beta^{(2)}(\tau) := \sum_{\beta_1+\beta_2=\beta} C(\beta_1,\beta_2)(\partial_\varphi^{\beta_2}\partial_i a[\widehat{\imath}])|D|^{\frac12}\partial_\varphi^{\beta_1}\Phi(\tau)\,. \tag{9.105}$$

We argue by induction on β. The proof of (9.102) for $\beta = 0$ follows as a particular case of the estimate below for the term in (9.105).

Estimate of (9.104). For any $\beta_1+\beta_2=\beta$, $|\beta_1|<|\beta|$ we have

$$\begin{aligned}
&\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\Phi(t-\tau)(\partial_\varphi^{\beta_2}a)|D|^{\frac12}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau)\langle D\rangle \\
&= \big(\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\Phi(t-\tau)\langle D\rangle^{\frac{|\beta|+1}{2}+1}\big)\big(\langle D\rangle^{-\frac{|\beta|+1}{2}-1}(\partial_\varphi^{\beta_2}a)\langle D\rangle^{\frac{|\beta|+1}{2}+1}\big) \\
&\quad |D|^{\frac12}\langle D\rangle^{-\frac12}\langle D\rangle^{-\frac{|\beta|}{2}-1}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau)\langle D\rangle\,.
\end{aligned} \tag{9.106}$$

By (9.53), $s_0 \le s \le s_1$, (9.100) one has

$$\|\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\Phi(t-\tau)\langle D\rangle^{\frac{|\beta|+1}{2}+1}h\|_s \le_s \|h\|_s + \|a\|_{s+s_0+2+\frac{|\beta|+1}{2}+1}\|h\|_{s_0} \le_s \|h\|_s\,. \tag{9.107}$$

Lemma 2.6, 2.7, and (2.39), (2.40), imply

$$|\!|\!|\langle D\rangle^{-\frac{|\beta|+1}{2}-1}(\partial_\varphi^{\beta_2}a)\langle D\rangle^{\frac{|\beta|+1}{2}+1}|\!|\!|_{0,s,0} \le_s \|\partial_\varphi^{\beta_2}a\|_{s+\frac{|\beta|+1}{2}+1} \le_s \|a\|_{s+\frac32|\beta|+\frac32} \overset{s\le s_1\,,\,(9.100)}{\le_s} 1\,. \tag{9.108}$$

Since $|\beta_1|<|\beta|$ the inductive hyphothesis implies

$$\begin{aligned}
\||D|^{\frac12}\langle D\rangle^{-\frac12}\langle D\rangle^{-\frac{|\beta|}{2}-1}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau)\langle D\rangle h\|_s &\le_s \|\langle D\rangle^{-\frac{|\beta_1|+1}{2}-1}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau)\langle D\rangle h\|_s \\
&\le_s \|\partial_i a[\widehat{\imath}]\|_{s+\frac32|\beta|+\frac32}\|h\|_s\,.
\end{aligned} \tag{9.109}$$

Then (9.104), (9.106), (9.107), (9.108), (9.109) imply

$$\|\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\Phi(t-\tau)F_\beta^{(1)}(\tau)\langle D\rangle h\|_s \le_s \|\partial_i a[\widehat{\imath}]\|_{s+\frac32|\beta|+\frac32}\|h\|_s\,. \tag{9.110}$$

Estimate of (9.105). For any $\beta_1+\beta_2=\beta$, $t,\tau\in[0,1]$, we have

$$\begin{aligned}
&\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\Phi(t-\tau)(\partial_\varphi^{\beta_2}\partial_i a[\widehat{\imath}])|D|^{\frac12}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle \\
&= \big(\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\Phi(t-\tau)\langle D\rangle^{\frac{|\beta|+1}{2}+1}\big)\big(\langle D\rangle^{-\frac{|\beta|+1}{2}-1}(\partial_\varphi^{\beta_2}\partial_i a[\widehat{\imath}])\langle D\rangle^{\frac{|\beta|+1}{2}+1}\big) \\
&\quad |D|^{\frac12}\langle D\rangle^{-\frac12}\langle D\rangle^{-\frac{|\beta|}{2}-1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle\,.
\end{aligned} \tag{9.111}$$

Lemma 2.6, 2.7, and (2.39), (2.40), imply (as for (9.108))

$$|\!|\!|\langle D\rangle^{-\frac{|\beta|+1}{2}-1}(\partial_\varphi^{\beta_2}\partial_i a[\widehat{\imath}])\langle D\rangle^{\frac{|\beta|+1}{2}+1}|\!|\!|_{0,s,0} \le_s \|\partial_i a[\widehat{\imath}]\|_{s+\frac32|\beta|+\frac32}\,. \tag{9.112}$$

By (9.93), $s_0\le s\le s_1$, and (9.100) we get

$$\||D|^{\frac12}\langle D\rangle^{-\frac12}\langle D\rangle^{-\frac{|\beta|}{2}-1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle h\|_s \le_s \|h\|_s\,. \tag{9.113}$$

Finally (9.105), (9.111), (9.107), (9.112), (9.113) imply

$$\|\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\Phi(t-\tau)F^{(2)}_\beta(\tau)\langle D\rangle h\|_s \leq_s \|\partial_i a[\widehat{\imath}]\|_{s+\frac{3}{2}|\beta|+\frac{3}{2}}\|h\|_s\,. \tag{9.114}$$

In conclusion the estimate (9.102) follows by (9.110), (9.114). If $h\in H^s$, $\widehat{\imath}\in\mathcal{Y}^{s+\frac{3}{2}|\beta|+\frac{3}{2}+\sigma}$, then (9.102) follows by density. □

Proposition 9.14. *Let $s_1>s_0$ and assume*

$$\|a\|_{s_1+s_0+\frac{5}{2}+\beta_0}\leq 1\,,\quad \|a\|_{s_1+s_0+\beta_0+1}\leq\delta(s_1)\,, \tag{9.115}$$

for some $\delta(s_1)>0$ small. Then for all $|\beta|\leq\beta_0$,

$$\|\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_s \leq_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{1}{2}+|\beta|}\|h\|_s\,,\qquad \forall s\in[0,s_1]\,, \tag{9.116}$$

$$\|\langle D\rangle\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_s \leq_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{3}{2}+|\beta|}\|h\|_s\,,\qquad \forall s\in[0,s_1-1]\,. \tag{9.117}$$

We first provide the estimate in $\|\cdot\|_{L^2_\varphi H^s_x}$ for all $s\in[0,s_1]$.

Lemma 9.15. *Assume* (9.115)*. Then for all $\varphi\in\mathbb{T}^\nu$, the following estimate holds*

$$\|\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{H^s_x} \leq_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{1}{2}+|\beta|}\|h\|_{H^s_x}\,,\quad \forall s\in[0,s_1]\,. \tag{9.118}$$

Proof. Let us suppose that $\widehat{\imath}$ and h are $\mathcal{C}^\infty$. We argue by induction on β, supposing that we have already proved (9.118) for $|\beta_1|<|\beta|$. We use the integral representation of $\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]$ in (9.103). For all $\beta_1+\beta_2=\beta$, $|\beta_1|<|\beta|$, $t,\tau\in[0,1]$, by (9.18), (9.19), (9.115), and the inductive hyphothesis,

$$\begin{aligned}&\|\Phi(t-\tau)(\partial_\varphi^{\beta_2}a)|D|^{\frac{1}{2}}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{H^s_x}\\&\leq_s \|a\|_{\mathcal{C}^{s+|\beta|}}\|\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{H_x^{s+\frac{1}{2}}} \leq_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{1}{2}+|\beta|+1}\|h\|_{H^s_x}\,.\end{aligned} \tag{9.119}$$

Similarly, for all $\beta_1+\beta_2=\beta$, by (9.18), (9.19), (9.115)

$$\begin{aligned}&\|\Phi(t-\tau)(\partial_\varphi^{\beta_2}\partial_i a[\widehat{\imath}])|D|^{\frac{1}{2}}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{H^s_x}\\&\leq_s \|\partial_i a[\widehat{\imath}]\|_{\mathcal{C}^{s+|\beta|}}\|\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{H_x^{s+\frac{1}{2}}} \overset{(9.62),(9.63),(9.115)}{\leq_s} \|\partial_i a[\widehat{\imath}]\|_{s+s_0+|\beta|}\|h\|_{H^s_x}\,.\end{aligned} \tag{9.120}$$

By (9.103), (9.119), (9.120) we deduce (9.118). If $h\in H^s_x$ and $\widehat{\imath}\in\mathcal{Y}^{s+s_0+\frac{1}{2}+|\beta|+\sigma}$ it follows by density. □

Then, integrating in φ, we get the following corollary

Lemma 9.16. *Let $s_1>s_0$ and assume* (9.115)*. Then for all $|\beta|\leq\beta_0$*

$$\|\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{L^2_\varphi H^s_x} \leq_s \|\partial_i a[\widehat{\imath}]\|^{k_0,\gamma}_{s+s_0+\frac{1}{2}+|\beta|}\|h\|_{L^2_\varphi H^s_x}\,,\qquad \forall s\in[0,s_1]\,, \tag{9.121}$$

$$\|\langle D\rangle\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_{L^2_\varphi H^s_x} \leq_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{3}{2}+|\beta|}\|h\|_{L^2_\varphi H^s_x}\,,\forall s\in[0,s_1-1]\,. \tag{9.122}$$

Proof of Proposition 9.14. Let h and $\widehat{\imath}$ be $\mathcal{C}^\infty$ with respect to the variables φ and x.

Proof of (9.116). We argue by induction $|\beta|$. For $\beta=0$ the proof of (9.116) is a particular case of the estimate of (9.126), (9.129) (with $k=0,\beta+\alpha=0$) in (9.133). Assume that we have proved (9.116) for $\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}$ for all $|\beta|<n$, and let us prove it for $|\beta|=n$. Then we estimate $\|\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_s$ for all $|\beta|=n$, for all $s\in[0,s_1]$. For $s=0$ one has

$$\|\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_0 = \|\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{L^2_\varphi L^2_x} \overset{(9.121)}{\lessdot} \|\partial_i a[\widehat{\imath}]\|_{s_0+\frac{1}{2}+|\beta|}\|h\|_0\,.$$

Then, assume that (9.116) holds up to the Sobolev index $s < s_1$ and we prove it for $s+1 \leq s_1$. We have

$$\|\partial_\varphi^\beta \partial_i \Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}} h\|_{s+1} \simeq \|\partial_\varphi^\beta \partial_i \Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}} h\|_{L^2_\varphi H^{s+1}_x} + \|\partial_\varphi^\beta \partial_i \Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}} h\|_{H^{s+1}_\varphi L^2_x}.$$

By (9.121) we have

$$\|\partial_\varphi^\beta \partial_i \Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}} h\|_{L^2_\varphi H^{s+1}_x} \leq_s \|\partial_i a[\widehat{\imath}]\|_{s+1+s_0+\frac{1}{2}+|\beta|}\|h\|_{s+1}. \tag{9.123}$$

Then

$$\begin{aligned}\|\partial_\varphi^\beta \partial_i \Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}} h\|_{H^{s+1}_\varphi L^2_x} &\simeq \|\partial_\varphi^\beta \partial_i \Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}} h\|_0 \\ &+ \sup_{\alpha\in\mathbb{N}^\nu, |\alpha|=1}\|\partial_\varphi^\beta \partial_i \Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}} \partial_\varphi^\alpha h\|_{H^s_\varphi L^2_x} + \sup_{\alpha\in\mathbb{N}^\nu, |\alpha|=1}\|\partial_\varphi^{\beta+\alpha} \partial_i \Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}} h\|_{H^s_\varphi L^2_x}.\end{aligned} \tag{9.124}$$

The inductive hyphothesis implies

$$\|\partial_\varphi^\beta \partial_i \Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}} \partial_\varphi^\alpha h\|_s \leq_s \|\partial_i a[\widehat{\imath}]\|^{k_0,\gamma}_{s+s_0+\frac{1}{2}+|\beta|}\|h\|_{s+1}. \tag{9.125}$$

We estimate the last term in (9.124). Differentiating (6.131) and using the Duhamel principle we get

$$\partial_\varphi^{\beta+\alpha}\partial_i\Phi[\widehat{\imath}] = \int_0^t \Phi(t-\tau)F_\beta(\tau)\,d\tau\,, \quad F_\beta := F_\beta^{(1)} + F_\beta^{(2)} + F_\beta^{(3)} + F_\beta^{(4)}\,, \tag{9.126}$$

with

$$F_\beta^{(1)}(\tau) := \sum_{\beta_1+\beta_2=\beta+\alpha, |\beta_1|=|\beta|} C(\beta_1,\beta_2)\partial_\varphi^{\beta_2} a|D|^{\frac{1}{2}}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau) \tag{9.127}$$

$$F_\beta^{(2)}(\tau) := \sum_{\beta_1+\beta_2=\beta+\alpha, |\beta_1|<|\beta|} C(\beta_1,\beta_2)\partial_\varphi^{\beta_2} a|D|^{\frac{1}{2}}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau) \tag{9.128}$$

$$F_\beta^{(3)}(\tau) := \sum_{\beta_1+\beta_2=\beta+\alpha} C(\beta_1,\beta_2)(\partial_\varphi^{\beta_2}\partial_i a[\widehat{\imath}])|D|^{\frac{1}{2}}\partial_\varphi^{\beta_1}\Phi(\tau)\,. \tag{9.129}$$

We estimate separately the terms $\Phi(t-\tau)F_\beta^{(m)}(\tau)$, $m = 1,2,3$. We use that by (9.42), (9.43), (9.115)

$$\sup_{t\in[0,1]}\|\Phi(t)h\|_s \leq_s \|h\|_s \qquad \forall s\in[0,s_1]\,. \tag{9.130}$$

For all $t,\tau\in[0,1]$, $\beta_1+\beta_2 = \beta+\alpha$, $|\beta_1| = |\beta|$, one has by (9.130)

$$\|\Phi(t-\tau)\partial_\varphi^{\beta_2} a|D|^{\frac{1}{2}}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau)\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_s \leq_s \|a\|_{s+s_0+|\beta|+1}\|\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau)\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{s+1}. \tag{9.131}$$

For all $t,\tau\in[0,1]$, $\beta_1+\beta_2 = \beta+\alpha$, $|\beta_1| < |\beta|$, by (9.130), the inductive hyphothesis, and (9.115) we get

$$\begin{aligned}&\|\Phi(t-\tau)\partial_\varphi^{\beta_2} a|D|^{\frac{1}{2}}\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau)\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_s \\ &\leq_s \|a\|_{s+s_0+|\beta|+1}\|\partial_\varphi^{\beta_1}\partial_i\Phi[\widehat{\imath}](\tau)\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{s+1} \leq_s \|\partial_i a[\widehat{\imath}]\|_{s+1+s_0+\frac{1}{2}+|\beta|-1}\|h\|_{s+1}.\end{aligned} \tag{9.132}$$

For all $t,\tau\in[0,1]$, $\beta_1+\beta_2 = \beta+\alpha$, we have, by (9.130),

$$\begin{aligned}&\|\Phi(t-\tau)(\partial_\varphi^{\beta_2}\partial_i a[\widehat{\imath}])|D|^{\frac{1}{2}}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_s \\ &\leq_s \|\partial_i a[\widehat{\imath}]\|_{\mathcal{C}^{s+|\beta|+1}}\|\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{s+1} \leq_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+|\beta|+1}\|h\|_{s+1}.\end{aligned} \tag{9.133}$$

using (9.58), (9.59), (9.115). Collecting (9.123)-(9.133) we get

$$\begin{aligned}\sup_{|\beta|=n}\sup_{t\in[0,1]}\|\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{s+1} &\leq_s \|\partial_i a[\widehat{\imath}]\|_{s+1+s_0+\frac{1}{2}+|\beta|}\|h\|_{s+1} \\ &+ \|a\|_{s+1+s_0+|\beta|}\sup_{|\beta|=n}\sup_{t\in[0,1]}\|\partial_\varphi^\beta\partial_i\Phi[\widehat{\imath}]\langle D\rangle^{-\frac{|\beta|+1}{2}}h\|_{s+1}\end{aligned}$$

which, by (9.115), implies (9.116) with Sobolev index $s+1$. $\square$

Proof of (9.117). We argue by induction on s. For $s=0$ it follows by (9.122). Then assuming that (9.117) holds up to the Sobolev index $s<s_1-1$ and we prove it for $s+1$. We have

$$\|\langle D\rangle\partial_\varphi^\beta(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_{s+1}\simeq\|\langle D\rangle\partial_\varphi^\beta(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_{L^2_\varphi H^{s+1}_x} + \|\langle D\rangle\partial_\varphi^\beta(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_{H^{s+1}_\varphi L^2_x}\,. \tag{9.134}$$

By (9.122) we have

$$\|\langle D\rangle\partial_\varphi^\beta(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_{L^2_\varphi H^{s+1}_x}\leq_s\|\partial_i a[\widehat{\imath}]\|_{s+1+s_0+\frac{3}{2}+|\beta|}\|h\|_{s+1}\,. \tag{9.135}$$

We estimate the second term in (9.134). By the inductive hyphothesis and (9.122) one has

$$\begin{aligned}\|\langle D\rangle\partial_\varphi^\beta(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_{H^{s+1}_\varphi L^2_x}&\simeq\|\langle D\rangle\partial_\varphi^\beta(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_{L^2_\varphi L^2_x}\\&+\sup_{\alpha\in\mathbb{N}^\nu,|\alpha|=1}\|\langle D\rangle\partial_\varphi^\beta(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}\partial_\varphi^\alpha h\|_{H^s_\varphi L^2_x}\\&+\sup_{\alpha\in\mathbb{N}^\nu,|\alpha|=1}\|\langle D\rangle\partial_\varphi^{\beta+\alpha}(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_{H^s_\varphi L^2_x}\\&\leq_s\|h\|_{s+1}+\sup_{\alpha\in\mathbb{N}^\nu,|\alpha|=1}\|\langle D\rangle\partial_\varphi^{\beta+\alpha}(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_s\,.\end{aligned} \tag{9.136}$$

Finally, for all $\alpha\in\mathbb{N}^\nu$, $|\alpha|=1$, we have, by (9.116),

$$\begin{aligned}\|\langle D\rangle\partial_\varphi^{\beta+\alpha}(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+1}{2}-1}h\|_s&\leq_s\|\partial_\varphi^{\beta+\alpha}(\partial_i\Phi[\widehat{\imath}])\langle D\rangle^{-\frac{|\beta|+2}{2}}[\langle D\rangle^{-\frac12}h]\|_{s+1}\\&\leq_s\|\partial_i a[\widehat{\imath}]\|_{s+1+s_0+\frac{3}{2}+|\beta|}\|h\|_{s+1}\,.\end{aligned} \tag{9.137}$$

Hence (9.134)-(9.137) imply the estimate (9.117) with Sobolev index $s+1$. If $h\in H^s$ and $\widehat{\imath}\in\mathcal{Y}^{s+s_0+|\beta|+\frac12+\sigma}$ (resp. $\widehat{\imath}\in\mathcal{Y}^{s+s_0+|\beta|+\frac32+\sigma}$), the estimate (9.116) (resp. (9.117)) follows by density. $\square$

We now estimate the adjoint Φ^* of the time-1 flow $\Phi=\Phi(\varphi,1)$. As in [8] (Lemma 8.2) we represent the adjoint $\Phi^*=\Psi=\Psi(\varphi,0)$ with the backward flow $\Psi(\varphi,t)$ of

$$\partial_t\Psi(\varphi,t)=\mathrm{i}|D|^{\frac12}a\Psi(\varphi,t)\,,\qquad\Psi(\varphi,1)=\mathrm{Id}\,. \tag{9.138}$$

Indeed, since $\Phi(\varphi,t)$ solves (6.131) and $\Psi(\varphi,t)$ solves (9.138), we have, for all $u_0,v_0\in L^2_x(\mathbb{T})$, that

$$\partial_t\big(\Phi(\varphi,t)[u_0]\,,\,\Psi(\varphi,t)[v_0]\big)_{L^2_x}=0\,,\quad\forall t\in[0,1]\,.$$

Therefore $(\Phi(\varphi,1)[u_0]\,,\,v_0)_{L^2_x}=(u_0\,,\,\Psi(\varphi,0)[v_0])_{L^2_x}$, namely

$$\Psi(\varphi,0)=\Phi(\varphi,1)^*=\Phi(\varphi)^*\,. \tag{9.139}$$

The adjoint operator, since it is the flow of (9.138), satisfies properties like those stated in Lemma 9.3.

Proposition 9.17. (Adjoint) *Assume that*

$$\|a\|^{k_0,\gamma}_{2s_0+\frac52+k_0}\leq1\,,\qquad\|a\|_{2s_0+1}\leq\delta(s) \tag{9.140}$$

for some $\delta(s)>0$ small enough. Then for any $k\in\mathbb{N}^{\nu+1}$, $|k|\leq k_0$, for all $s\geq s_0$,

$$\|(\partial_\lambda^k\Phi^*)h\|_s\leq_s\gamma^{-|k|}\big(\|h\|_{s+\frac{|k|}{2}}+\|a\|^{k_0,\gamma}_{s+s_0+|k|+\frac32}\|h\|_{s_0+\frac{|k|}{2}}\big) \tag{9.141}$$

$$\|\partial_\lambda^k(\Phi^*-\mathrm{Id})h\|_s\leq_s\gamma^{-|k|}\big(\|a\|^{k_0,\gamma}_{s_0}\|h\|_{s+\frac{|k|+1}{2}}+\|a\|^{k_0,\gamma}_{s+s_0+|k|+2}\|h\|_{s_0+\frac{|k|+1}{2}}\big)\,. \tag{9.142}$$

Proof. First we take $h \in \mathcal{C}^\infty$.

PROOF OF (9.141). The equation (9.138) can be written as

$$\partial_t \Psi(\varphi,t) = \mathrm{i}a|D|^{\frac{1}{2}}\Psi(\varphi,t) + \mathrm{i}[|D|^{\frac{1}{2}}, a]\Psi(\varphi,t)\,, \qquad \Psi(\varphi,1) = \mathrm{Id}\,,$$

and, by Duhamel principle, one gets

$$\Psi(t) = \Phi(t)\Phi(1)^{-1} - \mathrm{i}\int_t^1 \Phi(t-\tau)[|D|^{\frac{1}{2}}, a]\Psi(\tau)\,d\tau\,. \tag{9.143}$$

By (9.139) the estimate (9.141) follows by proving that, for all $|k| \leq k_0$, $s \geq s_0$,

$$\sup_{t\in[0,1]} \|\partial_\lambda^k \Psi(t)h\|_s \leq_s \gamma^{-|k|}\big(\|h\|_{s+\frac{|k|}{2}} + \|a\|^{k_0,\gamma}_{s+s_0+|k|+\frac{3}{2}}\|h\|_{s_0+\frac{|k|}{2}}\big)\,. \tag{9.144}$$

For $k = 0$, the estimate (9.144) follows by the same proof below (using only (9.143), (9.43), and (9.150) with $k_1 = k_2 = 0$). Then we argue by induction. We assume that (9.144) holds for $k_1 \prec k$ with $|k| \leq k_0$ and we prove it for k. Differentiating (9.143) we get

$$\partial_\lambda^k \Psi(t) = F_1^{(k)}(t) + F_2^{(k)}(t) \tag{9.145}$$

where

$$F_1^{(k)}(t) := \partial_\lambda^k(\Phi(t)\Phi(1)^{-1}) - \mathrm{i}\sum_{k_1+k_2+k_3=k, k_3\prec k}\int_t^1 \partial_\lambda^{k_1}\Phi(t-\tau)[|D|^{\frac{1}{2}}, \partial_\lambda^{k_2}a]\partial_\lambda^{k_3}\Psi(\tau)\,d\tau\,, \tag{9.146}$$

$$F_2^{(k)}(t) := -\mathrm{i}\int_t^1 \Phi(t-\tau)[|D|^{\frac{1}{2}}, a]\partial_\lambda^k\Psi(\tau)\,d\tau\,. \tag{9.147}$$

ESTIMATE OF $F_1^{(k)}(t)$. By (9.98), (9.43) (for $\Phi(1)^{-1}$), and (9.140), we get

$$\|\partial_\lambda^k(\Phi(t)\Phi(1)^{-1})h\|_s \leq_s \gamma^{-|k|}\big(\|h\|_{s+\frac{|k|}{2}} + \|a\|^{k_0,\gamma}_{s+s_0+|k|+1}\|h\|_{s_0+\frac{|k|}{2}}\big) \tag{9.148}$$

and, for all $k_1 + k_2 + k_3 = k$, $k_3 \prec k$,

$$\begin{aligned}\|\partial_\lambda^{k_1}\Phi(t-\tau)[|D|^{\frac{1}{2}}, \partial_\lambda^{k_2}a]\partial_\lambda^{k_3}\Psi(\tau)h\|_s \leq_s \gamma^{-|k_1|}\|[|D|^{\frac{1}{2}}, \partial_\lambda^{k_2}a]\partial_\lambda^{k_3}\Psi(\tau)h\|_{s+\frac{|k_1|}{2}} \\ + \gamma^{-|k_1|}\|a\|^{k_0,\gamma}_{s+s_0+|k_1|+1}\|[|D|^{\frac{1}{2}}, \partial_\lambda^{k_2}a]\partial_\lambda^{k_3}\Psi(\tau)h\|_{s_0+\frac{|k_1|}{2}}\,.\end{aligned} \tag{9.149}$$

By (2.58) we have

$$|\!|\!|[|D|^{\frac{1}{2}}, \partial_\lambda^{k_2}a]|\!|\!|_{-\frac{1}{2},s+\frac{|k_1|}{2},0} \leq_s \|\partial_\lambda^{k_2}a\|_{s+\frac{|k_1|}{2}+\frac{5}{2}} \leq_s \gamma^{-|k_2|}\|a\|^{k_0,\gamma}_{s+\frac{|k_1|}{2}+\frac{5}{2}} \tag{9.150}$$

and, by (9.140), and the inductive hypothesis for $k_3 \prec k$, we get

$$\|[|D|^{\frac{1}{2}}, \partial_\lambda^{k_2}a]\partial_\lambda^{k_3}\Psi(\tau)h\|_{s+\frac{|k_1|}{2}} \leq_s \gamma^{-(|k_2|+|k_3|)}\big(\|h\|_{s+\frac{|k_1|+|k_3|}{2}} + \|a\|^{k_0,\gamma}_{s+s_0+|k|+\frac{3}{2}}\|h\|_{s_0+\frac{|k_1|+|k_3|}{2}}\big)\,. \tag{9.151}$$

Hence (9.146), (9.148), (9.149), (9.151) imply

$$\|F_1^{(k)}(t)h\|_s \leq_s \gamma^{-|k|}\big(\|h\|_{s+\frac{|k|}{2}} + \|a\|_{s+s_0+|k|+\frac{3}{2}}\|h\|_{s_0+\frac{|k|}{2}}\big)\,. \tag{9.152}$$

ESTIMATE OF $F_2^{(k)}(t)$. For all $t, \tau \in [0,1]$, using (9.43), the bound $|\!|\!|[|D|^{\frac{1}{2}}, a]|\!|\!|_{-\frac{1}{2},s,0} \leq_s \|a\|_{s+\frac{5}{2}}$ (see (9.150) with $k_1 = k_2 = 0$), and (9.140) we get

$$\|F_2^{(k)}(t)h\|_s \leq_s \|a\|_{s_0+\frac{5}{2}}\sup_{\tau\in[0,1]}\|\partial_\lambda^k\Psi(\tau)h\|_s + \|a\|_{s+s_0+1}\sup_{\tau\in[0,1]}\|\partial_\lambda^k\Psi(\tau)h\|_{s_0}\,. \tag{9.153}$$

ESTIMATE OF $\partial_\lambda^k \Psi(t)$. By (9.145), (9.152), (9.153) we get

$$
\begin{aligned}
\|\partial_\lambda^k \Psi(t) h\|_s \le_s \gamma^{-|k|} \big(\|h\|_{s+\frac{|k|}{2}} + \|a\|_{s+s_0+|k|+\frac{3}{2}} \|h\|_{s_0+\frac{|k|}{2}} \big) \\
+ \|a\|_{s_0+\frac{5}{2}} \sup_{\tau \in [0,1]} \|\partial_\lambda^k \Psi(\tau) h\|_s + \|a\|_{s+s_0+1} \sup_{\tau\in[0,1]} \|\partial_\lambda^k \Psi(\tau) h\|_{s_0} \, .
\end{aligned} \tag{9.154}
$$

Then, for $s = s_0$, using that, by (9.140), $\|a\|_{2s_0+1} \le \delta(s)$ is small enough, we get

$$
\sup_{t \in [0,1]} \|\partial_\lambda^k \Psi(t) h\|_{s_0} \lessdot \gamma^{-|k|} \big(\|h\|_{s_0+\frac{|k|}{2}} + \|a\|_{2s_0+|k|+\frac{3}{2}} \|h\|_{s_0+\frac{|k|}{2}} \big) \overset{(9.140)}{\lessdot} \gamma^{-|k|} \|h\|_{s_0+\frac{|k|}{2}} \, ,
$$

and therefore, by (9.154), for all $s \ge s_0$,

$$
\sup_{t\in[0,1]} \|\partial_\lambda^k \Psi(t) h\|_s \le_s \gamma^{-|k|} \big(\|h\|_{s+\frac{|k|}{2}} + \|a\|_{s+s_0+|k|+\frac{3}{2}} \|h\|_{s_0+\frac{|k|}{2}} \big) + \|a\|_{s_0+\frac{5}{2}} \sup_{t\in[0,1]} \|\partial_\lambda^k \Psi(t) h\|_s
$$

which yields the estimate (9.144) for $\partial_\lambda^k \Psi(t)$ (using again (9.140) and $\delta(s)$ small enough).
PROOF OF (9.142). By (9.138) we have $\Psi(\varphi,t) - \mathrm{Id} = -\mathrm{i} \int_t^1 |D|^{\frac{1}{2}} a \Psi(\varphi,\tau)\, d\tau$, then it is enough to apply (9.144). If $h \in H^{s+\frac{|k|}{2}}$ (resp. $h \in H^{s+\frac{|k|+1}{2}}$), the estimate (9.141) (resp. (9.142)) follows by density. $\square$

Finally we estimate the variation of the adjoint operator Φ^* with respect to the torus $i(\varphi)$.

Proposition 9.18. *Let $s_1 > s_0$ and assume the condition*

$$
\|a\|_{s_1+s_0+3} \le 1 \, , \quad \|a\|_{s_1+s_0+1} \le \delta(s_1) \, , \tag{9.155}
$$

for some $\delta(s_1) > 0$ small. Then, for all $s \in [s_0, s_1]$,

$$
\|\partial_i \Phi^* [\widehat{\imath}] h\|_s \le_s \|\partial_i a [\widehat{\imath}]\|_{s+s_0+\frac{1}{2}} \|h\|_{s+\frac{1}{2}} \, . \tag{9.156}
$$

Proof. First, we prove that the map $\Psi(t)$ defined in (9.143) satisfies (9.156) for h and $\widehat{\imath}$ which are $\mathcal{C}^\infty$ with respect to φ and x. By differentiating (9.143) we get

$$
\begin{aligned}
\partial_i \Psi(t)[\widehat{\imath}] = \partial_i (\Phi(t)\Phi(1)^{-1})[\widehat{\imath}] - \mathrm{i} \int_t^1 \partial_i \Phi(t-\tau)[\widehat{\imath}] [|D|^{\frac{1}{2}}, a] \Psi(\tau)\, d\tau - \mathrm{i} \int_t^1 \Phi(t-\tau) [|D|^{\frac{1}{2}}, \partial_i a[\widehat{\imath}]] \Psi(\tau)\, d\tau \\
- \mathrm{i} \int_t^1 \Phi(t-\tau)[|D|^{\frac{1}{2}}, a] \partial_i \Psi(\tau)[\widehat{\imath}]\, d\tau \, .
\end{aligned} \tag{9.157}
$$

By (9.116) applied with $\beta = 0$ we get

$$
\|\partial_i \Phi(t)[\widehat{\imath}] h\|_s \le_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{1}{2}} \|h\|_{s+\frac{1}{2}} \, . \tag{9.158}
$$

Moreover by (2.58)

$$
|\!|\!| [|D|^{\frac{1}{2}}, a] |\!|\!|_{-\frac{1}{2},s,0} \le_s \|a\|_{s+\frac{5}{2}} \, , \quad |\!|\!| [|D|^{\frac{1}{2}}, \partial_i a[\widehat{\imath}]] |\!|\!|_{-\frac{1}{2},s,0} \le_s \|\partial_i a[\widehat{\imath}]\|_{s+\frac{5}{2}} \, . \tag{9.159}
$$

Then for all $t \in [0,1]$, by (9.158), (9.43), (9.155),

$$
\|\partial_i (\Phi(t)\Phi(1)^{-1})[\widehat{\imath}] h\|_s \le_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{1}{2}} \|h\|_{s+\frac{1}{2}} \tag{9.160}
$$

and for all $t, \tau \in [0,1]$, by (9.144) (applied for $k = 0$), (9.158), (9.116), (9.159), (9.43) and (9.155) we get, for any $s \in [s_0, s_1]$,

$$
\|\partial_i \Phi(t-\tau)[\widehat{\imath}][|D|^{\frac{1}{2}}, a]\Psi(\tau) h\|_s \, , \, \|\Phi(t-\tau)[|D|^{\frac{1}{2}}, \partial_i a[\widehat{\imath}]] \Psi(\tau) h\|_s \le_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{1}{2}} \|h\|_{s+\frac{1}{2}} \, , \tag{9.161}
$$

$$
\|\Phi(t-\tau)[|D|^{\frac{1}{2}}, a] \partial_i \Psi(\tau)[\widehat{\imath}] h\|_s \le_s \|a\|_{s+\frac{5}{2}} \|\partial_i \Psi(\tau)[\widehat{\imath}] h\|_s \le_s \delta(s_1) \|\partial_i \Psi(\tau)[\widehat{\imath}] h\|_s \, . \tag{9.162}
$$

Therefore (9.157), (9.160), (9.161), (9.162) imply, for all $s \in [s_0, s_1]$,

$$
\sup_{t\in[0,1]} \|\partial_i \Psi(t)[\widehat{\imath}] h\|_s \le_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{1}{2}} \|h\|_{s+\frac{1}{2}} + \delta(s_1) \sup_{t\in[0,1]} \|\partial_i \Psi(t)[\widehat{\imath}] h\|_s
$$

and therefore, taking $\delta(s_1)$ small, $\sup_{t\in[0,1]} \|\partial_i \Psi(t)[\widehat{\imath}] h\|_s \le_s \|\partial_i a[\widehat{\imath}]\|_{s+s_0+\frac{1}{2}} \|h\|_{s+\frac{1}{2}}$, proving (9.156). If $h \in H^{s+\frac{1}{2}}$ and $\widehat{\imath} \in \mathcal{Y}^{s+s_0+\frac{1}{2}+\sigma}$, then the estimate follows by density. $\square$

References

[1] Alazard T., Baldi P. Gravity capillary standing water waves. *Arch. Rat. Mech. Anal*, 217, 3, 741-830, 2015.

[2] Alazard T., Burq N., Zuily C. On the water-wave equations with surface tension. *Duke Math. J.*, 158, 413-499, 2011.

[3] Alazard T., Delort J-M. Sobolev estimates for two dimensional gravity water waves. *Astérisque*, 374, viii + 241, 2015.

[4] Alazard T., Delort J-M. Global solutions and asymptotic behavior for two dimensional gravity water waves. *Ann. Sci. Éc. Norm. Supér.*, 48, no. 5, 1149-1238, 2015.

[5] Baldi P. *Private Communication.*

[6] Baldi P. Periodic solutions of fully nonlinear autonomous equations of Benjamin-Ono type. *Ann. I. H. Poincaré (C) Anal. Non Linéaire*, 30, no. 1, 33-77, 2013.

[7] Baldi P., Berti M., Montalto R. KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation. *Math. Annalen*, 359, 1-2, 471-536, 2014.

[8] Baldi P., Berti M., Montalto R. KAM for autonomous quasi-linear perturbations of KdV. *Ann. I. H. Poincaré (C) Anal. Non Linéaire*, to appear, printed online 3 August 2015.

[9] Baldi P., Berti M., Montalto R. KAM for autonomous quasi-linear perturbations of mKdV. *Bollettino Unione Matematica Italiana*, to appear.

[10] Bambusi D., Berti M., Magistrelli E. Degenerate KAM theory for partial differential equations. *Journal Diff. Equations*, 250, 8, 3379-3397, 2011.

[11] Berti M., Biasco L. Branching of Cantor manifolds of elliptic tori and applications to PDEs. *Comm. Math. Phys.*, 305, 3, 741-796, 2011.

[12] Berti M., Biasco L., Procesi M. KAM theory for the Hamiltonian derivative wave equation. *Ann. Sci. Éc. Norm. Supér. (4)*, 46(2):301-373, 2013.

[13] Berti M., Biasco L., Procesi M. KAM for Reversible Derivative Wave Equations. *Arch. Rat. Mech. Anal.*, 212(3):905-955, 2014.

[14] Berti M., Bolle P. Cantor families of periodic solutions for completely resonant nonlinear wave. equations. *Duke Mathematical Journal*, 134, issue 2, 359-419, 2006.

[15] Berti M., Bolle P. A Nash-Moser approach to KAM theory. *Fields Institute Communications*, 255-284, special volume "Hamiltonian PDEs and Applications", 2015.

[16] Berti M., Bolle P. Quasi-periodic solutions for nonlinear wave equations with a multiplicative potential, in preparation.

[17] Berti M., Corsi L., Procesi M. An Abstract NashMoser Theorem and Quasi-Periodic Solutions for NLW and NLS on Compact Lie Groups and Homogeneous Manifolds. *Comm. Math. Phys.* 334, no. 3, 1413-1454, 2015.

[18] Bourgain J. Green's function estimates for lattice Schrödinger operators and applications, *Annals of Mathematics Studies* 158, Princeton University Press, Princeton, 2005.

[19] Craig W., Nicholls D. Travelling two and three dimensional capillary gravity water waves. *SIAM J. Math. Anal.*, 32(2):323–359 (electronic), 2000.

[20] Craig W., Sulem C. Numerical simulation of gravity waves. *J. Comput. Phys.*, 108(1):73-83, 1993.

[21] Craig W., Sulem C. Normal form transformations for capillary-gravity water waves. *Field Institute Communications*, 73-110, special volume "Hamiltonian PDEs and Applications", 2015.

[22] Craig W., Worfolk P. An integrable normal form for water waves in infinite depth. *Phys. D*, 84, no. 3-4, 513-531, 1995.

[23] Delort, J.-M., Szeftel, J. Long-time existence for small data nonlinear Klein-Gordon equations on tori and spheres. *Int. Math. Res. Not.* 2004, no. 37, 1897-1966.

[24] Eliasson L.H., Kuksin S.. KAM for non-linear Schrödinger equation, *Annals of Math.*, 172, 371-435, 2010.

[25] Fejoz J. Démonstration du théoréme d' Arnold sur la stabilité du systéme planétaire (d' aprés Herman). *Ergodic Theory Dynam. Systems* 24 (5), 1521-1582, 2004.

[26] Feola R., Procesi M. Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations, *J. Diff. Eq.*, 259, no. 7, 3389-3447, 2015.

[27] Hörmander L. The analysis of linear partial differential operators III. Pseudo-differential operartors. *Springer-Verlag*, Berlin, 1990.

[28] Iooss G., Plotnikov P. Existence of multimodal standing gravity waves. *J. Math. Fluid Mech.*, 7, 349–364, 2005.

[29] Iooss G., Plotnikov P. Multimodal standing gravity waves: a completely resonant system. *J. Math. Fluid Mech.*, 7(suppl. 1), 110-126, 2005.

[30] Iooss G., Plotnikov P. Small divisor problem in the theory of three-dimensional water gravity waves. *Mem. Amer. Math. Soc.*, 200(940):viii+128, 2009.

[31] Iooss G., Plotnikov P. Asymmetrical tridimensional travelling gravity waves. *Arch. Rat. Mech. Anal.*, 200(3):789-880, 2011.

[32] Iooss G., Plotnikov P., Toland J. Standing waves on an infinitely deep perfect fluid under gravity. *Arch. Rat. Mech. Anal.*, 177(3):367–478, 2005.

[33] Lannes D. The water waves problem: mathematical analysis and asymptotics. *Mathematical Surveys and Monographs*, 188, 2013.

[34] Levi-Civita T. Détermination rigoureuse des ondes permanentes dampleur finie. *Math. Ann.*, 93 , pp. 264-314, 1925.

[35] Liu J., Yuan X. A KAM theorem for Hamiltonian partial differential equations with unbounded perturbations. *Comm. Math. Phys.*, 307(3), 629–673, 2011.

[36] Kappeler T., Pöschel J. KAM and KdV, *Springer*, 2003.

[37] Kuksin S. Analysis of Hamiltonian PDEs, volume 19 of Oxford Lecture Series in Mathematics and its Applications. *Oxford University Press*, Oxford, 2000.

[38] Métivier G. Para-differential Calculus and Applications to the Cauchy Problem for Nonlinear Systems. *Pubblicazioni Scuola Normale Pisa*, 5, 2008.

[39] Plotnikov P., Toland J. Nash-Moser theory for standing water waves. *Arch. Rat. Mech. Anal.*, 159(1):1–83, 2001.

[40] Pöschel J. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Applied. Math.*, XXXV, 653-695, 1982.

[41] Rüssmann H. Invariant tori in non-degenerate nearly integrable Hamiltonian systems. *Regul. Chaotic Dyn.* 6 (2), 119-204, 2001.

[42] Saranen J., Vainikko G. Periodic Integral and Pseudodifferential Equations with Numerical Approximation. *Springer Monographs in Mathematics*, 2002.

[43] Taylor M. E. Pseudodifferential Operators and Nonlinear PDEs, *Progress in Mathematics*, Birkhäuser, 1991.

[44] Zakharov V. Stability of periodic waves of finite amplitude on the surface of a deep fluid. *Journal of Applied Mechanics and Technical Physics*, 9(2):190–194, 1968.

[45] Zhang J., Gao M., Yuan X. KAM tori for reversible partial differential equations. *Nonlinearity*, 24(4):1189-1228, 2011.

Massimiliano Berti, SISSA, Via Bonomea 265, 34136, Trieste, Italy, *E-mail:* `berti@sissa.it`,
Riccardo Montalto, University of Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland,
E-mail: `riccardo.montalto@math.uzh.ch`