Time quasi-periodic gravity water waves in finite depth

Pietro Baldi, Massimiliano Berti, Emanuele Haus, Riccardo Montalto

Abstract: We prove the existence and the linear stability of Cantor families of small amplitude time *quasi-periodic* standing water wave solutions — namely periodic and even in the space variable x — of a bi-dimensional ocean with finite depth under the action of pure gravity. Such a result holds for all the values of the depth parameter in a Borel set of asymptotically full measure. This is a small divisor problem. The main difficulties are the quasi-linear nature of the gravity water waves equations and the fact that the linear frequencies grow just in a sublinear way at infinity. We overcome these problems by first reducing the linearized operators obtained at each approximate quasi-periodic solution along the Nash-Moser iteration to constant coefficients up to smoothing operators, using pseudo-differential changes of variables that are quasi-periodic in time. Then we apply a KAM reducibility scheme which requires very weak Melnikov non-resonance conditions (losing derivatives both in time and space), which we are able to verify for most values of the depth parameter using degenerate KAM theory arguments.

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

MSC 2010: 76B15, 37K55 (37C55, 35S05).

Contents

1 Introduction and main result

We consider the Euler equations of hydrodynamics for a 2-dimensional perfect, incompressible, inviscid, irrotational fluid under the action of gravity, filling an ocean with finite depth h and with space periodic boundary conditions, namely the fluid occupies the region

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

In this paper we prove the existence and the linear stability of small amplitude quasi-periodic in time solutions of the *pure gravity* water waves system

$$\begin{cases} \partial_t \Phi + \frac{1}{2} |\nabla \Phi|^2 + g\eta = 0 & \text{at } y = \eta(x) \\ \Delta \Phi = 0 & \text{in } \mathcal{D}_\eta \\ \partial_y \Phi = 0 & \text{at } y = -h \\ \partial_t \eta = \partial_y \Phi - \partial_x \eta \cdot \partial_x \Phi & \text{at } y = \eta(x) \end{cases} \tag{1.2}$$

where $g > 0$ is the acceleration of gravity. 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 (1.2) is the Bernoulli condition stating the continuity of the pressure at the free surface. The last equation in (1.2) expresses that the fluid particles on the free surface always remain part of it.

Following Zakharov [61] and Craig-Sulem [26], the evolution problem (1.2) may be written as an infinite-dimensional Hamiltonian system in the unknowns $(\eta(x), \psi(x))$ where, at each instant t,

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

is the trace at the free boundary of the velocity potential. Given the shape $\eta(t, x)$ of the domain top boundary and the Dirichlet value $\psi(t, x)$ of the velocity potential at the top boundary, there is a unique solution $\Phi(t, x, y; h)$ of the elliptic problem

$$\begin{cases} \Delta\Phi = 0 & \text{in } \{-h < y < \eta(t, x)\} \\ \partial_y \Phi = 0 & \text{on } y = -h \\ \Phi = \psi & \text{on } \{y = \eta(t, x)\}\,. \end{cases} \tag{1.3}$$

As proved in [26], system (1.2) is then equivalent to the Craig-Sulem-Zakharov system

$$\begin{cases} \partial_t \eta = G(\eta, h)\psi \\ \partial_t \psi = -g\eta - \dfrac{\psi_x^2}{2} + \dfrac{1}{2(1+\eta_x^2)}\big(G(\eta, h)\psi + \eta_x \psi_x\big)^2 \end{cases} \tag{1.4}$$

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

$$G(\eta, h)\psi := \big\{\Phi_y - \eta_x \Phi_x\big\}_{|y=\eta(t,x)} \tag{1.5}$$

(we denote by η_x the space derivative $\partial_x \eta$). The reason of the name "Dirichlet-Neumann" is that $G(\eta, h)$ maps the Dirichlet datum ψ into the (normalized) normal derivative $G(\eta, h)\psi$ at the top boundary (Neumann datum). The operator $G(\eta, h)$ is linear in ψ, self-adjoint with respect to the L^2 scalar product and positive-semidefinite, and its kernel contains only the constant functions. The Dirichlet-Neumann operator is a *pseudo-differential* operator with principal symbol $D\tanh(hD)$, with the property

$$G(\eta, h) - D\tanh(hD) \in OPS^{-\infty}$$

when $\eta(x) \in \mathcal{C}^\infty$, see Section 3.

Furthermore, equations (1.4) are the Hamiltonian system (see [61], [26])

$$\begin{gathered} \partial_t \eta = \nabla_\psi H(\eta, \psi)\,, \quad \partial_t \psi = -\nabla_\eta H(\eta, \psi) \\ \partial_t u = J\nabla_u H(u)\,, \quad u := \begin{pmatrix} \eta \\ \psi \end{pmatrix}, \quad J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}, \end{gathered} \tag{1.6}$$

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

$$H(\eta, \psi) := H(\eta, \psi, h) := \frac{1}{2}\int_{\mathbb{T}} \psi\, G(\eta, h)\psi\, dx + \frac{g}{2}\int_{\mathbb{T}} \eta^2\, dx \tag{1.7}$$

is the sum of the kinetic and potential energies expressed in terms of the variables (η, ψ). The symplectic structure induced by (1.6) is the standard Darboux 2-form

$$\mathcal{W}(u_1, u_2) := (u_1, Ju_2)_{L^2(\mathbb{T}_x)} = (\eta_1, \psi_2)_{L^2(\mathbb{T}_x)} - (\psi_1, \eta_2)_{L^2(\mathbb{T}_x)} \tag{1.8}$$

for all $u_1 = (\eta_1, \psi_1)$, $u_2 = (\eta_2, \psi_2)$. In the paper we will often write $G(\eta)$, $H(\eta, \psi)$ instead of $G(\eta, h)$, $H(\eta, \psi, h)$, omitting for simplicity to denote the dependence on the depth parameter h.

The phase space of (1.4) is

$$(\eta, \psi) \in H_0^1(\mathbb{T}) \times \dot{H}^1(\mathbb{T}) \qquad \text{where} \qquad \dot{H}^1(\mathbb{T}) := H^1(\mathbb{T})/_\sim \tag{1.9}$$

is the homogeneous space obtained by the equivalence relation $\psi_1(x) \sim \psi_2(x)$ if and only if $\psi_1(x) - \psi_2(x) = c$ is a constant. For simplicity of notation we denote the equivalence class $[\psi]$ by ψ. Note that the second equation in (1.4) is in $\dot{H}^1(\mathbb{T})$, as it is natural because only the gradient of the velocity potential has a physical meaning. Since the quotient map induces an isometry of $\dot{H}^1(\mathbb{T})$ onto $H^1_0(\mathbb{T})$, it is often convenient to identify ψ with a function with zero average.

The water waves system (1.4)-(1.6) exhibits several symmetries. First of all, the mass $\int_{\mathbb{T}} \eta \, dx$ is a first integral of (1.4). In addition, the subspace of functions that are even in x,

$$\eta(x) = \eta(-x) \, , \quad \psi(x) = \psi(-x) \, , \tag{1.10}$$

is invariant under (1.4). 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 $x = k\pi$, for all $k \in \mathbb{Z}$. Hence there is no flow of fluid through the lines $x = k\pi$, $k \in \mathbb{Z}$, and a solution of (1.4) satisfying (1.10) describes the motion of a liquid confined between two vertical walls.

Another important symmetry of the water waves system is reversibility, namely equations (1.4)-(1.6) 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 (1.4) satisfying

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

namely η is even in time and ψ is odd in time. Solutions of the water waves equations (1.4) satisfying (1.10) and (1.12) are called gravity *standing water waves*.

The existence of standing water waves is a small divisor problem, which is particularly difficult because (1.4) is a quasi-linear system of PDEs. The existence of small amplitude time-periodic gravity standing wave solutions for bi-dimensional fluids has been first proved by Plotinkov and Toland [52] in finite depth and by Iooss, Plotnikov and Toland in [41] in infinite depth, see also [37], [38]. More recently, the existence of time periodic gravity-capillary standing wave solutions has been proved by Alazard and Baldi [1]. Next, both the existence and the linear stability of time *quasi-periodic* gravity-capillary standing wave solutions have been proved by Berti and Montalto in [21], see also the expository paper [20].

We also mention that the bifurcation of small amplitude one-dimensional *traveling* gravity water wave solutions (namely traveling waves in bi-dimensional fluids like (1.4)) dates back to Levi-Civita [47]; note that *standing* waves are not traveling because they are even in space, see (1.10). For three-dimensional fluids, the existence of small amplitude traveling water wave solutions with space periodic boundary conditions has been proved by Craig and Nicholls [24] for the gravity-capillary case (which is not a small divisor problem) and by Iooss and Plotinikov [39]-[40] in the pure gravity case (which is a small divisor problem).

The dynamics of the pure gravity and gravity-capillary water waves equations is very different, since in the first case the linear frequencies grow at infinity as $\sim \sqrt{j}$, see (1.19), while in the presence of surface tension they grow as $\sim j^{3/2}$. The sub/super linear growth of the dispersion relation at high frequencies induces quite a relevant difference for the development of KAM theory. As is well known, the abstract infinite-dimensional KAM theorems available in literature, e.g. [43], [44], [53], require that the eigenvalues of the linear constant coefficient differential operator grow as j^α, $\alpha \geq 1$. The reason is that, in presence of a sublinear ($\alpha < 1$) growth of the linear frequencies, one may impose only very weak Melnikov non-resonance conditions, see e.g. (1.36), which produce strong losses of derivatives along the iterative KAM scheme. We overcome this difficulty by a regularization procedure performed on the linearized PDE at each approximate quasi-periodic solution. This a very general idea, which can be applied in a broad class of situations. We shall explain below in detail this key step of the proof.

The main result of this paper — see Theorem 1.1 — proves the existence of small amplitude time quasi-periodic solutions of (1.4) for most values of the depth parameter h. Actually, from a physical point of view, it is also natural to consider the depth h of the ocean as a fixed physical quantity and to look for quasi-periodic solutions for most values of the space wavelength. This can be achieved by rescaling time and space as

$$\tau := \mu t, \quad \tilde{x} := \lambda x \,,$$

and the amplitude of (η, ψ) as

$$\tilde{\eta}(\tau, \tilde{x}) := \lambda \eta(\mu^{-1}\tau, \lambda^{-1}\tilde{x}) = \lambda \eta(t, x)\,, \qquad \tilde{\psi}(\tau, \tilde{x}) := \alpha \psi(\mu^{-1}\tau, \lambda^{-1}\tilde{x}) = \alpha \psi(t, x)\,.$$

Thus $\eta(t, x)$, $\psi(t, x)$ satisfy (1.4) if and only if $\tilde{\eta}(\tau, \tilde{x})$, $\tilde{\psi}(\tau, \tilde{x})$ satisfy

$$\begin{cases} \partial_\tau \tilde{\eta} = \dfrac{\lambda^2}{\alpha\mu} G(\tilde{\eta}, \lambda h)\tilde{\psi} \\ \partial_\tau \tilde{\psi} = -\dfrac{g\alpha}{\lambda\mu}\tilde{\eta} - \dfrac{\lambda^2 \tilde{\psi}_{\tilde{x}}^2}{\alpha\mu 2} + \dfrac{\lambda^2}{\alpha\mu 2(1+\tilde{\eta}_{\tilde{x}}^2)}\Big(G(\tilde{\eta}, \lambda h)\tilde{\psi} + \tilde{\eta}_{\tilde{x}}\tilde{\psi}_{\tilde{x}}\Big)^2 . \end{cases}$$

Choosing the scaling parameters λ, μ, α such that

$$\frac{\lambda^2}{\alpha\mu} = 1\,, \qquad \frac{g\alpha}{\lambda\mu} = 1\,,$$

we obtain system (1.4) where the gravity constant g has been replaced by 1 and the depth parameter h by

$$\mathtt{h} := \lambda h\,. \tag{1.13}$$

The previous scaling implies that, given a fixed value of the depth h, for many values of the parameter λ there exist time quasi-periodic solutions to (1.4) whose space period is $2\pi\lambda$. In this sense, changing the parameter $\mathtt{h}$ can be interpreted as changing the space period of solutions and not the depth of water.

Summarizing, in the sequel of the paper we shall look for time quasi-periodic solutions of the water waves system

$$\begin{cases} \partial_t \eta = G(\eta, \mathtt{h})\psi \\ \partial_t \psi = -\eta - \dfrac{\psi_x^2}{2} + \dfrac{1}{2(1+\eta_x^2)}\big(G(\eta, \mathtt{h})\psi + \eta_x\psi_x\big)^2 \end{cases} \tag{1.14}$$

with $\eta(t) \in H_0^1(\mathbb{T}_x)$ and $\psi(t) \in \dot{H}^1(\mathbb{T}_x)$.

We look for small amplitude solutions of (1.14). Of main importance is therefore the dynamics of the system obtained linearizing (1.14) at the equilibrium $(\eta, \psi) = (0, 0)$, namely

$$\begin{cases} \partial_t \eta = G(0, \mathtt{h})\psi, \\ \partial_t \psi = -\eta \end{cases} \tag{1.15}$$

where $G(0, \mathtt{h}) = D\tanh(\mathtt{h}D)$ is the Dirichlet-Neumann operator at the flat surface $\eta = 0$, namely

$$G(0, \mathtt{h})\cos(jx) = j\tanh(\mathtt{h}j)\cos(jx), \quad G(0, \mathtt{h})\sin(jx) = j\tanh(\mathtt{h}j)\sin(jx), \quad \forall j \in \mathbb{N}\,.$$

In the compact Hamiltonian form as in (1.6), system (1.15) reads

$$\partial_t u = J\Omega u\,, \quad \Omega := \begin{pmatrix} 1 & 0 \\ 0 & G(0, \mathtt{h}) \end{pmatrix}, \tag{1.16}$$

which is the Hamiltonian system generated by the quadratic Hamiltonian (see (1.7))

$$H_L := \frac{1}{2}(u, \Omega u)_{L^2} = \frac{1}{2}\int_{\mathbb{T}} \psi\, G(0, \mathtt{h})\psi\, dx + \frac{1}{2}\int_{\mathbb{T}} \eta^2\, dx\,. \tag{1.17}$$

The solutions of the linear system (1.15), i.e. (1.16), even in x, satisfying (1.12), are

$$\eta(t, x) = \sum_{j\geq 1} a_j \cos(\omega_j t)\cos(jx), \quad \psi(t, x) = -\sum_{j\geq 1} a_j \omega_j^{-1}\sin(\omega_j t)\cos(jx)\,, \tag{1.18}$$

with linear frequencies of oscillation

$$\omega_j := \omega_j(\mathtt{h}) := \sqrt{j\tanh(\mathtt{h}j)}\,, \quad j \geq 1\,. \tag{1.19}$$

Note that, since $j \mapsto j\tanh(\mathtt{h}j)$ is monotone increasing, all the linear frequencies are simple.

The main result of the paper proves that most solutions (1.18) of the linear system (1.15) can be continued to solutions of the nonlinear water waves Hamiltonian system (1.14) for most values of the parameter $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$. More precisely we look for quasi-periodic solutions $u(\tilde{\omega}t) = (\eta, \psi)(\tilde{\omega}t)$ of (1.14), with frequency $\tilde{\omega} \in \mathbb{R}^\nu$ (to be determined), close to some solutions (1.18) of (1.15), in the Sobolev spaces of functions

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

$$H^s := H^s(\mathbb{T}^{\nu+1}, \mathbb{R}) = \Big\{ f = \sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} f_{\ell j}\, e^{\mathrm{i}(\ell\cdot\varphi + jx)} : \ \|f\|_s^2 := \sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} |f_{\ell j}|^2 \langle \ell, j\rangle^{2s} < \infty \Big\}, \tag{1.20}$$

where $\langle \ell, j\rangle := \max\{1, |\ell|, |j|\}$. For

$$s \geq s_0 := \Big[\frac{\nu+1}{2}\Big] + 1 \in \mathbb{N} \tag{1.21}$$

one has $H^s(\mathbb{T}^{\nu+1}, \mathbb{R}) \subset L^\infty(\mathbb{T}^{\nu+1}, \mathbb{R})$, and $H^s(\mathbb{T}^{\nu+1}, \mathbb{R})$ is an algebra.

Fix an arbitrary finite subset $\mathbb{S}^+ \subset \mathbb{N}^+ := \{1, 2, \ldots\}$ (tangential sites) and consider the solutions of the linear equation (1.15)

$$\eta(t,x) = \sum_{j\in\mathbb{S}^+} \sqrt{\xi_j}\cos(\omega_j t)\cos(jx), \quad \psi(t,x) = -\sum_{j\in\mathbb{S}^+} \sqrt{\xi_j}\,\omega_j^{-1}\sin(\omega_j t)\cos(jx)\,, \quad \xi_j > 0\,, \tag{1.22}$$

which are Fourier supported on $\mathbb{S}^+$. We denote by $\nu := |\mathbb{S}^+|$ the cardinality of $\mathbb{S}^+$.

Theorem 1.1. (KAM for gravity water waves in finite depth) *For every choice of the tangential sites $\mathbb{S}^+ \subset \mathbb{N}\setminus\{0\}$, there exists $\bar{s} > \frac{|\mathbb{S}^+|+1}{2}$, $\varepsilon_0 \in (0,1)$ such that for every $|\xi| \leq \varepsilon_0^2$, $\xi := (\xi_j)_{j\in\mathbb{S}^+}$, $\xi_j > 0$ for all $j \in \mathbb{S}^+$, there exists a Cantor-like set $\mathcal{G} \subset [\mathtt{h}_1, \mathtt{h}_2]$ with asymptotically full measure as $\xi \to 0$, i.e.*

$$\lim_{\xi\to 0} |\mathcal{G}| = \mathtt{h}_2 - \mathtt{h}_1\,,$$

such that, for any $\mathtt{h} \in \mathcal{G}$, the gravity water waves system (1.14) *has a time quasi-periodic solution $u(\tilde{\omega}t, x) = (\eta(\tilde{\omega}t, x), \psi(\tilde{\omega}t, x))$, with Sobolev regularity $(\eta, \psi) \in H^{\bar{s}}(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R}^2)$, of the form*

$$\begin{aligned} \eta(t,x) &= \sum_{j\in\mathbb{S}^+} \sqrt{\xi_j}\cos(\tilde{\omega}_j t)\cos(jx) + r_1(\tilde{\omega}t, x), \\ \psi(t,x) &= -\sum_{j\in\mathbb{S}^+} \sqrt{\xi_j}\,\omega_j^{-1}\sin(\tilde{\omega}_j t)\cos(jx) + r_2(\tilde{\omega}t, x) \end{aligned} \tag{1.23}$$

with a Diophantine frequency vector $\tilde{\omega} := (\tilde{\omega}_j)_{j\in\mathbb{S}^+} \in \mathbb{R}^\nu$ satisfying $\tilde{\omega} \to \vec{\omega}(\mathtt{h}) := (\omega_j(\mathtt{h}))_{j\in\mathbb{S}^+}$ as $\xi \to 0$, and the functions $r_1(\varphi, x), r_2(\varphi, x)$ are $o(\sqrt{|\xi|})$-small in $H^{\bar{s}}(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$, i.e. $\|r_i\|_{\bar{s}}/\sqrt{|\xi|} \to 0$ as $|\xi| \to 0$ for $i = 1, 2$. The solution (η, ψ) is even in x, η is even in t and ψ is odd in t. In addition these quasi-periodic solutions are linearly stable.

Let us make some comments on the result.

1. The parameter $\mathtt{h}$ varies in the finite interval $[\mathtt{h}_1, \mathtt{h}_2]$ with $0 < \mathtt{h}_1 < \mathtt{h}_2 < +\infty$, and all the estimates depend on $\mathtt{h}_1, \mathtt{h}_2$. The result does not pass to the limit of zero ($\mathtt{h}_1 \to 0^+$) nor infinite ($\mathtt{h}_2 \to +\infty$) *rescaled* depth parameter $\mathtt{h}$ (recall (1.13)). In those limit regimes different phenomena arise.

2. Note that the linear frequencies (1.19) admit the asymptotic expansion

$$\sqrt{j\tanh(\mathtt{h}j)} = \sqrt{j} + r(j, \mathtt{h}) \quad \text{where} \quad \big|\partial_{\mathtt{h}}^k r(j, \mathtt{h})\big| \leq C_k e^{-\mathtt{h}j} \quad \forall k \in \mathbb{N},\ \forall j \geq 1, \tag{1.24}$$

   uniformly in $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, where the constant C_k depends only on k and $\mathtt{h}_1$. Despite the fact that $\mathtt{h}$ changes the frequencies of exponentially small terms, we shall use the finite depth parameter $\mathtt{h}$ to impose the required Melnikov non-resonance conditions.

3. No global in time existence results concerning the initial value problem of the water waves equations (1.4) under *periodic* boundary conditions are known so far. Global existence results have been proved for smooth Cauchy data rapidly decaying at infinity in $\mathbb{R}^d$, $d = 1, 2$, exploiting the dispersive properties of the flow. For three dimensional fluids (i.e. $d = 2$) it has been proved independently by Germain-Masmoudi-Shatah [31] and Wu [60]. In the more difficult case of bi-dimensional fluids (i.e. $d = 1$) it has been proved by Alazard-Delort [4] and Ionescu-Pusateri [36].

   In the case of periodic boundary conditions, Ifrim-Tataru [35] proved for small initial data a cubic life span time of existence, which is longer than the one just provided by the local existence theory, see for example [3]. For longer times, we mention the almost global existence result in Berti-Delort [19] for gravity-capillary space periodic water waves.

   The present Nash-Moser-KAM iterative procedure selects many values of the parameter $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ that give rise to the quasi-periodic solutions (1.23), which are defined for all times. Clearly, by a Fubini-type argument it also results that, for most values of $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, there exist quasi-periodic solutions of (1.14) for most values of the amplitudes $|\xi| \leq \varepsilon_0^2$. The fact that we find quasi-periodic solutions only restricting to a proper subset of parameters is not a technical issue, because the gravity water waves equations (1.4) are expected to be not integrable, see [27], [28] in the case of infinite depth.

4. The quasi-periodic solutions (1.23) are mainly supported in Fourier space on the tangential sites $\mathbb{S}^+$. The dynamics of the water waves equations (1.4) on the symplectic subspaces

$$H_{\mathbb{S}^+} := \Big\{ v = \sum_{j \in \mathbb{S}^+} \begin{pmatrix} \eta_j \\ \psi_j \end{pmatrix} \cos(jx) \Big\}, \quad H_{\mathbb{S}^+}^{\perp} := \Big\{ z = \sum_{j \in \mathbb{N} \setminus \mathbb{S}^+} \begin{pmatrix} \eta_j \\ \psi_j \end{pmatrix} \cos(jx) \in H_0^1(\mathbb{T}_x) \Big\}, \tag{1.25}$$

   is quite different. We shall 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 shall describe the dynamics by introducing the action-angle variables $(\theta, I) \in \mathbb{T}^\nu \times \mathbb{R}^\nu$ in Section 5.

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], [37], [38], [41]). 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 (6.17) and [16]) in which the water waves Hamiltonian reads

$$\omega \cdot y + \frac{1}{2} K_{20}(\phi) y \cdot y + \big(K_{11}(\phi) y, w \big)_{L^2(\mathbb{T}_x)} + \frac{1}{2} \big(K_{02}(\phi) w, w \big)_{L^2(\mathbb{T}_x)} + K_{\geq 3}(\phi, y, w), \tag{1.26}$$

where $K_{\geq 3}$ collects the terms at least cubic in the variables (y, w) (see (6.19) and note that, at a solution, one has $\partial_\phi K_{00} = 0$, $K_{10} = \omega$, $K_{01} = 0$ by Lemma 6.5). 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

$$\begin{cases} \dot\phi = K_{20}(\omega t)[y] + K_{11}^T(\omega t)[w] \\ \dot y = 0 \\ \dot w = J K_{02}(\omega t)[w] + J K_{11}(\omega t)[y] \, . \end{cases} \tag{1.27}$$

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

$$\dot w = J K_{02}(\omega t)[w] + J K_{11}(\omega t)[y(0)] \, . \tag{1.28}$$

The self-adjoint operator $K_{02}(\omega t)$ (defined in (6.19)) 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 written in (1.38), up to a finite dimensional remainder, see Lemma 7.1.

In Sections 7-15 we prove the existence of a bounded and invertible "symmetrizer" map, see (14.2), (15.105), such that, for all $\varphi \in \mathbb{T}^\nu$,

$$\mathbf{W}_\infty(\varphi) : \big(H^s(\mathbb{T}_x, \mathbb{C}) \times H^s(\mathbb{T}_x, \mathbb{C})\big) \cap H^\perp_{\mathbb{S}^+} \to \big(H^{s-\frac14}(\mathbb{T}_x, \mathbb{R}) \times H^{s+\frac14}(\mathbb{T}_x, \mathbb{R})\big) \cap H^\perp_{\mathbb{S}^+} , \tag{1.29}$$

$$\mathbf{W}^{-1}_\infty(\varphi) : \big(H^{s-\frac14}(\mathbb{T}_x, \mathbb{R}) \times H^{s+\frac14}(\mathbb{T}_x, \mathbb{R})\big) \cap H^\perp_{\mathbb{S}^+} \to \big(H^s(\mathbb{T}_x, \mathbb{C}) \times H^s(\mathbb{T}_x, \mathbb{C})\big) \cap H^\perp_{\mathbb{S}^+} , \tag{1.30}$$

and, under the change of variables

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

equation (1.28) transforms into the (complex) diagonal system

$$\partial_t w_\infty = -\mathrm{i}\mathbf{D}_\infty w_\infty + f_\infty(\omega t) , \quad f_\infty(\omega t) := \mathbf{W}^{-1}_\infty(\omega t) J K_{11}(\omega t)[y(0)] = \begin{pmatrix} \mathtt{f}_\infty(\omega t) \\ \overline{\mathtt{f}}_\infty(\omega t) \end{pmatrix} \tag{1.31}$$

where i is the imaginary unit and, denoting $\mathbb{S}_0 := \mathbb{S}^+ \cup (-\mathbb{S}^+) \cup \{0\} \subseteq \mathbb{Z}$ and $\mathbb{S}^c_0 := \mathbb{Z} \setminus \mathbb{S}_0$,

$$\mathbf{D}_\infty := \begin{pmatrix} D_\infty & 0 \\ 0 & -D_\infty \end{pmatrix} , \quad D_\infty := \mathrm{diag}_{j \in \mathbb{S}^c_0} \{\mu^\infty_j\} , \quad \mu^\infty_j \in \mathbb{R} , \tag{1.32}$$

is a Fourier multiplier operator of the form (see (16.38), (15.23), (15.8), (13.78), (13.79))

$$\mu^\infty_j := \mathtt{m}^\infty_{\frac12} |j|^{\frac12} \tanh^{\frac12}(\mathtt{h}|j|) + r^\infty_j , \ j \in \mathbb{S}^c_0 , \quad r^\infty_j = r^\infty_{-j} , \tag{1.33}$$

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

$$\mathtt{m}^\infty_{\frac12} = 1 + O(|\xi|^{\mathtt{a}}) , \quad \sup_{j \in \mathbb{S}^c_0} |j|^{\frac12} |r^\infty_j| = O(|\xi|^{\mathtt{a}}) .$$

Actually by (5.21)-(5.22) and (5.25) we also have a control of the derivatives of $\mathtt{m}^\infty_{\frac12}$ and r^∞_j with respect to $(\omega, \mathtt{h})$. The purely imaginary numbers $\mathrm{i}\mu^\infty_j$ are the *Floquet exponents* of the quasi-periodic solution. The second equation of system (1.31) is, in fact, the complex conjugate of the first one, and (1.31) reduces to the infinitely many decoupled scalar equations

$$\partial_t \mathtt{w}_{\infty,j} = -\mathrm{i}\mu^\infty_j \mathtt{w}_{\infty,j} + \mathtt{f}_{\infty,j}(\omega t) , \quad \forall j \in \mathbb{S}^c_0 .$$

By variation of constants the solutions are

$$\mathtt{w}_{\infty,j}(t) = c_j e^{-\mathrm{i}\mu^\infty_j t} + \mathtt{v}_{\infty,j}(t) \qquad \text{where} \qquad \mathtt{v}_{\infty,j}(t) := \sum_{\ell \in \mathbb{Z}^\nu} \frac{\mathtt{f}_{\infty,j,\ell}\, e^{\mathrm{i}\omega\cdot\ell t}}{\mathrm{i}(\omega\cdot\ell + \mu^\infty_j)} , \quad \forall j \in \mathbb{S}^c_0 . \tag{1.34}$$

Note that the first Melnikov conditions (5.23) hold at a solution, so that $\mathtt{v}_{\infty,j}(t)$ in (1.34) is well defined. Moreover (1.29) and (1.31) imply that $\|f_\infty(\omega t)\|_{H^s_x \times H^s_x} \le C|y(0)|$ for all t. As a consequence, the Sobolev norm of the solution of (1.31) with initial condition $w_\infty(0) \in H^{\mathfrak{s}_0}(\mathbb{T}_x) \times H^{\mathfrak{s}_0}(\mathbb{T}_x)$, for some $\mathfrak{s}_0 \in (s_0, s)$ (in a suitable range of values), satisfies

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

and, for all $t \in \mathbb{R}$, using (1.29), (1.30), we get

$$\|(\eta, \psi)(t)\|_{H^{\mathfrak{s}_0 - \frac14}_x \times H^{\mathfrak{s}_0 + \frac14}_x} \le C \|(\eta(0), \psi(0))\|_{H^{\mathfrak{s}_0 - \frac14}_x \times H^{\mathfrak{s}_0 + \frac14}_x} ,$$

which proves the linear stability of the torus. Note that the profile $\eta \in H^{\mathfrak{s}_0 - \frac14}(\mathbb{T}_x)$ is less regular than the velocity potential $\psi \in H^{\mathfrak{s}_0 + \frac14}(\mathbb{T}_x)$, as it happens for pure gravity waves, see [2].

Clearly a crucial point is the diagonalization of (1.28) into (1.32). With respect to the pioneering works of Plotnikov-Toland [52] and Iooss-Plotnikov-Toland [41] dealing with time periodic solutions, this requires to analyze more in detail the linearized operator in two respects:

1. We have to perform a reduction of the linearized operator into a constant coefficient pseudo-differential operator, up to smoothing remainders, via changes of variables that are quasi-periodic transformations of the phase space, so that the dynamical system nature of the transformed systems is preserved. We shall perform such reductions in Sections 7-14 by changes of variables generated by pseudo-differential operators, diffeomorphisms of the torus, and "semi-Fourier integral operators" (namely pseudo-differential operators of type $(\frac{1}{2}, \frac{1}{2})$ in the notation of Hörmander [34]), inspired by [1], [21].

2. Once the above regularization has been performed, we implement in Section 15 a KAM iterative scheme which completes the diagonalization of the linearized operator. This scheme uses very weak second order Melnikov non-resonance conditions which lose derivatives. This loss is compensated by the smoothing nature of the variable coefficients remainders.

   This diagonalization is not required for the construction of time-periodic solutions, as in [1], [41], [37], [38], [52]. 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.

We shall explain these steps in detail in Section 1.1.

Literature about KAM for PDEs. KAM theory for PDEs has been developed to a large extent for perturbations that are bounded and with linear frequencies growing in a superlinear way, as j^α, $\alpha \geq 1$. The case $\alpha = 1$, which corresponds to Klein-Gordon equations, is more delicate. In the sublinear case $\alpha < 1$, as far as we know, there are no KAM results in literature, since the second order Melnikov conditions lose derivatives. Of course we can regard the existence results for PDEs in higher space dimension under this respect because the eigenvalues grow, according to the Weyl law, like $\sim j^{2/d}$ (which is a strictly sublinear rate if the space dimension d is larger than 2), and the known results use the fact that one has a PDE on a torus or a Lie group. In such cases one proves specific properties of clustering of the eigenvalues, according to a different counting, and uses properties of "localization with respect to the exponentials" of the corresponding eigenfunctions, see for example [22], [32], [15], [18], [54]. In the present case the linear frequencies grow as $\sqrt{j}$ and we perform a very detailed analysis of the water waves nonlinearity.

The existence of quasi-periodic solutions of PDEs (which we shall call, in a broad sense, KAM theory) with *unbounded* perturbations (i.e. the nonlinearity contains derivatives) has been first proved by Kuksin [44] and Kappeler-Pöschel [42] for KdV, then by Liu-Yuan [48], Zhang-Gao-Yuan [63] for derivative NLS, and by Berti-Biasco-Procesi [13]-[14] for derivative NLW. All these previous results still refer to semilinear perturbations, i.e. where 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 [30] for quasi-linear perturbations of Schrödinger equations and by Montalto [50] for the Kirchhoff equation.

Acknowledgements. This research was supported by PRIN 2015 "Variational methods, with applications to problems in mathematical physics and geometry", by the European Research Council under FP7, project no. 306414 "Hamiltonian PDEs and small divisor problem: a dynamical systems approach" (HamPDEs), partially by the Swiss National Science Foundation, and partially by the Programme STAR, funded by Compagnia di San Paolo and UniNA.

1.1 Ideas of the proof

There are three major difficulties for proving the existence of time quasi-periodic solutions of the gravity water waves equations (1.14):

1. Equations (1.14) are a quasi-linear system.

2. The dispersion relation (1.19) of the linear water waves equations is sublinear, i.e. $\omega_j \sim \sqrt{j}$ for $j \to \infty$.

3. One has to verify all the Melnikov non-resonance conditions required on the frequencies by the KAM scheme.

We present below the key ideas of the paper to solve these three major problems. We start by the last one, i.e. how to verify the non-resonance conditions which play a key role for the perturbation theory of quasi-periodic solutions.

1. *Bifurcation analysis and degenerate KAM theory.* The first key observation is that we can use effectively the depth parameter $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ to impose all the required Melnikov non-resonance conditions. Indeed we can prove that, for most values of $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, the unperturbed linear frequencies (1.19) are Diophantine and they satisfy also first and second order Melnikov non-resonance conditions: more precisely the unperturbed tangential frequency vector $\vec{\omega}(\mathtt{h}) := (\omega_j(\mathtt{h}))_{j \in \mathbb{S}^+}$ satisfies

$$|\vec{\omega}(\mathtt{h}) \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau}, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \quad \langle \ell \rangle := \max\{1, |\ell|\}, \tag{1.35}$$

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

$$\begin{aligned} |\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h})| &\geq \gamma j^{\frac{1}{2}} \langle \ell \rangle^{-\tau}, \, \forall \ell \in \mathbb{Z}^\nu, \, j \in \mathbb{N}^+ \setminus \mathbb{S}^+ \, , \\ |\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) + \Omega_{j'}(\mathtt{h})| &\geq \gamma (j^{\frac{1}{2}} + j'^{\frac{1}{2}}) \langle \ell \rangle^{-\tau}, \forall \ell \in \mathbb{Z}^\nu, j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \, , \\ |\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})| &\geq \gamma j^{-\mathtt{d}} j'^{-\mathtt{d}} \langle \ell \rangle^{-\tau}, \forall (\ell, j, j') \neq (0, j, j) \, . \end{aligned} \tag{1.36}$$

The verification of (1.35)-(1.36) is a problem of Diophantine approximation on submanifolds as in [55]. It can be solved by degenerate KAM theory (explained below), exploiting the fact that the linear frequencies $\mathtt{h} \mapsto \omega_j(\mathtt{h})$ are *analytic*, simple (in the subspace of functions even in x), they grow asymptotically like $\sqrt{j}$ for $j \to \infty$, and they are *non-degenerate* in the sense of Bambusi-Berti-Magistrelli [11].

For such values of $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, the solutions (1.22) of the linear equation (1.15) are already sufficiently good approximate quasi-periodic solutions of the nonlinear water waves system (1.4). Since the parameter space $[\mathtt{h}_1, \mathtt{h}_2]$ is fixed, independently of the $O(\varepsilon)$-neighborhood of the origin where we look for the solutions, the small divisor constant γ in (1.35)-(1.36) can be taken $\gamma = o(1)$ as $\varepsilon \to 0$. Actually for simplicity we take $\gamma = o(\varepsilon^a)$ with $a > 0$ small as needed, see (5.25). As a consequence, in order to prove the continuation of the solutions (1.22) of the linearized PDE (1.15) to solutions of the nonlinear water waves system (1.14), all the terms which are at least quadratic in (1.14) are already perturbative. The precise meaning is that in (5.1) it is sufficient to regard the vector field $\varepsilon X_{P_\varepsilon}$ as a perturbation of the linear vector field $J\Omega$.

Along the Nash-Moser-KAM iteration we need to verify that the *perturbed* frequencies, and not only the unperturbed linear ones, are Diophantine and satisfy first and second order Melnikov non-resonance conditions, see the explicit conditions in (5.23). It is for this purpose that we find it convenient to develop degenerate KAM theory as in [11], [21], and to formulate the problem as a Nash-Moser theorem of "hypothetical conjugation" as in [21].

Notice that in the case of infinite depth $\mathtt{h} = +\infty$ the linear frequencies (1.19) are exactly $\sqrt{j}$ and therefore some of the unperturbed Melnikov non-resonance conditions (1.36) are certainly violated. As a consequence, the corresponding perturbed non-resonance conditions can hold only with a small constant $\gamma = o(\varepsilon^2)$. In this case, existence of pure gravity quasi-periodic solutions is still an open problem.

Regarding second order Melnikov non-resonance conditions, two relevant differences with respect to the capillary-gravity case studied in [21] are the following:

   (a) The linear frequencies $\omega_j(\mathtt{h}) = \sqrt{j \tanh(\mathtt{h}j)}$ in (1.19) grow in a sublinear way as $\sqrt{j}$ as $j \to \infty$, and not as $\sim j^{3/2}$ as for the gravity-capillary dispersion relation $\sqrt{(1 + \kappa j^2) j \tanh(\mathtt{h}j)}$.

   (b) The parameter $\mathtt{h}$ moves the frequencies $\omega_j(\mathtt{h})$ of exponentially small quantities of order $O(e^{-\mathtt{h}j})$ (on the contrary, the surface tension parameter κ moves the frequencies of polynomial quantities $O(j^{3/2})$).

As a consequence, we can prove that the second Melnikov non-resonance conditions in (1.36), and the corresponding ones in (5.23), hold for most values of the parameter $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ only if $\mathtt{d}$ is large enough, i.e. $\mathtt{d} > \frac{3}{4} k_0^*$ in Theorem 5.2. The larger is $\mathtt{d}$, the weaker are such Melnikov conditions, and the stronger will be the loss of derivatives due to the small divisors in the reducibility scheme of Section 15. In order to guarantee the convergence of such a KAM reducibility scheme, these losses of derivatives will be compensated by the regularization procedure of Sections 7-14, where we reduce the linearized operator to constant coefficients up to very regularizing terms $O(|D_x|^{-M})$ for some $M := M(\mathtt{d}, \tau)$ large enough, fixed in (15.16), which is large with respect to $\mathtt{d}$ and τ by (15.10). We shall explain in detail this procedure below.

2. *A Nash-Moser Theorem of hypothetical conjugation.* The expected quasi-periodic solutions of the autonomous Hamiltonian system (1.14) will have shifted frequencies $\tilde{\omega}_j$ – to be found – close to the linear frequencies $\omega_j(\mathtt{h})$ in (1.19). The perturbed frequencies depend on the nonlinearity and on the amplitudes ξ_j. Since the Melnikov non-resonance conditions are naturally imposed on ω, it is convenient to use the functional setting of Theorem 5.1 where the parameters are the *frequencies* $\omega \in \mathbb{R}^\nu$ and we introduce a "counter-term" $\alpha \in \mathbb{R}^\nu$ in the family of Hamiltonians H_α defined in (5.12).

   Then the goal is to prove that, for ε small enough, for "most" parameters $(\omega, \mathtt{h})$, there exists a value of the constants $\alpha := \alpha_\infty(\omega, \mathtt{h}, \varepsilon) = \omega + O(\varepsilon\gamma^{-k_0})$ and a ν-dimensional embedded torus $\mathcal{T} = i(\mathbb{T}^\nu)$, close to $\mathbb{T}^\nu \times \{0\} \times \{0\}$, that is invariant for the Hamiltonian vector field $X_{H_{\alpha_\infty(\omega,\mathtt{h},\varepsilon)}}$ and supports quasi-periodic solutions with frequency ω. This is equivalent to looking for a zero of the nonlinear operator $\mathcal{F}(i, \alpha, \omega, \mathtt{h}, \varepsilon) = 0$ defined in (5.13). This equation is solved in Theorem 5.1 by a Nash-Moser iterative scheme. The value of $\alpha := \alpha_\infty(\omega, \mathtt{h}, \varepsilon)$ is adjusted along the iteration in order to control the average of the first component of the Hamilton equation (5.13), especially for solving the linearized equation (6.36), in particular (6.40).

   The set $\mathcal{C}_\infty^\gamma$ of parameters $(\omega, \mathtt{h})$ for which the invariant torus exists is the explicit set defined in (5.23), where we require ω to satisfy, in addition to the Diophantine property

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

   the first and second Melnikov non-resonance conditions stated in (5.23).

   Note that the set $\mathcal{C}_\infty^\gamma$ is defined in terms of the "final torus" i_∞ (see (5.20)) and the "final eigenvalues" in (5.21) which are defined for *all* the values of the frequency $\omega \in \mathbb{R}^\nu$ and $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ by a Whitney extension argument (we shall use the abstract Whitney extension theorem reported in Appendix A). This formulation completely decouples the Nash-Moser iteration (which provides the torus $i_\infty(\omega, \mathtt{h}, \varepsilon)$ and the constant $\alpha_\infty(\omega, \mathtt{h}, \varepsilon) \in \mathbb{R}^\nu$) from the discussion about the measure of the set of parameters where all the non-resonance conditions are indeed verified. This simplifies the analysis of the measure estimates, which are verified once and for all in Section 5.2.

   In order to prove the existence of quasi-periodic solutions of the water waves equations (1.14), and not only of the system with modified Hamiltonian H_α with $\alpha := \alpha_\infty(\omega, \mathtt{h}, \varepsilon)$, we have then to prove that the curve of the unperturbed linear frequencies

$$[\mathtt{h}_1, \mathtt{h}_2] \ni \mathtt{h} \mapsto \vec{\omega}(\mathtt{h}) := (\sqrt{j \tanh(\mathtt{h}j)})_{j \in \mathbb{S}^+} \in \mathbb{R}^\nu$$

   intersects the image $\alpha_\infty(\mathcal{C}_\infty^\gamma)$ of the set $\mathcal{C}_\infty^\gamma$ under the map α_∞, for "most" values of $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$. This is proved in Theorem 5.2 by degenerate KAM theory. For such values of $\mathtt{h}$ we have found a quasi-periodic solution of (1.14) with Diophantine frequency $\omega_\varepsilon(\mathtt{h}) := \alpha_\infty^{-1}(\vec{\omega}(\mathtt{h}), \mathtt{h})$, where $\alpha_\infty^{-1}(\cdot, \mathtt{h})$ is the inverse of the function $\alpha_\infty(\cdot, \mathtt{h})$ at a fixed $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$.

   The above perspective is in the spirit of the Theorem of hypothetical conjugation of Herman proved by Féjoz [29] for finite dimensional Hamiltonian systems. A relevant difference is that in [29], in addition to α, also the normal frequencies are introduced as independent parameters, unlike in Theorem 5.1. Actually for PDEs the present formulation seems to be more convenient: it is a major point of the work to deduce the asymptotic expansion (1.33) of the Floquet exponents.

3. *Degenerate KAM theory and measure estimates.* In Theorem 5.2 we prove that for all the values of $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ except a set of small measure $O(\gamma^{1/k_0^*})$ (the value of $k_0^* \in \mathbb{N}$ is fixed once and for all in Section 4) the vector $(\alpha_\infty^{-1}(\vec{\omega}(\mathtt{h}), \mathtt{h}), \mathtt{h})$ belongs to the set $\mathcal{C}_\infty^\gamma$, see the set $\mathcal{G}_\varepsilon$ in (5.26). As already said, we use in an essential way the fact that the unperturbed frequencies $\mathtt{h} \mapsto \omega_j(\mathtt{h})$ are *analytic* and simple (on the subspace of the even functions), they grow asymptotically as $j^{1/2}$ and they are *non-degenerate* in the sense of [11]. This is verified in Lemma 4.2 as in [11] by analyticity and a generalized Van der Monde determinant. Then we develop degenerate KAM theory which reduces this qualitative non-degeneracy condition to a quantitative one, which is sufficient to estimate effectively the measure of the set $\mathcal{G}_\varepsilon$ by the classical Rüssmann lemma. We deduce in Proposition 4.4 that there exist $k_0^* > 0$, $\rho_0 > 0$ such that, for all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$,

$$\max_{0 \le k \le k_0} \left| \partial_{\mathtt{h}}^k \big(\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h}) \big) \right| \ge \rho_0 \langle \ell \rangle \,, \quad \forall \ell \ne 0 \,, \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \,, \tag{1.37}$$

and similarly for the 0-th, 1-st and 2-nd order Melnikov non-resonance condition with the + sign. Note that the restriction to the subspace of functions with zero average in x eliminates the zero frequency $\omega_0 = 0$, which is trivially resonant (this is used also in [27]). Property (1.37) implies that for "most" parameters $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ the unperturbed linear frequencies $(\vec{\omega}(\mathtt{h}), \vec{\Omega}(\mathtt{h}))$ satisfy the Melnikov conditions of 0-th, 1-st and 2-nd order (but we do not use it explicitly). Actually, condition (1.37) is stable under perturbations that are small in $\mathcal{C}^{k_0}$-norm, see Lemma 5.4. Since the perturbed Floquet exponents in (5.29) are small perturbations of the unperturbed linear frequencies $\sqrt{j \tanh(\mathtt{h} j)}$ in $\mathcal{C}^{k_0}$-norm, with $k_0 := k_0^* + 2$, the "transversality" property (1.37) still holds for the perturbed frequencies $\omega_\varepsilon(\mathtt{h})$ defined in (5.27). As a consequence, by applying the classical Rüssmann lemma (Theorem 17.1 in [57]) we prove that the set of non-resonant parameters $\mathcal{G}_\varepsilon$ has a large measure, see Lemma 5.5 and the end of the proof of Theorem 5.2.

We conclude this discussion underlining two important points (that we have already mentioned):

(a) It is possible to use effectively $\mathtt{h}$ as a parameter to impose the second order Melnikov non-resonance conditions, even though $\mathtt{h}$ moves the linear frequencies $\omega_j(\mathtt{h}) = \sqrt{j \tanh(\mathtt{h} j)}$ in (1.19) just of exponentially small terms.

(b) The second Melnikov conditions that we (can) impose are very weak. The loss of derivatives that they produce will be compensated by the reduction to constant coefficients up to very regularizing remainders as we explain below.

Analysis of the linearized operators. The other crucial point is to prove that the linearized operators obtained at any approximate solution along the Nash-Moser iterative scheme are, for most values of the parameters, invertible, and that their inverse satisfies *tame* estimates in Sobolev spaces (with, of course, loss of derivatives). This is the key assumption to implement in Section 16 a convergent differentiable Nash-Moser iterative scheme in scales of Sobolev spaces.

Linearizing the water waves equations (1.14) at a time-quasi-periodic approximate solution $(\eta, \psi)(\omega t, x)$, and changing ∂_t into the directional derivative $\omega \cdot \partial_\varphi$, we obtain (see (7.6)) the operator

$$\mathcal{L} = \omega \cdot \partial_\varphi + \begin{pmatrix} \partial_x V + G(\eta) B & -G(\eta) \\ (1 + B V_x) + B G(\eta) B & V \partial_x - B G(\eta) \end{pmatrix} \tag{1.38}$$

where the functions B, V are given in (3.2). It turns out that $(V, B) = \nabla_{x,y} \Phi$ is the velocity field evaluated at the free surface $(x, \eta(\omega t, x))$.

By the symplectic procedure developed in Berti-Bolle [16] for autonomous PDEs, and implemented in [8]-[9], [21], it is sufficient to prove the invertibility of (a finite rank perturbation of) the operator $\mathcal{L}$ restricted to the normal subspace $\Pi_{\mathbb{S}^+}^\perp$ introduced in (1.25), see (7.5). We refer to [23] for a similar reduction which applies also to PDEs which are not Hamiltonian, but for example reversible.

In Sections 7-15 we conjugate the operator $\mathcal{L}$ in (1.38) to a diagonal system of infinitely many decoupled, constant coefficients, scalar linear equations, see (1.40) below. Our approach involves two well separated procedures that we shall describe in detail:

1. *Symmetrization and diagonalization of $\mathcal{L}$ up to smoothing operators.* The goal of Sections 7-14 is to conjugate $\mathcal{L}$ to an operator of the form

$$\omega \cdot \partial_\varphi + \mathrm{i}\mathtt{m}_{\frac{1}{2}} |D|^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h}|D|) + \mathrm{i}r(D) + \mathcal{R}_0(\varphi) \tag{1.39}$$

where $\mathtt{m}_{\frac{1}{2}} \approx 1$ is a real constant, independent of φ, the symbol $r(\xi)$ is real and independent of (φ, x), of order $S^{-1/2}$, and the remainder $\mathcal{R}_0(\varphi)$, as well as $\partial_\varphi^\beta \mathcal{R}_0$ for all $|\beta| \leq \beta_0$ large enough, is a small, still variable coefficient operator, which is regularizing at a sufficiently high order, and satisfies tame estimates in Sobolev spaces.

2. *KAM reducibility.* In Section 15 we implement an iterative diagonalization scheme to reduce quadratically the size of the perturbation $\mathcal{R}_0(\varphi)$ in (1.39), completing the conjugation of $\mathcal{L}$ to a diagonal, constant coefficient system of the form

$$\omega \cdot \partial_\varphi + \mathrm{i}\mathrm{Op}(\mu_j) \tag{1.40}$$

where $\mu_j = \mathtt{m}_{\frac{1}{2}} |j|^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h}|j|) + r(j) + \tilde{r}(j)$ are real and $\tilde{r}(j)$ are small. The numbers $\mathrm{i}\mu_j$ are the perturbed Floquet exponents of the quasi-periodic solution.

We underline that all the transformations performed in Sections 7-15 are quasi-periodically-time-dependent changes of variables acting in phase spaces of functions of x (quasi-periodic Floquet operators). Therefore, they preserve the dynamical system structure of the conjugated linear operators.

All these changes of variables are bounded and satisfy tame estimates between Sobolev spaces. As a consequence, the estimates that we shall obtain on the final system (1.40) directly provide good tame estimates for the inverse of the operator (1.38) in the original physical coordinates.

We also note that the original system $\mathcal{L}$ is reversible and even and that all the transformations that we perform are reversibility preserving and even. The preservation of these properties ensures that in the final system (1.40) the μ_j are real valued. Under this respect, the linear stability of the quasi-periodic standing wave solutions proved in Theorem 1.1 is obtained as a consequence of the reversible nature of the water waves equations. We could also preserve the Hamiltonian nature of $\mathcal{L}$ performing symplectic transformations, but it would be more complicated.

The above procedure – which we explain in detail below – is quite different from the approach developed in the pioneering works of Plotnikov-Toland [52] and Iooss-Plotnikov-Toland [41] for time periodic gravity waves. There are two main differences. The first one is that not all the transformations used in these works are periodically-time-dependent changes of variables acting in the phase space of functions on x, and therefore the dynamical system structure of the final conjugated system is lost. The second difference is that, when searching for time periodic solutions, it is sufficient to invert the linearized operator simply by a Neumann argument, as it is done in [1], [41], [37], [38], [52]. This approach does not work in the quasi-periodic case. The key difference is that, in the time 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.

We now explain in detail the steps for the conjugation of the quasi-periodic linear operator (1.38) to an operator of the form (1.40). We underline that all the coefficients of the linearized operator $\mathcal{L}$ in (1.38) 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 (16.1)) and the water waves vector field is analytic. This allows us to work in the usual framework of $\mathcal{C}^\infty$ pseudo-differential symbols, as recalled in Section 2.3.

1. *Linearized good unknown of Alinhac.* The first step is to introduce in Section 7.1 the linearized good unknown of Alinhac, as in [1] and [21]. This is indeed the same change of variable introduced by Lannes [45] for the local existence theory, see also [46] and Alazard-Metivier [5]. The outcome is the more symmetric system in (7.13)

$$\mathcal{L}_0 = \omega \cdot \partial_\varphi + \begin{pmatrix} \partial_x V & -G(\eta) \\ a & V\partial_x \end{pmatrix} = \omega \cdot \partial_\varphi + \begin{pmatrix} V\partial_x & 0 \\ 0 & V\partial_x \end{pmatrix} + \begin{pmatrix} V_x & -G(\eta) \\ a & 0 \end{pmatrix}, \tag{1.41}$$

where the Dirichlet-Neumann operator admits the expansion

$$G(\eta) = |D| \tanh(\mathtt{h}|D|) + \mathcal{R}_G$$

and $\mathcal{R}_G$ is an $OPS^{-\infty}$ smoothing operator. In Section 3 we provide a self-contained proof of such a representation, by transforming the elliptic problem (1.3), which is defined in the variable fluid domain $\{-\mathtt{h} \leq y \leq \eta(x)\}$, into the elliptic problem (3.45), which is defined on the straight strip $\{-\mathtt{h} - c \leq Y \leq 0\}$ and can be solved by an explicit integration.

2. *Straightening the first order vector field* $\omega \cdot \partial_\varphi + V(\varphi, x)\partial_x$. The next step is to conjugate the variable coefficients vector field (we regard equivalently a vector field as a differential operator)

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

to the constant coefficient vector field $\omega \cdot \partial_\varphi$ on the torus $\mathbb{T}^\nu_\varphi \times \mathbb{T}_x$ for $V(\varphi, x)$ small. This a perturbative problem of rectification of a close to constant vector field on a torus, which is a classical small divisor problem. For perturbation of a Diophantine vector field this problem was solved at the beginning of KAM theory, we refer e.g. to [62] and references therein. Notice that, despite the fact that $\omega \in \mathbb{R}^\nu$ is Diophantine, the constant vector field $\omega \cdot \partial_\varphi$ is resonant on the higher dimensional torus $\mathbb{T}^\nu_\varphi \times \mathbb{T}_x$. We exploit in a crucial way the reversibility property of $V(\varphi, x)$, i.e $V(\varphi, x)$ is odd in φ, to prove that it is possible to conjugate $\omega \cdot \partial_\varphi + V(\varphi, x)\partial_x$ to the constant vector field $\omega \cdot \partial_\varphi$ without changing the frequency ω.

From a functional point of view we have to solve a linear transport equation which depends on time in quasi-periodic way, see equation (8.5). Actually we solve equation (8.7) for the inverse diffeomorphism. This problem amounts to prove that all the solutions of the quasi periodically time-dependent scalar characteristic equation $\dot{x} = V(\omega t, x)$ are quasi-periodic in time with frequency ω, see Remark 8.1, [52], [41] and [51]. We solve this problem in Section 8 using a Nash-Moser implicit function theorem. Actually, after having inverted the linearized operator at an approximate solution (Lemma 8.2), we apply the Nash-Moser-Hörmander Theorem B.1, proved in Baldi-Haus [10]. The main advantage of this approach is to provide in Theorem 8.3 the optimal higher order regularity estimates (8.16) of the solution in terms of V.

Finally we remark that, when searching for time periodic solutions as in [41], [52], the corresponding transport equation is not a small-divisor problem and has been solved in [52] by a direct ODE analysis.

Applying this change of variable to the whole operator $\mathcal{L}_0$ in (1.41), the new conjugated system has the form, see (8.32),

$$\mathcal{L}_1 = \omega \cdot \partial_\varphi + \begin{pmatrix} a_1 & -a_2|D| \tanh(\mathtt{h}|D|) + \mathcal{R}_1 \\ a_3 & 0 \end{pmatrix}$$

where the remainder $\mathcal{R}_1$ is in $OPS^{-\infty}$.

3. *Change of the space variable.* In Section 9 we introduce a change of variable induced by a diffeomorphism of $\mathbb{T}_x$ of the form (independent of φ)

$$y = x + \alpha(x) \qquad \Leftrightarrow \qquad x = y + \breve{\alpha}(y) \, . \tag{1.42}$$

Conjugating $\mathcal{L}_1$ by the change of variable $u(x) \mapsto u(x + \alpha(x))$, we obtain an operator of the same form

$$\mathcal{L}_2 = \omega \cdot \partial_\varphi + \begin{pmatrix} a_4 & -a_5|D|T_\mathtt{h} + \mathcal{R}_2 \\ a_6 & 0 \end{pmatrix}, \qquad T_\mathtt{h} := \tanh(\mathtt{h}|D|) \, ,$$

see (9.5), where $\mathcal{R}_2$ is in $OPS^{-\infty}$, and the functions a_5, a_6 are given by

$$a_5 = \big[a_2(\varphi, x)(1 + \alpha_x(x))\big]_{|x = y + \breve{\alpha}(y)} \, , \quad a_6 = a_3(\varphi, y + \breve{\alpha}(y)) \, .$$

We shall choose in Section 12 the function $\alpha(x)$ in order to eliminate the space dependence from the highest order coefficients, see (12.25). The advantage to introduce at this step the diffeomorphism (1.42) is that it is easy to study the conjugation under this change of variable of differentiation and multiplication operators, Hilbert transform, and integral operators in $OPS^{-\infty}$, see Section 2.4.

4. *Symmetrization of the highest order.* In Section 10 we apply two simple conjugations (with a Fourier multiplier and a multiplication operator) whose goal is to obtain a new operator of the form

$$\mathcal{L}_3 = \omega \cdot \partial_\varphi + \begin{pmatrix} \breve{a}_4 & -a_7 |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} \\ a_7 |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} & 0 \end{pmatrix} + \dots ,$$

see (10.10)-(10.14), up to lower order operators. The function a_7 is close to 1 and $\breve{a}_4$ is small in ε, see (10.17). In the complex unknown $h = \eta + \mathrm{i}\psi$ the first component of such an operator reads

$$(h, \bar{h}) \mapsto \omega \cdot \partial_\varphi h + \mathrm{i} a_7 |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} h + a_8 h + P_5 h + Q_5 \bar{h}$$

(which corresponds to (11.1) neglecting the projector $\mathrm{i}\Pi_0$) where $P_5(\varphi)$ is a φ-dependent families of pseudo-differential operators of order $-1/2$, and $Q_5(\varphi)$ of order 0. We shall call the former operator "diagonal", and the latter "off-diagonal", with respect to the variables $(h, \bar{h})$.

5. *Symmetrization of the lower orders.* In Section 11 we reduce the off-diagonal term Q_5 to a pseudo-differential operator with very negative order, i.e. we conjugate the above operator to another one of the form (see Lemma 11.3)

$$(h, \bar{h}) \mapsto \omega \cdot \partial_\varphi h + \mathrm{i} a_7(\varphi, x) |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} h + a_8 h + P_6 h + Q_6 \bar{h} , \tag{1.43}$$

where P_6 is in $OPS^{-\frac{1}{2}}$ and $Q_6 \in OPS^{-M}$ for a constant M large enough fixed in Section 15, in view of the reducibility scheme.

6. *Time and space reduction at the highest order.* In Section 12, we eliminate the φ- and the x-dependence from the coefficient of the leading operator $\mathrm{i} a_7(\varphi, x) |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}}$. We conjugate the operator (1.43) by the time-1 flow of the pseudo-PDE

$$\partial_\tau u = \mathrm{i} \beta(\varphi, x) |D|^{\frac{1}{2}} u$$

where $\beta(\varphi, x)$ is a small function to be chosen. This kind of transformations – which are "semi-Fourier integral operators", namely pseudo-differential operators of type $(\frac{1}{2}, \frac{1}{2})$ in Hörmander's notation – has been introduced in [1] and studied as flows in [21].

Choosing appropriately the functions $\beta(\varphi, x)$ and $\alpha(x)$ (introduced in Section 9), see formulas (12.21) and (12.25), the final outcome is a linear operator of the form, see (12.33),

$$(h, \bar{h}) \mapsto \omega \cdot \partial_\varphi h + \mathrm{i} \mathtt{m}_{\frac{1}{2}} |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} h + (a_8 + a_9 \mathcal{H}) h + P_7 h + \mathcal{T}_7(h, \bar{h}) ,$$

where $\mathcal{H}$ is the Hilbert transform. This linear operator has the constant coefficient $\mathtt{m}_{\frac{1}{2}} \approx 1$ at the highest order, while P_7 is in $OPS^{-1/2}$ and the operator $\mathcal{T}_7$ is small, smoothing and satisfies tame estimates in Sobolev spaces, see (12.41). The constant $\mathtt{m}_{\frac{1}{2}}$ collects the quasi-linear effects of the non-linearity at the highest order.

7. *Reduction of the lower orders.* In Section 13 we further diagonalize the linear operator, reducing it to constant coefficients up to regularizing smoothing operators of very negative order $|D|^{-M}$. This is realized by applying an iterative sequence of pseudo-differential transformations that eliminate the φ- and the x-dependence of the diagonal symbols. The final system has the form

$$(h, \bar{h}) \mapsto \omega \cdot \partial_\varphi h + \mathrm{i} \mathtt{m}_{\frac{1}{2}} |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} h + \mathrm{i} r(D) h + \mathcal{R}_0(\varphi)(h, \bar{h}) \tag{1.44}$$

where the constant Fourier multiplier $r(\xi)$ is real, even $r(\xi) = r(-\xi)$, it satisfies (see (13.79))

$$\sup_{j \in \mathbb{Z}} |j|^{\frac{1}{2}} |r_j|^{k_0, \gamma} \lesssim_M \varepsilon \gamma^{-(2M+1)} ,$$

and the variable coefficient operator $\mathcal{R}_0(\varphi)$ is regularizing and satisfies tame estimates, see more precisely properties (1.45). We also remark that this final operator (1.44) is reversible and even, since all the previous transformations that we performed are reversibility preserving and even.

Our next goal is to diagonalize the operator (1.44); actually, it is sufficient to "almost-diagonalize" it by the KAM iterative scheme of Section 15. The expression "almost-diagonalize" refers to the fact that in Theorem 15.5 the remainders $\mathcal{R}_n$ that are left in (15.45) are not zero, but they are as small as $O(\varepsilon\gamma^{-2(M+1)}N_{n-1}^{-\mathtt{a}})$ (and this is because we only require the finitely many Diophantine conditions (15.44)).

8. *KAM-reducibility scheme.* In order to decrease quadratically the size of the perturbation $\mathcal{R}_0$ we apply the KAM diagonalization iterative scheme of Section 15 to the linear operator (1.44). Such a scheme converges because the operators

$$\langle D\rangle^{\mathtt{m}+\mathtt{b}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+\mathtt{b}+1}, \quad \partial_{\varphi_i}^{s_0+\mathtt{b}}\langle D\rangle^{\mathtt{m}+\mathtt{b}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+\mathtt{b}+1}, \ i=1,\ldots,\nu, \tag{1.45}$$

satisfy tame estimates for some $\mathtt{b} := \mathtt{b}(\tau, k_0) \in \mathbb{N}$ and $\mathtt{m} := \mathtt{m}(k_0)$ which are large enough (independently of s), fixed in (15.10), see precisely conditions (15.13)-(15.15). Such conditions are verified to hold in Lemma 15.3, under the assumption that M (the order of regularization of the remainder) is chosen large enough as in (15.16) (essentially $M = O(\mathtt{m}+\mathtt{b})$). This is the property that compensates, along the KAM iteration, the loss of derivatives in φ and x produced by the small divisors in the second order Melnikov non-resonance conditions.

The big difference of the KAM reducibility scheme of Section 15 with respect to the one developed in [21] is that the second order Melnikov non-resonance conditions that we impose are very weak, see (15.29), in particular they lose regularity, not only in the φ-variable, but also in the space variable x. For this reason we apply at each iterative step a smoothing procedure also in the space variable.

After the above diagonalization of the linearized operator we invert it, by imposing the first order Melnikov non-resonance conditions, see Lemma 15.11. Since all the changes of variables that we performed in the diagonalization process satisfy tame estimates in Sobolev spaces, we finally conclude the existence of an approximate inverse of the linearized operator which satisfies tame estimates, see Theorem 15.12.

Finally, in Section 16, we implement a differentiable Nash-Moser iterative scheme (Theorem 16.2) that provides an embedded torus which is invariant under the flow of the Hamiltonian vector field $X_{H_{\alpha_\infty(\omega,\mathtt{h})}}$ for most values of the parameters $(\omega,\mathtt{h})$. Section 16.1 concludes the proof of the Nash-Moser Theorem 5.1 of hypothetical conjugation.

1.2 Notation

We organize in this subsection the most important notation used in the paper.

We denote by $\mathbb{N} := \{0,1,2,\ldots\}$ the natural numbers including $\{0\}$ and $\mathbb{N}^+ := \{1,2,\ldots\}$. We denote the "tangential" sites by

$$\mathbb{S}^+ \subset \mathbb{N}^+ \quad \text{and we set} \quad \mathbb{S} := \mathbb{S}^+ \cup (-\mathbb{S}^+), \quad \mathbb{S}_0 := \mathbb{S}^+ \cup (-\mathbb{S}^+) \cup \{0\} \subseteq \mathbb{Z}, \quad \mathbb{S}_0^c := \mathbb{Z} \setminus \mathbb{S}_0 . \tag{1.46}$$

The cardinality of the set $\mathbb{S}^+$ is also denoted by $|\mathbb{S}^+| = \nu$. We look for quasi-periodic solutions with frequency $\omega \in \mathbb{R}^\nu$. The depth parameter $\mathtt{h}$ is in the interval $[\mathtt{h}_1, \mathtt{h}_2]$ with $\mathtt{h}_1 > 0$. In the paper all the functions, operators, transformations, etc ..., depend on the parameter

$$\lambda = (\omega, \mathtt{h}) \in \Lambda_0 \subset \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2],$$

in a k_0 times differentiable way, either in a classical or in a Whitney sense, as discussed in Section 2.1 and in Appendix A. We will often not specify the domain Λ_0 which is understood from the context. Given a set B we denote by $\mathcal{N}(B,\eta)$ the open neighborhood of B of width η (which is empty if B is empty) in $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, namely

$$\mathcal{N}(B,\eta) := \big\{\lambda \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] : \mathrm{dist}(B,\lambda) \le \eta\big\}. \tag{1.47}$$

We use the multi-index notation: if $k = (k_1,\ldots,k_{\nu+1}) \in \mathbb{N}^{\nu+1}$ and $\lambda = (\lambda_1,\ldots,\lambda_{\nu+1}) \in \mathbb{R}^{\nu+1}$, we denote the derivative $\partial_\lambda^k := \partial_{\lambda_1}^{k_1}\ldots\partial_{\lambda_{\nu+1}}^{k_{\nu+1}}$ and

$$|k| := k_1 + \ldots + k_{\nu+1}, \quad k! := k_1!\cdots k_{\nu+1}!, \quad \lambda^k := \lambda_1^{k_1}\cdots\lambda_{\nu+1}^{k_{\nu+1}}. \tag{1.48}$$

Given $j \in \mathbb{Z}$, we set $\langle j \rangle := \max\{1, |j|\}$ and for any vector $\ell = (\ell_1, \ldots, \ell_\nu) \in \mathbb{Z}^\nu$,

$$\langle \ell \rangle := \max\{1, |\ell|\}, \quad |\ell| = \max_{i=1,\ldots,\nu} |\ell_i| .$$

With a slight abuse of notation, given $\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}$, we write $\langle \ell, j \rangle := \max\{1, |\ell|, |j|\}$.

Sobolev spaces. We denote by $H^s(\mathbb{T}^{\nu+1})$ the Sobolev space of both real and complex valued functions defined by

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

see (1.20). In the paper we shall use H^s Sobolev spaces with index s in a finite range of values

$$s \in [s_0, S] , \quad \text{where} \quad s_0 := \Big[\frac{\nu + 1}{2} \Big] + 1 \in \mathbb{N} ,$$

see (1.21), and the value of S is fixed in the Nash-Moser iteration in Section 16, see (16.12).

We shall also use the notation $H^s_x := H^s(\mathbb{T}_x)$ for Sobolev spaces of functions of the space-variable $x \in \mathbb{T}$, and $H^s_\varphi = H^s(\mathbb{T}^\nu_\varphi)$ for Sobolev spaces of the periodic variable $\varphi \in \mathbb{T}^\nu$. Moreover we also define the subspace $H^1_0(\mathbb{T}_x)$ of $H^1(\mathbb{T}_x)$ of functions depending only on the space variable x with zero average, i.e.

$$H^1_0(\mathbb{T}_x) := \Big\{ u \in H^1(\mathbb{T}) : \int_{\mathbb{T}} u(x)\, dx = 0 \Big\} . \tag{1.49}$$

Given a function $u(\varphi, x)$ we write that it is even(φ)even(x), if it is even in φ for any x and, separately, even in x for any φ. With similar meaning we say that $u(\varphi, x)$ is even(φ)odd(x), odd(φ)even(x) and odd(φ)even(x).

Pseudo-differential operators and norms. A pseudo-differential operator with symbol $a(x, \xi)$ is denoted by $\mathrm{Op}(a)$ or $a(x, D)$, see Definition 2.8. The set of symbols $a(x, \xi)$ of order m is denoted by S^m and the class of the corresponding pseudo-differential operators by OPS^m. We also set

$$OPS^{-\infty} = \cap_{m \in \mathbb{R}} OPS^m .$$

We shall denote by OPS^m also matrix valued pseudo-differential operators with entries in OPS^m.

Along the paper we have to consider symbols $a(\lambda, \varphi, x, \xi)$ that depend on $\varphi \in \mathbb{T}^\nu$ and on a parameter $\lambda \in \Lambda_0 \subset \mathbb{R}^{\nu+1}$. The symbol a is k_0 times differentiable with respect to λ and $\mathcal{C}^\infty$ with respect to (φ, x, ξ). For the corresponding family of pseudo-differential operators $A(\lambda) = a(\lambda, \varphi, x, D)$ we introduce in Definition 2.9 the norms

$$|\!|\!| A |\!|\!|^{k_0,\gamma}_{m,s,\alpha} := \sum_{|k| \le k_0} \gamma^{|k|} \sup_{\lambda \in \Lambda_0} |\!|\!| \partial^k_\lambda A(\lambda) |\!|\!|_{m,s,\alpha} \tag{1.50}$$

indexed by $k_0 \in \mathbb{N}$, $\gamma \in (0, 1)$, $m \in \mathbb{R}$, $s \ge s_0$, $\alpha \in \mathbb{N}$, where

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

$\mathcal{D}^{k_0}$-tame and $\mathcal{D}^{k_0}$-modulo-tame operators. In Definition 2.25 we introduce the class of linear operators $A = A(\lambda)$ satisfying tame estimates of the form

$$\sup_{|k| \le k_0} \sup_{\lambda \in \Lambda_0} \gamma^{|k|} \|(\partial^k_\lambda A(\lambda)) u\|_s \le \mathfrak{M}_A(s_0) \|u\|_{s+\sigma} + \mathfrak{M}_A(s) \|u\|_{s_0+\sigma} ,$$

which we call $\mathcal{D}^{k_0}$-σ-tame operators. The constant $\mathfrak{M}_A(s)$ is called the tame constant of the operator A. When the "loss of derivatives" σ is zero, we simply write $\mathcal{D}^{k_0}$-tame instead of $\mathcal{D}^{k_0}$-0-tame.

In Definition 2.30 we introduce the subclass of $\mathcal{D}^{k_0}$-modulo tame operators $A = A(\lambda)$ such that for any $k \in \mathbb{N}^{\nu+1}$, $|k| \le k_0$, the majorant operator $|\partial^k_\lambda A|$ satisfies the tame estimates

$$\sup_{|k| \le k_0} \sup_{\lambda \in \Lambda_0} \gamma^{|k|} \||\partial^k_\lambda A| u\|_s \le \mathfrak{M}^\sharp_A(s_0) \|u\|_s + \mathfrak{M}^\sharp_A(s) \|u\|_{s_0} .$$

The majorant operator $|A|$ is introduced in Definition 2.7-1, by taking the modulus of the entries of the matrix which represents the operator A with respect to the exponential basis. We refer to $\mathfrak{M}^\sharp_A(s)$ as the modulo tame constant of the operator A.

Along the paper several functions, symbols and operators will depend on the torus embedding $\varphi \mapsto i(\varphi)$ (the point at which we linearize the nonlinear equation) and we shall use the notation

$$\Delta_{12} u := u(i_2) - u(i_1)$$

to denote the increment of such quantities with respect to i.

Finally we use the following notation: $a \lesssim_{s,\alpha,M} b$ means that $a \le C(s,\alpha,M) b$ for some constant $C(s,\alpha,M) > 0$ depending on the Sobolev index s and the constants α, M. Sometimes, along the paper, we omit to write the dependence $\lesssim_{s_0,k_0}$ with respect to s_0, k_0, because s_0 (defined in (1.21)) and k_0 (determined in Section 4) are considered as fixed constants. Similarly, the set $\mathbb{S}^+$ of tangential sites and its cardinality $\nu = |\mathbb{S}^+|$ are also considered as fixed along the paper.

2 Functional setting

2.1 Function spaces

In the paper we will use Sobolev norms for real or complex functions $u(\omega, \mathtt{h}, \varphi, x)$, $(\varphi, x) \in \mathbb{T}^\nu \times \mathbb{T}$, depending on parameters $(\omega, \mathtt{h}) \in F$ in a Lipschitz way together with their derivatives in the sense of Whitney, where F is a closed subset of $\mathbb{R}^{\nu+1}$. We use the compact notation $\lambda := (\omega, \mathtt{h})$ to collect the frequency ω and the depth $\mathtt{h}$ into a parameter vector.

Also recall that $\| \ \|_s$ denotes the norm of the Sobolev space $H^s(\mathbb{T}^{\nu+1}, \mathbb{C}) = H^s_{(\varphi,x)}$ introduced in (1.20). We now define the "Whitney-Sobolev" norm $\| \cdot \|^{k+1,\gamma}_{s,F}$.

Definition 2.1. (Whitney-Sobolev functions) *Let F be a closed subset of $\mathbb{R}^{\nu+1}$. Let $k \ge 0$ be an integer, $\gamma \in (0,1]$, and $s \ge s_0 > (\nu+1)/2$. We say that a function $u : F \to H^s_{(\varphi,x)}$ belongs to $\mathrm{Lip}(k+1, F, s, \gamma)$ if there exist functions*

$$u^{(j)} : F \to H^s_{(\varphi,x)}, \quad j \in \mathbb{N}^\nu, \ 0 \le |j| \le k,$$

with $u^{(0)} = u$, and a constant $M > 0$ such that, if $R_j(\lambda, \lambda_0) := R_j^{(u)}(\lambda, \lambda_0)$ is defined by

$$u^{(j)}(\lambda) = \sum_{\ell \in \mathbb{N}^{\nu+1} : |j+\ell| \le k} \frac{1}{\ell!} u^{(j+\ell)}(\lambda_0) \, (\lambda - \lambda_0)^\ell + R_j(\lambda, \lambda_0), \quad \lambda, \lambda_0 \in F, \tag{2.1}$$

(recall the multi-index notation (1.48)) then

$$\gamma^{|j|} \| u^{(j)}(\lambda) \|_s \le M, \quad \gamma^{k+1} \| R_j(\lambda, \lambda_0) \|_s \le M |\lambda - \lambda_0|^{k+1-|j|} \quad \forall \lambda, \lambda_0 \in F, \ |j| \le k. \tag{2.2}$$

An element of $\mathrm{Lip}(k+1, F, s, \gamma)$ is in fact the collection $\{u^{(j)} : |j| \le k\}$. The norm of $u \in \mathrm{Lip}(k+1, F, s, \gamma)$ is defined as

$$\|u\|^{k+1,\gamma}_{s,F} := \|u\|^{k+1,\gamma}_s := \inf\{M > 0 : (2.2) \text{ holds}\}. \tag{2.3}$$

If $F = \mathbb{R}^{\nu+1}$ by $\mathrm{Lip}(k+1, \mathbb{R}^{\nu+1}, s, \gamma)$ we shall mean the space of the functions $u = u^{(0)}$ for which there exist $u^{(j)} = \partial_\lambda^j u$, $|j| \le k$, satisfying (2.2), with the same norm (2.3).

We make some remarks.

1. If $F = \mathbb{R}^{\nu+1}$, and $u \in \mathrm{Lip}(k+1, F, s, \gamma)$ the $u^{(j)}$, $|j| \ge 1$, are uniquely determined as the partial derivatives $u^{(j)} = \partial_\lambda^j u$, $|j| \le k$, of $u = u^{(0)}$. Moreover all the derivatives $\partial_\lambda^j u$, $|j| = k$ are Lipschitz. Since H^s is a Hilbert space we have that $\mathrm{Lip}(k+1, \mathbb{R}^{\nu+1}, s, \gamma)$ coincides with the Sobolev space $W^{k+1,\infty}(\mathbb{R}^{\nu+1}, H^s)$.

2. The Whitney-Sobolev norm of u in (2.3) is equivalently given by

$$\|u\|_{s,F}^{k+1,\gamma} := \|u\|_s^{k+1,\gamma} = \max_{|j|\le k}\Big\{\gamma^{|j|}\sup_{\lambda\in F}\|u^{(j)}(\lambda)\|_s, \gamma^{k+1}\sup_{\lambda\neq\lambda_0}\frac{\|R_j(\lambda,\lambda_0)\|_s}{|\lambda-\lambda_0|^{k+1-|j|}}\Big\}. \qquad (2.4)$$

3. The exponent of γ in (2.2) gives the number of "derivatives" of u that are involved in the Taylor expansion (taking into account that in the remainder there is one derivative more than in the Taylor polynomial); on the other hand the exponent of $|\lambda-\lambda_0|$ gives the order of the Taylor expansion of $u^{(j)}$ with respect to λ. This is the reason for the difference of $|j|$ between the two exponents. The factor γ is normalized by the rescaling (A.7).

Theorem A.2 and (A.10) provide an extension operator which associates to an element $u \in \mathrm{Lip}(k+1, F, s, \gamma)$ an extension $\tilde u \in \mathrm{Lip}(k+1, \mathbb{R}^{\nu+1}, s, \gamma)$. As already observed, the space $\mathrm{Lip}(k+1, \mathbb{R}^{\nu+1}, s, \gamma)$ coincides with $W^{k+1,\infty}(\mathbb{R}^{\nu+1}, H^s)$, with equivalence of the norms (see (A.9))

$$\|u\|_{s,F}^{k+1,\gamma} \sim_{\nu,k} \|\tilde u\|_{W^{k+1,\infty,\gamma}(\mathbb{R}^{\nu+1},H^s)} := \sum_{|\alpha|\le k+1}\gamma^{|\alpha|}\|\partial_\lambda^\alpha \tilde u\|_{L^\infty(\mathbb{R}^{\nu+1},H^s)}\,.$$

By Lemma A.3, the extension $\tilde u$ is independent of the Sobolev space H^s.

We can identify any element $u \in \mathrm{Lip}(k+1, F, s, \gamma)$ (which is a collection $u = \{u^{(j)} : |j| \le k\}$) with the equivalence class of functions $f \in W^{k+1,\infty}(\mathbb{R}^{\nu+1}, H^s)/\!\sim$ with respect to the equivalence relation $f \sim g$ when $\partial_\lambda^j f(\lambda) = \partial_\lambda^j g(\lambda)$ for all $\lambda \in F$, for all $|j| \le k+1$.

For any $N > 0$, we introduce the smoothing operators

$$(\Pi_N u)(\varphi, x) := \sum_{\langle \ell,j\rangle \le N} u_{\ell j} e^{\mathrm{i}(\ell\cdot\varphi + jx)} \qquad \Pi_N^\perp := \mathrm{Id} - \Pi_N. \qquad (2.5)$$

Lemma 2.2. (Smoothing) *Consider the space* $\mathrm{Lip}(k+1, F, s, \gamma)$ *defined in Definition 2.1. The smoothing operators* $\Pi_N, \Pi_N^\perp$ *satisfy the estimates*

$$\|\Pi_N u\|_s^{k+1,\gamma} \le N^\alpha \|u\|_{s-\alpha}^{k+1,\gamma}, \quad 0 \le \alpha \le s, \qquad (2.6)$$

$$\|\Pi_N^\perp u\|_s^{k+1,\gamma} \le N^{-\alpha} \|u\|_{s+\alpha}^{k+1,\gamma}, \quad \alpha \ge 0. \qquad (2.7)$$

Proof. See Appendix A. □

Lemma 2.3. (Interpolation) *Consider the space* $\mathrm{Lip}(k+1, F, s, \gamma)$ *defined in Definition 2.1.*

(i) Let $s_1 < s_2$. *Then for any* $\theta \in (0,1)$ *one has*

$$\|u\|_s^{k+1,\gamma} \le (\|u\|_{s_1}^{k+1,\gamma})^\theta (\|u\|_{s_2}^{k+1,\gamma})^{1-\theta}, \quad s := \theta s_1 + (1-\theta)s_2\,. \qquad (2.8)$$

(ii) 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*

$$\|u\|_{a_0+p}^{k+1,\gamma}\|v\|_{b_0+q}^{k+1,\gamma} \le \epsilon\|u\|_{a_0+p+q}^{k+1,\gamma}\|v\|_{b_0}^{k+1,\gamma} + C(\epsilon)\|u\|_{a_0}^{k+1,\gamma}\|v\|_{b_0+p+q}^{k+1,\gamma}\,. \qquad (2.9)$$

Proof. See Appendix A. □

Lemma 2.4. (Product and composition) *Consider the space* $\mathrm{Lip}(k+1, F, s, \gamma)$ *defined in Definition 2.1. For all* $s \ge s_0 > (\nu+1)/2$, *we have*

$$\|uv\|_s^{k+1,\gamma} \le C(s,k)\|u\|_s^{k+1,\gamma}\|v\|_{s_0}^{k+1,\gamma} + C(s_0,k)\|u\|_{s_0}^{k+1,\gamma}\|v\|_s^{k+1,\gamma}\,. \qquad (2.10)$$

Let $\|\beta\|_{2s_0+1}^{k+1,\gamma} \le \delta(s_0,k)$ *small enough. Then the composition operator*

$$\mathcal{B} : u \mapsto \mathcal{B}u, \quad (\mathcal{B}u)(\varphi, x) := u(\varphi, x + \beta(\varphi, x))\,,$$

satisfies the following tame estimates: for all $s \geq s_0$,

$$\|\mathcal{B}u\|_s^{k+1,\gamma} \lesssim_{s,k} \|u\|_{s+k+1}^{k+1,\gamma} + \|\beta\|_s^{k+1,\gamma}\|u\|_{s_0+k+2}^{k+1,\gamma} . \tag{2.11}$$

Let $\|\beta\|_{2s_0+k+2}^{k+1,\gamma} \leq \delta(s_0,k)$ small enough. The function $\breve{\beta}$ defined by the inverse diffeomorphism $y = x+\beta(\varphi,x)$ if and only if $x = y + \breve{\beta}(\varphi,y)$, satisfies

$$\|\breve{\beta}\|_s^{k+1,\gamma} \lesssim_{s,k} \|\beta\|_{s+k+1}^{k+1,\gamma} . \tag{2.12}$$

Proof. See Appendix A. □

If ω belongs to the set of Diophantine vectors $\mathtt{DC}(\gamma,\tau)$, where

$$\mathtt{DC}(\gamma,\tau) := \Big\{\omega \in \mathbb{R}^\nu : |\omega\cdot\ell| \geq \frac{\gamma}{|\ell|^\tau} \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\Big\} , \tag{2.13}$$

the equation $\omega\cdot\partial_\varphi v = u$, where $u(\varphi,x)$ has zero average with respect to φ, has the periodic solution

$$(\omega\cdot\partial_\varphi)^{-1}u := \sum_{\ell\in\mathbb{Z}^\nu\setminus\{0\}, j\in\mathbb{Z}} \frac{u_{\ell,j}}{\mathrm{i}\omega\cdot\ell} e^{\mathrm{i}(\ell\cdot\varphi+jx)} . \tag{2.14}$$

For all $\omega \in \mathbb{R}^\nu$ we define its extension

$$(\omega\cdot\partial_\varphi)_{ext}^{-1}u(\varphi,x) := \sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} \frac{\chi(\omega\cdot\ell\gamma^{-1}\langle\ell\rangle^\tau)}{\mathrm{i}\omega\cdot\ell}\, u_{\ell,j}\, e^{\mathrm{i}(\ell\cdot\varphi+jx)}, \tag{2.15}$$

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

$$\chi(\xi) = \begin{cases} 0 & \text{if} \quad |\xi| \leq \frac{1}{3} \\ 1 & \text{if} \quad |\xi| \geq \frac{2}{3} , \end{cases} \qquad \partial_\xi\chi(\xi) > 0 \quad \forall \xi \in \Big(\frac{1}{3},\frac{2}{3}\Big) . \tag{2.16}$$

Note that $(\omega\cdot\partial_\varphi)_{ext}^{-1}u = (\omega\cdot\partial_\varphi)^{-1}u$ for all $\omega \in \mathtt{DC}(\gamma,\tau)$.

Lemma 2.5. (Diophantine equation) *For all $u \in W^{k+1,\infty,\gamma}(\mathbb{R}^{\nu+1}, H^{s+\mu})$, we have*

$$\|(\omega\cdot\partial_\varphi)_{ext}^{-1}u\|_{s,\mathbb{R}^{\nu+1}}^{k+1,\gamma} \leq C(k)\gamma^{-1}\|u\|_{s+\mu,\mathbb{R}^{\nu+1}}^{k+1,\gamma}, \qquad \mu := k+1+\tau(k+2). \tag{2.17}$$

Moreover, for $F \subseteq \mathtt{DC}(\gamma,\tau)\times\mathbb{R}$ one has

$$\|(\omega\cdot\partial_\varphi)^{-1}u\|_{s,F}^{k+1,\gamma} \leq C(k)\gamma^{-1}\|u\|_{s+\mu,F}^{k+1,\gamma} . \tag{2.18}$$

Proof. See Appendix A. □

We finally state a standard Moser tame estimate for the nonlinear composition operator

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

Since the variables $(\varphi,x) := y$ have the same role, we state it for a generic Sobolev space $H^s(\mathbb{T}^d)$.

Lemma 2.6. (Composition operator) *Let $f \in \mathcal{C}^\infty(\mathbb{T}^d\times\mathbb{R},\mathbb{C})$ and $C_0 > 0$. Consider the space* Lip$(k+1,F,s,\gamma)$ *given in Definition 2.1. If $u(\lambda) \in H^s(\mathbb{T}^d,\mathbb{R})$, $\lambda \in F$ is a family of Sobolev functions satisfying $\|u\|_{s_0,F}^{k+1,\gamma} \leq C_0$, then, for all $s \geq s_0 > (d+1)/2$,*

$$\|\mathtt{f}(u)\|_{s,F}^{k+1,\gamma} \leq C(s,k,f,C_0)(1+\|u\|_{s,F}^{k+1,\gamma}) . \tag{2.19}$$

The constant $C(s,k,f,C_0)$ depends on s, k and linearly on $\|f\|_{\mathcal{C}^m(\mathbb{T}^d\times B)}$, where m is an integer larger than $s+k+1$, and $B \subset \mathbb{R}$ is a bounded interval such that $u(\lambda,y) \in B$ for all $\lambda \in F$, $y \in \mathbb{T}^d$, for all $\|u\|_{s_0,F}^{k+1,\gamma} \leq C_0$.

Proof. See Appendix A. □

2.2 Linear operators

Along the paper we consider φ-dependent families of linear operators $A : \mathbb{T}^\nu \mapsto \mathcal{L}(L^2(\mathbb{T}_x))$, $\varphi \mapsto A(\varphi)$ acting on functions $u(x)$ of the space variable x, i.e. on subspaces of $L^2(\mathbb{T}_x)$, either real or complex valued. We also regard A as an operator (which for simplicity we denote by A as well) that acts on functions $u(\varphi, x)$ of space-time, i.e. we consider the corresponding 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

$$\mathcal{R} \begin{pmatrix} \eta \\ \psi \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \eta \\ \psi \end{pmatrix} \tag{2.20}$$

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 scalar function $u := u(\varphi, x) \in L^2(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{C})$ that we expand in Fourier series as

$$u(\varphi, x) = \sum_{j \in \mathbb{Z}} u_j(\varphi) e^{\mathrm{i} jx} = \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}} u_{\ell, j} e^{\mathrm{i}(\ell \cdot \varphi + jx)} \tag{2.21}$$

is

$$Au(\varphi, x) = \sum_{j, j' \in \mathbb{Z}} A_j^{j'}(\varphi) u_{j'}(\varphi) e^{\mathrm{i} jx} = \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}} \ \sum_{\ell' \in \mathbb{Z}^\nu, j' \in \mathbb{Z}} A_j^{j'}(\ell - \ell') u_{\ell', j'} e^{\mathrm{i}(\ell \cdot \varphi + jx)} \, . \tag{2.22}$$

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 $\partial_{\varphi_m} A(\varphi)$, $m = 1, \ldots, \nu$, is represented by the matrix with 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 $\mathrm{i}(j - j') A_j^{j'}(\ell - \ell')$.

Also note that the operator norm $\|A\|_{\mathcal{L}(H^s)} := \sup\{\|Ah\|_s : \|h\|_s = 1\}$ of a bounded operator $A : H^s \to H^s$ and its matrix entries $A_j^{j'}(\ell - \ell')$ satisfy

$$\sum_{\ell, j} |A_j^{j'}(\ell - \ell')|^2 \langle \ell, j \rangle^{2s} \leq \|A\|^2_{\mathcal{L}(H^s)} \langle \ell', j' \rangle^{2s} \, , \quad \forall (\ell', j') \in \mathbb{Z}^{\nu+1}. \tag{2.23}$$

To prove (2.23), consider $h = e^{\mathrm{i}(\ell', j') \cdot (\varphi, x)}$.

Definition 2.7. *Given a linear operator A as in* (2.22) *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*

$$(\Pi_N A)_j^{j'}(\ell - \ell') := \begin{cases} A_j^{j'}(\ell - \ell') & \text{if } \ \langle \ell - \ell', j - j' \rangle \leq N \\ 0 & \text{otherwise} \, . \end{cases} \tag{2.24}$$

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

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

Given linear operators A, B we have that (see Lemma 2.4 in [21])

$$\| \, |A + B| u \|_s \leq \| \, |A| \, |\!|u|\!| \, \|_s + \| \, |B| \, |\!|u|\!| \, \|_s \, , \quad \| \, |AB| u \|_s \leq \| \, |A||B| \, |\!|u|\!| \, \|_s \, , \tag{2.25}$$

where, for a given a function $u(\varphi, x)$ expanded in Fourier series as in (2.21), we define the majorant function

$$|\!|u|\!|(\varphi, x) := \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}} |u_{\ell, j}| e^{\mathrm{i}(\ell \cdot \varphi + jx)} \, . \tag{2.26}$$

Note that the Sobolev norms of u and $|\!|u|\!|$ are the same, i.e.

$$\|u\|_s = \| |\!|u|\!| \|_s. \tag{2.27}$$

2.3 Pseudo-differential operators

In this section we recall the main properties of pseudo-differential operators on the torus that we shall use in the paper, similarly to [1], [21]. Pseudo-differential operators on the torus may be seen as a particular case of the theory on $\mathbb{R}^n$, as developed for example in [34]. They can also be directly expressed through Fourier series, for which we refer to [58].

Definition 2.8. (ΨDO) *A linear operator A is called a* pseudo-differential *operator of order 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*

$$\big|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)\big| \leq C_{\alpha,\beta}\langle\xi\rangle^{m-\beta}\,, \quad \forall \alpha,\beta\in\mathbb{N}\,. \tag{2.28}$$

We call $a(x,\xi)$ the symbol of the operator A, which 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 (2.28)*, and by OPS^m the associated set of pseudo-differential operators of order m. We set $OPS^{-\infty} := \cap_{m\in\mathbb{R}} OPS^m$.*

For a matrix of pseudo differential operators

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, \quad A_i \in OPS^m, \quad i = 1,\dots,4 \tag{2.29}$$

we say that $A\in OPS^m$.

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

We underline that we regard any operator $\mathrm{Op}(a)$ as an operator acting only on 2π-periodic functions $u(x)=\sum_{j\in\mathbb{Z}} u_j e^{\mathrm{i}jx}$ as

$$(Au)(x) := \mathrm{Op}(a)[u](x) := \sum_{j\in\mathbb{Z}} a(x,j) u_j e^{\mathrm{i}jx}\,.$$

We recall some fundamental properties of pseudo-differential operators.

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

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

where $\widehat{\ }$ 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

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

that is, for all $N\geq 1$,

$$\sigma_{AB}(x,\xi) = \sum_{\beta=0}^{N-1} \frac{1}{\mathrm{i}^\beta\beta!}\partial_\xi^\beta a(x,\xi)\,\partial_x^\beta b(x,\xi) + r_N(x,\xi) \qquad \text{where} \qquad r_N := r_{N,AB}\in S^{m+m'-N}\,. \tag{2.30}$$

The remainder r_N has the explicit formula

$$r_N(x,\xi) := r_{N,AB}(x,\xi) := \frac{1}{\mathrm{i}^N(N-1)!}\int_0^1 (1-\tau)^{N-1}\sum_{j\in\mathbb{Z}} (\partial_\xi^N a)(x,\xi+\tau j)\widehat{(\partial_x^N b)}(j,\xi)e^{\mathrm{i}jx}\,d\tau\,. \tag{2.31}$$

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

$$A^* = \mathrm{Op}(a^*) \qquad \text{with symbol} \qquad a^*(x, \xi) := \overline{\sum_{j \in \mathbb{Z}} \widehat{a}(j, \xi - j) e^{\mathrm{i} jx}} \,. \tag{2.32}$$

Along the paper we consider φ-dependent families of pseudo-differential operators

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

where the symbol $a(\varphi, x, \xi)$ is $\mathcal{C}^\infty$-smooth also in φ. We still denote $A := A(\varphi) = \mathrm{Op}(a(\varphi, \cdot)) = \mathrm{Op}(a)$. Moreover we consider pseudo-differential operators $A(\lambda) := \mathrm{Op}(a(\lambda, \varphi, x, \xi))$ that are k_0 times differentiable with respect to a parameter $\lambda := (\omega, \mathtt{h})$ in an open subset $\Lambda_0 \subseteq \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$. The regularity constant $k_0 \in \mathbb{N}$ is fixed once and for all in Section 4. Note that

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

We shall use the following notation, used also in [1], [21]. For any $m \in \mathbb{R} \setminus \{0\}$, we set

$$|D|^m := \mathrm{Op}\big(\chi(\xi)|\xi|^m\big) \,, \tag{2.33}$$

where χ is the even, positive $\mathcal{C}^\infty$ cut-off defined in (2.16). We also identify the Hilbert transform $\mathcal{H}$, acting on the 2π-periodic functions, defined by

$$\mathcal{H}(e^{\mathrm{i} jx}) := -\mathrm{i}\,\mathrm{sign}(j) e^{\mathrm{i} jx} \,, \ \forall j \neq 0 \,, \quad \mathcal{H}(1) := 0 \,, \tag{2.34}$$

with the Fourier multiplier $\mathrm{Op}\big(-\mathrm{i}\,\mathrm{sign}(\xi)\chi(\xi)\big)$, i.e.

$$\mathcal{H} \equiv \mathrm{Op}\big(-\mathrm{i}\,\mathrm{sign}(\xi)\chi(\xi)\big) \,.$$

We shall identify the projector π_0, defined on the 2π-periodic functions as

$$\pi_0 u := \frac{1}{2\pi} \int_{\mathbb{T}} u(x) \, dx \,, \tag{2.35}$$

with the Fourier multiplier $\mathrm{Op}\big(1 - \chi(\xi)\big)$, i.e.

$$\pi_0 \equiv \mathrm{Op}\big(1 - \chi(\xi)\big) \,,$$

where the cut-off $\chi(\xi)$ is defined in (2.16). We also define the Fourier multiplier $\langle D \rangle^m$, $m \in \mathbb{R} \setminus \{0\}$, as

$$\langle D \rangle^m := \pi_0 + |D|^m := \mathrm{Op}\left((1 - \chi(\xi)) + \chi(\xi)|\xi|^m\right), \quad \xi \in \mathbb{R} \,. \tag{2.36}$$

We now recall the pseudo-differential norm introduced in Definition 2.11 in [21] (inspired by Métivier [49], 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 2.9. (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*

$$|A|^{k_0, \gamma}_{m, s, \alpha} := \sum_{|k| \leq k_0} \gamma^{|k|} \sup_{\lambda \in \Lambda_0} |\partial_\lambda^k A(\lambda)|_{m, s, \alpha} \tag{2.37}$$

where

$$|A(\lambda)|_{m, s, \alpha} := \max_{0 \leq \beta \leq \alpha} \sup_{\xi \in \mathbb{R}} \|\partial_\xi^\beta a(\lambda, \cdot, \cdot, \xi)\|_s \langle \xi \rangle^{-m+\beta} \,. \tag{2.38}$$

For a matrix of pseudo differential operators $A \in OPS^m$ as in (2.29), we define its pseudo differential norm

$$|A|^{k_0, \gamma}_{m, s, \alpha} := \max_{i = 1, \ldots, 4} |A_i|^{k_0, \gamma}_{m, s, \alpha} \,.$$

For each k_0, γ, m fixed, the norm (2.37) is non-decreasing both in s and α, namely

$$\forall s \leq s', \, \alpha \leq \alpha' , \qquad | \, |_{m,s,\alpha}^{k_0,\gamma} \leq | \, |_{m,s',\alpha}^{k_0,\gamma} , \quad | \, |_{m,s,\alpha}^{k_0,\gamma} \leq | \, |_{m,s,\alpha'}^{k_0,\gamma} , \tag{2.39}$$

and it is non-increasing in m, i.e.

$$\forall m \leq m', \qquad | \, |_{m',s,\alpha}^{k_0,\gamma} \leq | \, |_{m,s,\alpha}^{k_0,\gamma} . \tag{2.40}$$

Given a function $a(\lambda, \varphi, x)$ that is $\mathcal{C}^\infty$ in (φ, x) and k_0 times differentiable in λ, the "weighted ΨDO norm" of the corresponding multiplication operator $\mathrm{Op}\,(a)$ is

$$|\mathrm{Op}\,(a)|_{0,s,\alpha}^{k_0,\gamma} = \sum_{|k| \leq k_0} \gamma^{|k|} \sup_{\lambda \in \Lambda_0} \|\partial_\lambda^k a(\lambda)\|_s = \|a\|_{W^{k_0,\infty,\gamma}(\Lambda_0, H^s)} \sim_{k_0} \|a\|_s^{k_0,\gamma} , \quad \forall \alpha \in \mathbb{N} , \tag{2.41}$$

see (A.9). For a Fourier multiplier $g(\lambda, D)$ with symbol $g \in S^m$, we simply have

$$|\mathrm{Op}(g)|_{m,s,\alpha}^{k_0,\gamma} = |\mathrm{Op}(g)|_{m,0,\alpha}^{k_0,\gamma} \leq C(m, \alpha, g, k_0) , \quad \forall s \geq 0 . \tag{2.42}$$

Given a symbol $a(\lambda, \varphi, x, \xi) \in S^m$, we define its averages

$$\langle a \rangle_\varphi(\lambda, x, \xi) := \frac{1}{(2\pi)^\nu} \int_{\mathbb{T}^\nu} a(\lambda, \varphi, x, \xi) \, d\varphi \;, \quad \langle a \rangle_{\varphi,x}(\lambda, \xi) := \frac{1}{(2\pi)^{\nu+1}} \int_{\mathbb{T}^{\nu+1}} a(\lambda, \varphi, x, \xi) \, d\varphi \, dx \, .$$

One has that $\langle a \rangle_\varphi$ and $\langle a \rangle_{\varphi,x}$ are symbols in S^m that satisfy

$$|\mathrm{Op}(\langle a \rangle_\varphi)|_{m,s,\alpha}^{k_0,\gamma} \lesssim |\mathrm{Op}(a)|_{m,s,\alpha}^{k_0,\gamma} \qquad \forall s \geq 0 , \tag{2.43}$$

$$|\mathrm{Op}(\langle a \rangle_{\varphi,x})|_{m,s,\alpha}^{k_0,\gamma} \lesssim |\mathrm{Op}(a)|_{m,0,\alpha}^{k_0,\gamma} \quad \forall s \geq 0 . \tag{2.44}$$

The norm $| \, |_{0,s,0}$ controls the action of a pseudo-differential operator on the Sobolev spaces H^s, see Lemma 2.29. The norm $| \, |_{m,s,\alpha}^{k_0,\gamma}$ is closed under composition and satisfies tame estimates.

Lemma 2.10. (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 $\alpha \in \mathbb{N}$, $s \geq s_0$,*

$$|AB|_{m+m',s,\alpha}^{k_0,\gamma} \lesssim_{m,\alpha,k_0} C(s)|A|_{m,s,\alpha}^{k_0,\gamma}|B|_{m',s_0+\alpha+|m|,\alpha}^{k_0,\gamma} + C(s_0)|A|_{m,s_0,\alpha}^{k_0,\gamma}|B|_{m',s+\alpha+|m|,\alpha}^{k_0,\gamma} . \tag{2.45}$$

Moreover, for any integer $N \geq 1$, the remainder $R_N := \mathrm{Op}(r_N)$ in (2.30) satisfies

$$\begin{aligned} |R_N|_{m+m'-N,s,\alpha}^{k_0,\gamma} \lesssim_{m,N,\alpha,k_0} & C(s)|A|_{m,s,N+\alpha}^{k_0,\gamma}|B|_{m',s_0+2N+|m|+\alpha,\alpha}^{k_0,\gamma} \\ & + C(s_0)|A|_{m,s_0,N+\alpha}^{k_0,\gamma}|B|_{m',s+2N+|m|+\alpha,\alpha}^{k_0,\gamma} . \end{aligned} \tag{2.46}$$

Both (2.45)-(2.46) hold with the constant $C(s_0)$ interchanged with $C(s)$.

Analogous estimates hold if A and B are matrix operators of the form (2.29).

Proof. See Lemma 2.13 in [21]. □

For a Fourier multiplier $g(\lambda, D)$ with symbol $g \in S^{m'}$ we have the simpler estimate

$$|A \circ g(D)|_{m+m',s,\alpha}^{k_0,\gamma} \lesssim_{k_0,\alpha} |A|_{m,s,\alpha}^{k_0,\gamma}|\mathrm{Op}(g)|_{m',0,\alpha}^{k_0,\gamma} \lesssim_{k_0,\alpha,m'} |A|_{m,s,\alpha}^{k_0,\gamma} . \tag{2.47}$$

By (2.30) 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 $a \star b$, namely

$$[A, B] = \mathrm{Op}(a \star b) . \tag{2.48}$$

By (2.30) the symbol $a \star b \in S^{m+m'-1}$ admits the expansion

$$a \star b = -\mathrm{i}\{a, b\} + \mathtt{r}_2(a, b) \qquad \text{where} \qquad \{a, b\} := \partial_\xi a \, \partial_x b - \partial_x a \, \partial_\xi b \in S^{m+m'-1} \tag{2.49}$$

is the Poisson bracket between $a(x, \xi)$ and $b(x, \xi)$, and

$$\mathtt{r}_2(a, b) := r_{2,AB} - r_{2,BA} \in S^{m+m'-2} . \tag{2.50}$$

By Lemma 2.10 we deduce the following corollary.

Lemma 2.11. (Commutator) *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*

$$\begin{aligned} |[A,B]|^{k_0,\gamma}_{m+m'-1,s,\alpha} \lesssim_{m,m',\alpha,k_0} & C(s) |A|^{k_0,\gamma}_{m,s+2+|m'|+\alpha,\alpha+1} |B|^{k_0,\gamma}_{m',s_0+2+|m|+\alpha,\alpha+1} \\ & + C(s_0) |A|^{k_0,\gamma}_{m,s_0+2+|m'|+\alpha,\alpha+1} |B|^{k_0,\gamma}_{m',s+2+|m|+\alpha,\alpha+1} . \end{aligned} \tag{2.51}$$

Proof. Use the expansion in (2.30) with $N = 1$ for both AB and BA, then use (2.46) and (2.39). ☐

Iterating estimate (2.51), given $A \in OPS^m$ and $B \in OPS^{m'}$, we get estimates of the operators $\mathrm{Ad}^n_A(B)$, $n \in \mathbb{N}$, defined inductively by

$$\mathrm{Ad}_A(B) := [A,B] , \qquad \mathrm{Ad}^{n+1}_A(B) := [A, \mathrm{Ad}^n_A(B)] , \ n \in \mathbb{N} .$$

Lemma 2.12. *Let $A \in OPS^m$, $B \in OPS^{m'}$, $m,m' \in \mathbb{R}$. Then for any $n,\alpha \in \mathbb{N}$, $s \geq s_0$,*

$$\begin{aligned} |\mathrm{Ad}^n_A(B)|^{k_0,\gamma}_{nm+m'-n,s,\alpha} \lesssim_{m,m',s,\alpha,k_0} & (|A|^{k_0,\gamma}_{m,s_0+\mathtt{c}_n(m,m',\alpha),\alpha+n})^n |B|^{k_0,\gamma}_{m',s+\mathtt{c}_n(m,m',\alpha),\alpha+n} \\ & + (|A|^{k_0,\gamma}_{m,s_0+\mathtt{c}_n(m,m',\alpha),\alpha+n})^{n-1} |A|^{k_0,\gamma}_{m,s+\mathtt{c}_n(m,m',\alpha),\alpha+n} |B|^{k_0,\gamma}_{m',s_0+\mathtt{c}_n(m,m',\alpha),\alpha+n} \end{aligned} \tag{2.52}$$

where the constants $\mathtt{c}_n(m,m',\alpha)$ are

$$\mathtt{c}_n(m,m',\alpha) := n(2+\alpha) + \frac{n(n-1)}{2} + (n-1)|m| + \max\{|m|,|m'|\}, \quad n \geq 1. \tag{2.53}$$

Proof. Estimate (2.52) follows by applying iteratively (2.51). Bound (2.51) gives (2.52) for $n = 1$ with $\mathtt{c}_1 = 2+\alpha+\max\{|m|,|m'|\}$. The induction step requires that $2+\alpha+|m|+\mathtt{c}_n(m,m',\alpha+1) \leq \mathtt{c}_{n+1}(m,m',\alpha)$ and $2+\alpha+|nm+m'-n| \leq \mathtt{c}_{n+1}(m,m',\alpha)$ for all $n \geq 1$, which is satisfied by (2.53). ☐

The pseudo-differential norm of the adjoint A^* of a pseudo-differential operator $A = \mathrm{Op}(a) \in OPS^m$ (see (2.32)) may be estimated in terms of that of A.

Lemma 2.13. (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*

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

The same estimate holds if A is a matrix operator of the form (2.29).

Proof. See Lemma 2.16 in [21]. ☐

Finally we report a lemma about inverse of pseudo-differential operators.

Lemma 2.14. (Invertibility) *Let $\Phi := \mathrm{Id} + A$ where $A := \mathrm{Op}(a(\lambda,\varphi,x,\xi)) \in OPS^0$. There exist constants $C(s_0,\alpha,k_0)$, $C(s,\alpha,k_0) \geq 1$, $s \geq s_0$, such that, if*

$$C(s_0,\alpha,k_0) |A|^{k_0,\gamma}_{0,s_0+\alpha,\alpha} \leq 1/2 , \tag{2.54}$$

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

$$|\Phi^{-1} - \mathrm{Id}|^{k_0,\gamma}_{0,s,\alpha} \leq C(s,\alpha,k_0) |A|^{k_0,\gamma}_{0,s+\alpha,\alpha} . \tag{2.55}$$

The same estimate holds for a matrix operator $\Phi = \mathbb{I}_2 + A$ where $\mathbb{I}_2 = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Id} \end{pmatrix}$ and A has the form (2.29).

Proof. By a Neumann series argument. See Lemma 2.17 in [21]. ☐

2.4 Integral operators and Hilbert transform

In this section we consider the integral operators with a $\mathcal{C}^\infty$ kernel, which are the operators in $OPS^{-\infty}$. As in the previous section, we deal with families of such operators that are k_0 times differentiable with respect to a parameter $\lambda := (\omega, \mathtt{h})$ in an open subset $\Lambda_0 \subseteq \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$.

Lemma 2.15. *Let $K := K(\lambda, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^\nu \times \mathbb{T} \times \mathbb{T})$. Then the integral operator*

$$(\mathcal{R}u)(\varphi, x) := \int_{\mathbb{T}} K(\lambda, \varphi, x, y) u(\varphi, y)\, dy \tag{2.56}$$

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

$$|\mathcal{R}|_{-m,s,\alpha}^{k_0,\gamma} \le C(m, s, \alpha, k_0) \|K\|_{\mathcal{C}^{s+m+\alpha}}^{k_0,\gamma} . \tag{2.57}$$

Proof. See Lemma 2.32 in [21]. □

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

$$Pu(\varphi, x) := u(\varphi, x + p(\varphi, x)) . \tag{2.58}$$

Lemma 2.16. *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 $\|p\|_{2s_0+k_0+1}^{k_0,\gamma} \le \delta$, then the integral operator $\mathcal{R}$ in* (2.56) *transforms into the integral operator*

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

with a $\mathcal{C}^\infty$ kernel

$$\breve{K}(\lambda, \varphi, x, z) := \big(1 + \partial_z q(\lambda, \varphi, z)\big) K(\lambda, \varphi, x + q(\lambda, \varphi, x), z + q(\lambda, \varphi, z)), \tag{2.59}$$

where $z \mapsto z + q(\lambda, \varphi, z)$ is the inverse diffeomorphism of $x \mapsto x + p(\lambda, \varphi, x)$. The function $\breve{K}$ satisfies the estimates

$$\|\breve{K}\|_s^{k_0,\gamma} \le C(s, k_0)\big(\|K\|_{s+k_0}^{k_0,\gamma} + \|p\|_{s+k_0+1}^{k_0,\gamma} \|K\|_{s_0+k_0+1}^{k_0,\gamma}\big) \qquad \forall s \ge s_0 . \tag{2.60}$$

Proof. See Lemma 2.34 in [21]. □

We now recall some properties of the Hilbert transform $\mathcal{H}$ defined as a Fourier multiplier in (2.34). The Hilbert transform also admits an integral representation. Given a 2π-periodic function u, its Hilbert transform is

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

The commutator between the Hilbert transform $\mathcal{H}$ and the multiplication operator by a smooth function a is a regularizing operator in $OPS^{-\infty}$, as stated for example in Lemma 2.35 in [21] (see also Lemma B.5 in [6], Appendices H and I in [41] for similar statements).

Lemma 2.17. *Let $a(\lambda, \cdot, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$. Then the commutator $[a, \mathcal{H}]$ is in $OPS^{-\infty}$ and satisfies, for all $m, s, \alpha \in \mathbb{N}$,*

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

We also report the following classical lemma, see e.g. Lemma 2.36 in [21] and Lemma B.5 in [6] (and Appendices H and I in [41] for similar statements).

Lemma 2.18. *Let $p = p(\lambda, \cdot)$ be in $\mathcal{C}^\infty(\mathbb{T}^{\nu+1})$ and $P := P(\lambda, \cdot)$ be the associated change of variable defined in* (2.58). *There exists $\delta(s_0, k_0) > 0$ such that, if $\|p\|^{k_0,\gamma}_{2s_0+k_0+1} \leq \delta(s_0, k_0)$, then the operator $P^{-1}\mathcal{H}P - \mathcal{H}$ is an integral operator of the form*

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

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

$$K(\lambda, \varphi, x, z) := -\frac{1}{\pi}\partial_z \log(1 + g(\lambda, \varphi, x, z))\,, \tag{2.61}$$

with

$$g(\lambda, \varphi, x, z) := \cos\Big(\frac{q(\lambda, \varphi, x) - q(\lambda, \varphi, z)}{2}\Big) - 1 + \cos\Big(\frac{x - z}{2}\Big)\frac{\sin(\frac{1}{2}(q(\lambda, \varphi, x) - q(\lambda, \varphi, z)))}{\sin(\frac{1}{2}(x - z))} \tag{2.62}$$

where $z \mapsto q(\lambda, \varphi, z)$ is the inverse diffeomorphism of $x \mapsto x + p(\lambda, \varphi, x)$. The kernel K satisfies the estimate

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

We finally provide an estimate for the integral kernel of a family of Fourier multipliers in $OPS^{-\infty}$.

Lemma 2.19. *Let $g(\lambda, \varphi, \xi)$ be a family of Fourier multipliers with $\partial_\lambda^k g(\lambda, \varphi, \cdot) \in S^{-\infty}$, for all $k \in \mathbb{N}^{\nu+1}$, $|k| \leq k_0$. Then the operator $\mathrm{Op}(g)$ admits the integral representation*

$$\big[\mathrm{Op}(g)u\big](\varphi, x) = \int_{\mathbb{T}} K_g(\lambda, \varphi, x, y)u(\varphi, y)\, dy\,, \qquad K_g(\lambda, \varphi, x, y) := \frac{1}{2\pi}\sum_{j\in\mathbb{Z}} g(\lambda, \varphi, j)e^{\mathrm{i}j(x-y)}\,, \tag{2.63}$$

and the kernel K_g satisfies, for all $s \in \mathbb{N}$, the estimate

$$\|K_g\|^{k_0,\gamma}_{\mathcal{C}^s} \lesssim |\mathrm{Op}(g)|^{k_0,\gamma}_{-1,s+s_0,0} + |\mathrm{Op}(g)|^{k_0,\gamma}_{-s-s_0-1,0,0}\,. \tag{2.64}$$

Proof. It is straightforward to verify formula (2.63). For any $k \in \mathbb{N}^{\nu+1}$, $|k| \leq k_0$, $\alpha \in \mathbb{N}^\nu$ and for any $s, s_1, s_2 \in \mathbb{N}$ with $|\alpha| + s_1 + s_2 = s$, one has

$$\begin{aligned}\partial_\lambda^k\partial_\varphi^\alpha\partial_x^{s_1}\partial_y^{s_2}K_g(\lambda, x, y) &= \frac{1}{2\pi}\sum_{j\in\mathbb{Z}} \mathrm{i}^{s_1+s_2}(-1)^{s_2}j^{s_1+s_2}\partial_\lambda^k\partial_\varphi^\alpha g(\lambda, \varphi, j)e^{\mathrm{i}j(x-y)} \\ &= \frac{1}{2\pi}\sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \mathrm{i}^s(-1)^{s_2}j^{s_1+s_2}\ell^\alpha\partial_\lambda^k\widehat{g}(\lambda, \ell, j)e^{\mathrm{i}\ell\cdot\varphi}e^{\mathrm{i}j(x-y)}\,.\end{aligned}$$

Hence, using the Cauchy-Schwarz inequality, one gets

$$\begin{aligned}\gamma^{|k|}|\partial_\lambda^k\partial_x^{s_1}\partial_\varphi^\alpha\partial_y^{s_2}K_g(\lambda, x, y)| &\leq \gamma^{|k|}\sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} |j|^{s_1+s_2}|\ell|^{|\alpha|}|\partial_\lambda^k\widehat{g}(\lambda, \ell, j)| \lesssim \gamma^{|k|}\sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} \langle\ell, j\rangle^s|\partial_\lambda^k\widehat{g}(\lambda, \ell, j)| \\ &\lesssim \gamma^{|k|}\Big(\sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} \langle\ell, j\rangle^{2(s+s_0)}|\partial_\lambda^k\widehat{g}(\lambda, \ell, j)|^2\Big)^{\frac{1}{2}} \\ &\lesssim \gamma^{|k|}\Big(\sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} \langle\ell\rangle^{2(s+s_0)}|\partial_\lambda^k\widehat{g}(\lambda, \ell, j)|^2\Big)^{\frac{1}{2}} + \gamma^{|k|}\Big(\sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} \langle j\rangle^{2(s+s_0)}|\partial_\lambda^k\widehat{g}(\lambda, \ell, j)|^2\Big)^{\frac{1}{2}} \\ &\lesssim \gamma^{|k|}\Big(\sum_{j\in\mathbb{Z}}\frac{1}{\langle j\rangle^2}\sum_{\ell\in\mathbb{Z}^\nu} \langle\ell\rangle^{2(s+s_0)}\langle j\rangle^2|\partial_\lambda^k\widehat{g}(\lambda, \ell, j)|^2\Big)^{\frac{1}{2}} \\ &\quad + \gamma^{|k|}\Big(\sum_{j\in\mathbb{Z}}\frac{1}{\langle j\rangle^2}\sum_{\ell\in\mathbb{Z}^\nu} \langle j\rangle^{2(s+s_0+1)}|\partial_\lambda^k\widehat{g}(\lambda, \ell, j)|^2\Big)^{\frac{1}{2}} \\ &\lesssim |\mathrm{Op}(g)|^{k_0,\gamma}_{-1,s+s_0,0} + |\mathrm{Op}(g)|^{k_0,\gamma}_{-s-s_0-1,0,0}\,,\end{aligned}$$

which implies estimate (2.64). □

2.5 Reversible, Even, Real operators

We introduce now some algebraic properties that have a key role in the proof.

Definition 2.20. (Even operator) *A linear operator $A := A(\varphi)$ as in (2.22) 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 $u_{-j} = u_j$ for all $j \in \mathbb{Z}$, we have that

$$A \text{ is even} \quad \Longleftrightarrow \quad A_j^{j'}(\varphi) + A_j^{-j'}(\varphi) = A_{-j}^{j'}(\varphi) + A_{-j}^{-j'}(\varphi), \quad \forall j, j' \in \mathbb{Z}, \ \varphi \in \mathbb{T}^\nu. \tag{2.65}$$

Definition 2.21. (Reversibility) *An operator $\mathcal{R}$ as in (2.20) is*

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

The composition of a reversible operator with a reversibility preserving operator is reversible. It turns out that an operator $\mathcal{R}$ as in (2.20) 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.

We shall say that a linear operator of the form $\mathcal{L} := \omega \cdot \partial_\varphi + A(\varphi)$ is reversible, respectively even, if $A(\varphi)$ is reversible, respectively even. 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

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

It results that the conjugation of an even and reversible operator with an operator $\Phi(\varphi)$ that is even and reversibility preserving is even and reversible.

Lemma 2.22. *Let $A := \mathrm{Op}(a)$ be a pseudo-differential operator. Then the following holds:*

1. *If the symbol a satisfies $a(-x, -\xi) = a(x, \xi)$, then A is even.*
2. *If $A = \mathrm{Op}(a)$ is even, then the pseudo-differential operator $\mathrm{Op}(\tilde{a})$ with symbol*
$$\tilde{a}(x, \xi) := \frac{1}{2}\big(a(x, \xi) + a(-x, -\xi)\big) \tag{2.66}$$
coincides with $\mathrm{Op}(a)$ on the subspace $E := \{u(-x) = u(x)\}$ of the functions even in x, namely $\mathrm{Op}(\tilde{a})_{|E} = \mathrm{Op}(a)_{|E}$.
3. *A is real, i.e. it maps real functions into real functions, if and only if the symbol $\overline{a(x, -\xi)} = a(x, \xi)$.*
4. *Let $g(\xi)$ be a Fourier multiplier satisfying $g(\xi) = g(-\xi)$. If $A = \mathrm{Op}(a)$ is even, then the operator $\mathrm{Op}(a(x, \xi)g(\xi)) = \mathrm{Op}(a) \circ \mathrm{Op}(g)$ is an even operator. More generally, the composition of even operators is an even operator.*

We shall use the following remark.

Remark 2.23. *By item 2, we can replace an even pseudo-differential operator $\mathrm{Op}(a)$ acting on the subspace of functions even in x, with the operator $\mathrm{Op}(\tilde{a})$ where the symbol $\tilde{a}(x, \xi)$ defined in* (2.66) *satisfies $\tilde{a}(-x, -\xi) = \tilde{a}(x, \xi)$. The pseudo-differential norms of $\mathrm{Op}(a)$ and $\mathrm{Op}(\tilde{a})$ are equivalent. Moreover, the space average*

$$\langle \tilde{a} \rangle_x(\xi) := \frac{1}{2\pi} \int_{\mathbb{T}} \tilde{a}(x, \xi)\, dx \qquad \textit{satisfies} \qquad \langle \tilde{a} \rangle_x(-\xi) = \langle \tilde{a} \rangle_x(\xi),$$

and, therefore, the Fourier multiplier $\langle \tilde{a} \rangle_x(D)$ is even.

It is convenient to consider a real operator $\mathcal{R} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ as in (2.20), which acts on the real variables $(\eta, \psi) \in \mathbb{R}^2$, as a linear operator acting on the complex variables $(u, \bar{u})$ introduced by the linear change of coordinates $(\eta, \psi) = \mathcal{C}(u, \bar{u})$, where

$$\mathcal{C} := \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \end{pmatrix}, \quad \mathcal{C}^{-1} = \begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix}. \tag{2.67}$$

We get that the *real* operator $\mathcal{R}$ acting in the complex coordinates $(u, \bar{u}) = \mathcal{C}^{-1}(\eta, \psi)$ takes the form

$$\begin{gathered} \mathbf{R} = \mathcal{C}^{-1} \mathcal{R} \mathcal{C} := \begin{pmatrix} \mathcal{R}_1 & \mathcal{R}_2 \\ \overline{\mathcal{R}_2} & \overline{\mathcal{R}_1} \end{pmatrix}, \\ \mathcal{R}_1 := \frac{1}{2} \big\{ (A + D) - \mathrm{i}(B - C) \big\}, \quad \mathcal{R}_2 := \frac{1}{2} \big\{ (A - D) + \mathrm{i}(B + C) \big\} \end{gathered} \tag{2.68}$$

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

$$\overline{A}(u) := \overline{A(\bar{u})}. \tag{2.69}$$

We say that a matrix operator acting on the complex variables $(u, \bar{u})$ is REAL if it has the structure in (2.68) and it is EVEN if both $\mathcal{R}_1$, $\mathcal{R}_2$ are even. The composition of two real (resp. even) operators is a real (resp. even) operator.

The following properties of the conjugated operator hold:

1. $\overline{AB} = \overline{A}\,\overline{B}$.
2. If $(A_j^{j'})$ is the matrix of A, then the matrix entries of $\overline{A}$ are $(\overline{A})_j^{j'} = \overline{A_{-j}^{-j'}}$.
3. If $A = \mathrm{Op}(a(x, \xi))$ is a pseudo-differential operator, then its conjugate is $\overline{A} = \mathrm{Op}(\overline{a(x, -\xi)})$. The pseudo differential norms of A and $\overline{A}$ are equal, namely $|A|_{m,s,\alpha}^{k_0,\gamma} = |\overline{A}|_{m,s,\alpha}^{k_0,\gamma}$.

In the complex coordinates $(u, \bar{u}) = \mathcal{C}^{-1}(\eta, \psi)$ the involution ρ defined in (1.11) reads as the map $u \mapsto \bar{u}$.

Lemma 2.24. *Let* $\mathbf{R}$ *be a real operator as in* (2.68)*. One has*

1. $\mathbf{R}$ *is reversible if and only if* $\mathcal{R}_i(-\varphi) = -\overline{\mathcal{R}_i}(\varphi)$ *for all* $\varphi \in \mathbb{T}^\nu$, $i = 1, 2$, *or equivalently*

$$(\mathcal{R}_i)_j^{j'}(-\varphi) = -\overline{(\mathcal{R}_i)_{-j}^{-j'}(\varphi)} \quad \forall \varphi \in \mathbb{T}^\nu, \quad \textit{i.e.} \ (\mathcal{R}_i)_j^{j'}(\ell) = -\overline{(\mathcal{R}_i)_{-j}^{-j'}(\ell)} \quad \forall \ell \in \mathbb{Z}^\nu. \tag{2.70}$$

2. $\mathbf{R}$ *is reversibility preserving if and only if* $\mathcal{R}_i(-\varphi) = \overline{\mathcal{R}_i}(\varphi)$ *for all* $\varphi \in \mathbb{T}^\nu$, $i = 1, 2$, *or equivalently*

$$(\mathcal{R}_i)_j^{j'}(-\varphi) = \overline{(\mathcal{R}_i)_{-j}^{-j'}(\varphi)} \quad \forall \varphi \in \mathbb{T}^\nu, \quad \textit{i.e.} \ (\mathcal{R}_i)_j^{j'}(\ell) = \overline{(\mathcal{R}_i)_{-j}^{-j'}(\ell)} \quad \forall \ell \in \mathbb{Z}^\nu. \tag{2.71}$$

2.6 $\mathcal{D}^{k_0}$-tame and modulo-tame operators

In this section we recall the notion and the main properties of $\mathcal{D}^{k_0}$-tame and modulo-tame operators that will be used in the paper. For the proofs we refer to Section 2.2 of [21] where this notion was introduced.

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

Definition 2.25. **($\mathcal{D}^{k_0}$-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* $s_0 \leq s \leq S$, *possibly with* $S = +\infty$, *for all* $u \in H^{s+\sigma}$,

$$\sup_{|k| \leq k_0} \sup_{\lambda \in \Lambda_0} \gamma^{|k|} \| (\partial_\lambda^k A(\lambda)) u \|_s \leq \mathfrak{M}_A(s_0) \|u\|_{s+\sigma} + \mathfrak{M}_A(s) \|u\|_{s_0+\sigma}, \tag{2.72}$$

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" σ is zero, we simply write $\mathcal{D}^{k_0}$-tame instead of $\mathcal{D}^{k_0}$-0-tame.

For a real matrix operator (as in (2.68)*)*

$$A = \begin{pmatrix} A_1 & A_2 \\ \overline{A_2} & \overline{A_1} \end{pmatrix}, \tag{2.73}$$

we denote the tame constant $\mathfrak{M}_A(s) := \max\{\mathfrak{M}_{A_1}(s), \mathfrak{M}_{A_2}(s)\}$.

Remark 2.26. *In Sections 7-15 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 S, for instance $\mathfrak{M}_A(s) \leq C(S)(1 + \|\mathfrak{I}_0\|_{s+\mu}^{k_0,\gamma})$, for all $s_0 \leq s \leq S$.*

An immediate consequence of (2.72) (with $k = 0$, $s = s_0$) is that

$$\|A\|_{\mathcal{L}(H^{s_0+\sigma}, H^{s_0})} \leq 2\mathfrak{M}_A(s_0). \tag{2.74}$$

Also note 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 (2.22) we have, for all $|k| \leq k_0$, $j' \in \mathbb{Z}$, $\ell' \in \mathbb{Z}^\nu$,

$$\gamma^{2|k|} \sum_{\ell,j} \langle \ell, j\rangle^{2s} |\partial_\lambda^k A_j^{j'}(\ell - \ell')|^2 \leq 2\big(\mathfrak{M}_A(s_0)\big)^2 \langle \ell', j'\rangle^{2(s+\sigma)} + 2\big(\mathfrak{M}_A(s)\big)^2 \langle \ell', j'\rangle^{2(s_0+\sigma)}. \tag{2.75}$$

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

Lemma 2.27. (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 composition $A \circ B$ is $\mathcal{D}^{k_0}$-$(\sigma_A + \sigma_B)$-tame with tame constant*

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

The same estimate holds if A, B are matrix operators as in (2.73).

Proof. The proof is straightforward (see Lemma 2.20 in [21]). ☐

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

Lemma 2.28. (Action on H^s) *Let $A := A(\lambda)$ be a $\mathcal{D}^{k_0}$-σ-tame operator. Then, $\forall s \geq s_0$, 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:*

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

The same estimate holds if A is a matrix operator as in (2.73).

Proof. The proof is straightforward (see Lemma 2.22 in [21]). ☐

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

Lemma 2.29. *Let $A = a(\lambda, \varphi, x, D) \in OPS^0$ be a family of pseudo-differential operators that 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*

$$\mathfrak{M}_A(s) \leq C(s)|A|_{0,s,0}^{k_0,\gamma}. \tag{2.76}$$

As a consequence

$$\|Ah\|_s^{k_0,\gamma} \leq C(s_0, k_0)|A|_{0,s_0,0}^{k_0,\gamma}\|h\|_s^{k_0,\gamma} + C(s, k_0)|A|_{0,s,0}^{k_0,\gamma}\|h\|_{s_0}^{k_0,\gamma}. \tag{2.77}$$

The same statement holds if A is a matrix operator of the form (2.73).

Proof. See Lemma 2.21 in [21] for the proof of (2.76), then apply Lemma 2.28 to deduce (2.77). ☐

In view of the KAM reducibility scheme of Section 15, we also consider the stronger notion of $\mathcal{D}^{k_0}$-modulo-tame operator, which we need only for operators with loss of derivatives $\sigma = 0$.

Definition 2.30. **($\mathcal{D}^{k_0}$-modulo-tame)** *A linear operator $A := A(\lambda)$ is $\mathcal{D}^{k_0}$-modulo-tame if, for all $k \in \mathbb{N}^{\nu+1}$, $|k| \leq k_0$, the majorant operators $|\partial_\lambda^k A|$ (Definition 2.7) satisfy the following weighted tame estimates: for all $s_0 \leq s \leq S$, $u \in H^s$,*

$$\sup_{|k|\leq k_0} \sup_{\lambda\in\Lambda_0} \gamma^{|k|} \| \, |\partial_\lambda^k A| u\|_s \leq \mathfrak{M}_A^\sharp(s_0)\|u\|_s + \mathfrak{M}_A^\sharp(s)\|u\|_{s_0} \tag{2.78}$$

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

For a matrix operator as in (2.73) *we denote the modulo tame constant $\mathfrak{M}_A^\sharp(s) := \max\{\mathfrak{M}_{A_1}^\sharp(s), \mathfrak{M}_{A_2}^\sharp(s)\}$.*

If A, B are $\mathcal{D}^{k_0}$-modulo-tame operators, with $|A_j^{j'}(\ell)| \leq |B_j^{j'}(\ell)|$, then $\mathfrak{M}_A^\sharp(s) \leq \mathfrak{M}_B^\sharp(s)$.

Lemma 2.31. *An operator A that is $\mathcal{D}^{k_0}$-modulo-tame is also $\mathcal{D}^{k_0}$-tame and $\mathfrak{M}_A(s) \leq \mathfrak{M}_A^\sharp(s)$. The same holds if A is a matrix operator as in* (2.73).

Proof. For all $k \in \mathbb{N}^{\nu+1}$ with $|k| \leq k_0$ and for all $u \in H^s$, one has

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

Then the thesis follows by (2.27) and by Definitions 2.25 and 2.30. ☐

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

Lemma 2.32. **(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*

$$\mathfrak{M}_{A+B}^\sharp(s) \leq \mathfrak{M}_A^\sharp(s) + \mathfrak{M}_B^\sharp(s) \, . \tag{2.79}$$

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

$$\mathfrak{M}_{AB}^\sharp(s) \leq C(k_0)\big(\mathfrak{M}_A^\sharp(s)\mathfrak{M}_B^\sharp(s_0) + \mathfrak{M}_A^\sharp(s_0)\mathfrak{M}_B^\sharp(s)\big) \, . \tag{2.80}$$

Assume in addition that $\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A$, $\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} B$ (see Definition 2.7) are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constant respectively $\mathfrak{M}_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A}^\sharp(s)$ and $\mathfrak{M}_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} B}^\sharp(s)$. Then $\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}(AB)$ is $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constant satisfying

$$\begin{aligned} \mathfrak{M}_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}(AB)}^\sharp(s) \leq C(\mathtt{b})C(k_0)\Big(&\mathfrak{M}_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A}^\sharp(s)\mathfrak{M}_B^\sharp(s_0) + \mathfrak{M}_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A}^\sharp(s_0)\mathfrak{M}_B^\sharp(s) \\ &+ \mathfrak{M}_A^\sharp(s)\mathfrak{M}_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} B}^\sharp(s_0) + \mathfrak{M}_A^\sharp(s_0)\mathfrak{M}_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} B}^\sharp(s)\Big) \end{aligned} \tag{2.81}$$

for some constants $C(k_0), C(\mathtt{b}) \geq 1$. The same statement holds if A and B are matrix operators as in (2.73).

Proof. PROOF OF (2.79), (2.80). These estimates have been proved in Lemma 2.25 of [21].
PROOF OF (2.81). For all $|k| \leq k_0$ we have (use the first inequality in (2.25))

$$\| \, |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} [\partial_\lambda^k (AB)] \, |u\|_s \leq C(k_0) \sum\nolimits_{k_1+k_2=k} \| \, |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} [(\partial_\lambda^{k_1} A)(\partial_\lambda^{k_2} B)] \, | |\!|u|\!| \|_s \, . \tag{2.82}$$

Next, recalling Definition 2.7 of the operator $\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}$ and (2.26), we have

$$\begin{aligned} \Big\| \, |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} [(\partial_\lambda^{k_1} A)(\partial_\lambda^{k_2} B)] \, | |\!|u|\!| \Big\|_s^2 &= \sum_{\ell,j} \langle \ell, j\rangle^{2s} \Big(\sum_{\ell',j'} |\langle \ell-\ell', j-j'\rangle^{\mathtt{b}} [(\partial_\lambda^{k_1} A)(\partial_\lambda^{k_2} B)]_j^{j'}(\ell-\ell')||u_{\ell',j'}|\Big)^2 \\ &\leq \sum_{\ell,j} \langle \ell, j\rangle^{2s} \Big(\sum_{\ell',j',\ell_1,j_1} \langle \ell-\ell', j-j'\rangle^{\mathtt{b}} |(\partial_\lambda^{k_1} A)_j^{j_1}(\ell-\ell_1)||(\partial_\lambda^{k_2} B)_{j_1}^{j'}(\ell_1-\ell')||u_{\ell',j'}|\Big)^2 \, . \end{aligned} \tag{2.83}$$

Since $\langle \ell-\ell', j-j'\rangle^{\mathtt b} \lesssim_{\mathtt b} \langle \ell-\ell_1, j-j_1\rangle^{\mathtt b} + \langle \ell_1-\ell', j_1-j'\rangle^{\mathtt b}$, we deduce that

$$
\begin{aligned}
(2.83) \lesssim_{\mathtt b} & \sum_{\ell,j} \langle \ell,j\rangle^{2s} \Big(\sum_{\ell',j',\ell_1,j_1} |\langle \ell-\ell_1, j-j_1\rangle^{\mathtt b} (\partial_\lambda^{k_1} A)_j^{j_1}(\ell-\ell_1)| |(\partial_\lambda^{k_2} B)_{j_1}^{j'}(\ell_1-\ell')| |u_{\ell',j'}| \Big)^2 \\
& + \sum_{\ell,j} \langle \ell,j\rangle^{2s} \Big(\sum_{\ell',j',\ell_1,j_1} |(\partial_\lambda^{k_1} A)_j^{j_1}(\ell-\ell_1)| |\langle \ell_1-\ell', j_1-j'\rangle^{\mathtt b} (\partial_\lambda^{k_2} B)_{j_1}^{j'}(\ell_1-\ell')| |u_{\ell',j'}| \Big)^2 \\
\lesssim_{\mathtt b} & \Big\| \big| \langle \partial_{\varphi,x}\rangle^{\mathtt b} (\partial_\lambda^{k_1} A) \big| \big[|\partial_\lambda^{k_2} B| |u| \big] \Big\|_s^2 + \Big\| \big| \partial_\lambda^{k_1} A \big| \big[|\langle \partial_{\varphi,x}\rangle^{\mathtt b} (\partial_\lambda^{k_2} B)| |u| \big] \Big\|_s^2 . \qquad (2.84)
\end{aligned}
$$

Hence (2.82)-(2.84), (2.78) and (2.27) imply

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

which proves (2.81). ☐

By (2.80), if A is $\mathcal{D}^{k_0}$-modulo-tame, then, for all $n \ge 1$, each A^n is $\mathcal{D}^{k_0}$-modulo-tame and

$$
\mathfrak{M}^\sharp_{A^n}(s) \le \big(2C(k_0) \mathfrak{M}^\sharp_A(s_0) \big)^{n-1} \mathfrak{M}^\sharp_A(s) . \qquad (2.85)
$$

Moreover, by (2.81) and (2.85), if $\langle \partial_{\varphi,x}\rangle^{\mathtt b} A$ is $\mathcal{D}^{k_0}$-modulo-tame, then, for all $n \ge 2$, each $\langle \partial_{\varphi,x}\rangle^{\mathtt b} A^n$ is $\mathcal{D}^{k_0}$-modulo-tame with

$$
\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt b} A^n}(s) \le (4C(\mathtt b)C(k_0))^{n-1} \Big(\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt b} A}(s) \big[\mathfrak{M}^\sharp_A(s_0) \big]^{n-1} + \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt b} A}(s_0) \mathfrak{M}^\sharp_A(s) \big[\mathfrak{M}^\sharp_A(s_0) \big]^{n-2} \Big) . \quad (2.86)
$$

Estimates (2.85), (2.86) hold also in the case when A is a matrix operator of the form (2.73).

Lemma 2.33. (Invertibility) *Let $\Phi := \mathrm{Id} + A$, where A and $\langle \partial_{\varphi,x}\rangle^{\mathtt b} A$ are $\mathcal{D}^{k_0}$-modulo-tame. Assume the smallness condition*

$$
4C(\mathtt b) C(k_0) \mathfrak{M}^\sharp_A(s_0) \le 1/2 . \qquad (2.87)
$$

Then the operator Φ is invertible, $\check{A} := \Phi^{-1} - \mathrm{Id}$ is $\mathcal{D}^{k_0}$-modulo-tame, as well as $\langle \partial_{\varphi,x}\rangle^{\mathtt b} \check{A}$, and

$$
\mathfrak{M}^\sharp_{\check A}(s) \le 2 \mathfrak{M}^\sharp_A(s) , \quad \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt b} \check A}(s) \le 2 \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt b} A}(s) + 8C(\mathtt b) C(k_0) \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt b} A}(s_0)\, \mathfrak{M}^\sharp_A(s) .
$$

The same statement holds if A is a matrix operator of the form (2.73).

Proof. The lemma follows by a Neumann series argument, using (2.79) and (2.85)-(2.86). ☐

We also have the following consequence.

Corollary 2.34. *Let $m \in \mathbb{R}$, $\Phi := \mathrm{Id} + A$ where $\langle D\rangle^m A \langle D\rangle^{-m}$ and $\langle \partial_{\varphi,x}\rangle^{\mathtt b} \langle D\rangle^m A \langle D\rangle^{-m}$ are $\mathcal{D}^{k_0}$-modulo-tame. Assume the smallness condition*

$$
4C(\mathtt b) C(k_0) \mathfrak{M}^\sharp_{\langle D\rangle^m A \langle D\rangle^{-m}}(s_0) \le 1/2 . \qquad (2.88)
$$

Let $\check A := \Phi^{-1} - \mathrm{Id}$. Then the operators $\langle D\rangle^m \check A \langle D\rangle^{-m}$ and $\langle \partial_{\varphi,x}\rangle^{\mathtt b} \langle D\rangle^m \check A \langle D\rangle^{-m}$ are $\mathcal{D}^{k_0}$-modulo-tame, with

$$
\begin{aligned}
& \mathfrak{M}^\sharp_{\langle D\rangle^m \check A \langle D\rangle^{-m}}(s) \le 2 \mathfrak{M}^\sharp_{\langle D\rangle^m A \langle D\rangle^{-m}}(s) , \\
& \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt b} \langle D\rangle^m \check A \langle D\rangle^{-m}}(s) \le 2 \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt b} \langle D\rangle^m A \langle D\rangle^{-m}}(s) + 8C(\mathtt b) C(k_0) \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt b} \langle D\rangle^m A \langle D\rangle^{-m}}(s_0)\, \mathfrak{M}^\sharp_{\langle D\rangle^m A \langle D\rangle^{-m}}(s) .
\end{aligned}
$$

The same statement holds if A is a matrix operator of the form (2.73).

Proof. Let us write $\Phi_m := \langle D\rangle^m \Phi \langle D\rangle^{-m} = \mathrm{Id} + A_m$ with $A_m := \langle D\rangle^m A \langle D\rangle^{-m}$. The corollary follows by Lemma 2.33, since the smallness condition (2.88) is (2.87) with $A = A_m$, and $\Phi_m^{-1} = \mathrm{Id} + \langle D\rangle^m \check{A} \langle D\rangle^{-m}$. ☐

Lemma 2.35. (Smoothing) *Suppose that $\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A$, $\mathtt{b} \geq 0$, is $\mathcal{D}^{k_0}$-modulo-tame. Then the operator $\Pi_N^\perp A$ (see Definition 2.7) is $\mathcal{D}^{k_0}$-modulo-tame with tame constant*

$$\mathfrak{M}^\sharp_{\Pi_N^\perp A}(s) \leq N^{-\mathtt{b}} \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A}(s)\,, \quad \mathfrak{M}^\sharp_{\Pi_N^\perp A}(s) \leq \mathfrak{M}^\sharp_A(s)\,. \tag{2.89}$$

The same estimate holds when A is a matrix operator of the form (2.73)*.*

Proof. For all $|k| \leq k_0$ one has, recalling (2.24),

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

and, using (2.78), (2.27), we deduce the first inequality in (2.89). 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 (2.89). ☐

The next lemmata will be used in the proof of the reducibility Theorem 15.4.

Lemma 2.36. *Let A and B be linear operators such that $|A|, |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A|, |B|, |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} B| \in \mathcal{L}(H^{s_0})$. Then*

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

The same estimates hold if A and B are matrix operators of the form (2.73)*, by replacing in the estimates $\mathcal{L}(H^{s_0})$ by $\mathcal{L}(H^{s_0} \times H^{s_0})$.*

Proof. Items 1-2 are a direct consequence of (2.25) and (2.27). Items 3-4 are proved arguing as in Lemmata 2.32 and 2.35. ☐

Lemma 2.37. *Let $\Phi_i := \mathrm{Id} + \Psi_i$, $i = 1, 2$, satisfy*

$$|||\Psi_i|||_{\mathcal{L}(H^{s_0})} \leq 1/2\,, \quad i = 1, 2\,. \tag{2.90}$$

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

$$\begin{aligned}
\| \langle \partial_{\varphi,x}\rangle^{\mathtt{b}} |\check{\Psi}_1 - \check{\Psi}_2| \|_{\mathcal{L}(H^{s_0})} &\lesssim_{\mathtt{b}} \| \langle \partial_{\varphi,x}\rangle^{\mathtt{b}} |\Psi_1 - \Psi_2| \|_{\mathcal{L}(H^{s_0})} \\
&\quad + \big(1 + |||\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \check{\Psi}_1|||_{\mathcal{L}(H^{s_0})} + |||\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \check{\Psi}_2|||_{\mathcal{L}(H^{s_0})} \big) |||\Psi_1 - \Psi_2|||_{\mathcal{L}(H^{s_0})}\,.
\end{aligned}$$

The same statements hold if Ψ_1 and Ψ_2 are matrix operators of the form (2.73) *(where $\mathcal{L}(H^{s_0})$ stands for $\mathcal{L}(H^{s_0} \times H^{s_0})$).*

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 2.36, using (2.90). ☐

Lemma 2.38. *Let $m \in \mathbb{R}$, $\Phi_i := \mathrm{Id} + \Psi_i$, $i = 1, 2$, satisfy*

$$\| |\langle D\rangle^m \Psi_i \langle D\rangle^{-m}| \|_{\mathcal{L}(H^{s_0})} \leq 1/2\,, \quad i = 1, 2\,. \tag{2.91}$$

Then $\Phi_i^{-1} = \mathrm{Id} + \check{\Psi}_i$, $i = 1, 2$, satisfy

$$\| |\langle D\rangle^m (\check{\Psi}_1 - \check{\Psi}_2) \langle D\rangle^{-m}| \|_{\mathcal{L}(H^{s_0})} \leq 4 \| |\langle D\rangle^m (\Psi_1 - \Psi_2) \langle D\rangle^{-m}| \|_{\mathcal{L}(H^{s_0})}$$

and

$$\begin{aligned} &\| |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^m (\check{\Psi}_1 - \check{\Psi}_2) \langle D\rangle^{-m}| \|_{\mathcal{L}(H^{s_0})} \lesssim_{\mathtt{b}} \| |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^m (\Psi_1 - \Psi_2) \langle D\rangle^{-m}| \|_{\mathcal{L}(H^{s_0})} \\ &\quad + \big(1 + \| |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^m \check{\Psi}_1 \langle D\rangle^{-m}| \|_{\mathcal{L}(H^{s_0})} + \| |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^m \check{\Psi}_2 \langle D\rangle^{-m}| \|_{\mathcal{L}(H^{s_0})}\big) \| |\langle D\rangle^m (\Psi_1 - \Psi_2) \langle D\rangle^{-m}| \|_{\mathcal{L}(H^{s_0})}\,. \end{aligned}$$

The same statements hold if Ψ_1 and Ψ_2 are matrix operators of the form (2.73) *(where $\mathcal{L}(H^{s_0})$ stands for $\mathcal{L}(H^{s_0} \times H^{s_0})$).*

Proof. The lemma follows writing $\Phi_{i,m} = \mathrm{Id} + \Psi_{i,m}$ and $\Psi_{i,m} := \langle D\rangle^m \Psi_i \langle D\rangle^{-m}$, $i = 1, 2$, by applying Lemma 2.37. □

Lemma 2.39. *Let π_0 be the projector defined in* (2.35) *by $\pi_0 u := \frac{1}{2\pi} \int_{\mathbb{T}} u(x)\, dx$. Let A, B be φ-dependent families of operators as in* (2.22) *that, together with their adjoints A^*, B^* with respect to the L^2_x scalar product, are $\mathcal{D}^{k_0}$-σ-tame. Let $m_1, m_2 \geq 0$, $\beta_0 \in \mathbb{N}$. Then for any $\beta \in \mathbb{N}^\nu$, $|\beta| \leq \beta_0$, the operator $\langle D\rangle^{m_1} \big(\partial_\varphi^\beta (A\pi_0 B - \pi_0)\big) \langle D\rangle^{m_2}$ is $\mathcal{D}^{k_0}$-tame with tame constant satisfying, for all $s \geq s_0$,*

$$\begin{aligned} \mathfrak{M}_{\langle D\rangle^{m_1} (\partial_\varphi^\beta (A\pi_0 B - \pi_0)) \langle D\rangle^{m_2}}(s) &\lesssim_{m,s,\beta_0,k_0} \mathfrak{M}_{A - \mathrm{Id}}(s + \beta_0 + m_1)\big(1 + \mathfrak{M}_{B^* - \mathrm{Id}}(s_0 + m_2)\big) \\ &\quad + \mathfrak{M}_{B^* - \mathrm{Id}}(s + \beta_0 + m_2)\big(1 + \mathfrak{M}_{A - \mathrm{Id}}(s_0 + m_1)\big). \end{aligned} \tag{2.92}$$

The same estimate holds if A, B are matrix operators of the form (2.73) *and π_0 is replaced by the matrix operator Π_0 defined in* (11.2).

Proof. Writing $A\pi_0 B - \pi_0 = (A - \mathrm{Id})\pi_0 B + \pi_0(B - \mathrm{Id})$ and using the identity $\langle D\rangle^m \pi_0 = \pi_0$ we get

$$\langle D\rangle^{m_1} \big(A\pi_0 B - \pi_0\big) \langle D\rangle^{m_2}[h] = g_1 (h, g_2)_{L^2_x} + (h, g_3)_{L^2_x} \tag{2.93}$$

where g_1, g_2, g_3 are the functions

$$g_1 := \frac{1}{2\pi} \langle D\rangle^{m_1} (A - \mathrm{Id})[1]\,, \quad g_2 := \langle D\rangle^{m_2} B^*[1]\,, \quad g_3 := \frac{1}{2\pi} \langle D\rangle^{m_2} (B^* - \mathrm{Id})[1]$$

(thus the operator (2.93) has the "finite dimensional" form as in (7.3)). We estimate

$$\|g_1\|_s^{k_0,\gamma} \lesssim_{k_0} \mathfrak{M}_{A - \mathrm{Id}}(s + m_1), \quad \|g_2\|_s^{k_0,\gamma} \lesssim_{k_0} 1 + \mathfrak{M}_{B^* - \mathrm{Id}}(s + m_2), \quad \|g_3\|_s^{k_0,\gamma} \lesssim_{k_0} \mathfrak{M}_{B^* - \mathrm{Id}}(s + m_2)\,. \tag{2.94}$$

For $\beta \in \mathbb{N}^\nu$, $k \in \mathbb{N}^{\nu+1}$ with $|\beta| \leq \beta_0$, $|k| \leq k_0$, the operator obtained differentiating (2.93) is

$$\partial_\lambda^k \partial_\varphi^\beta \big(\langle D\rangle^{m_1} (A\pi_0 B - \pi_0) \langle D\rangle^{m_2}\big)[h] = \sum_{\substack{\beta_1 + \beta_2 = \beta \\ k_1 + k_2 = k}} C(\beta_1, \beta_2, k_1, k_2) \partial_\lambda^{k_1} \partial_\varphi^{\beta_1} g_1 \big(h, \partial_\lambda^{k_2} \partial_\varphi^{\beta_2} g_2\big)_{L^2_x} + \big(h, \partial_\lambda^k \partial_\varphi^\beta g_3\big)_{L^2_x}\,.$$

Bound (2.92) follows, recalling Definition 2.25, applying (2.10), (2.9), (2.94). □

2.7 Tame estimates for the flow of pseudo-PDEs

We report in this section several results concerning tame estimates for the flow Φ^t of the pseudo-PDE

$$\begin{cases} \partial_t u = \mathrm{i} a(\varphi, x) |D|^{\frac{1}{2}} u \\ u(0, x) = u_0(x)\,, \end{cases} \qquad \varphi \in \mathbb{T}^\nu\,, \quad x \in \mathbb{T}\,, \tag{2.95}$$

where $a(\varphi, x) = a(\lambda, \varphi, x)$ is a real valued function that is $\mathcal{C}^\infty$ with respect to the variables (φ, x) and k_0 times differentiable with respect to the parameters $\lambda = (\omega, \mathtt{h})$. The function $a := a(i)$ may depend also on the "approximate" torus $i(\varphi)$. Most of these results have been obtained in the Appendix of [21].

The flow operator $\Phi^t := \Phi(t) := \Phi(\lambda, \varphi, t)$ satisfies the equation

$$\begin{cases} \partial_t \Phi(t) = \mathrm{i} a(\varphi, x) |D|^{\frac{1}{2}} \Phi(t) \\ \Phi(0) = \mathrm{Id}\,. \end{cases} \tag{2.96}$$

Since the function $a(\varphi, x)$ is real valued, usual energy estimates imply that the flow $\Phi(t)$ is a bounded operator mapping H^s_x to H^s_x. In the Appendix of [21] it is proved that the flow $\Phi(t)$ satisfies also tame estimates in $H^s_{\varphi,x}$, see Proposition 2.40 below. Moreover, since (2.95) is an autonomous equation, its flow $\Phi(\varphi, t)$ satisfies the group property

$$\Phi(\varphi, t_1 + t_2) = \Phi(\varphi, t_1) \circ \Phi(\varphi, t_2)\,, \qquad \Phi(\varphi, t)^{-1} = \Phi(\varphi, -t)\,, \tag{2.97}$$

and, since $a(\lambda, \cdot)$ is k_0 times differentiable with respect to the parameter λ, then $\Phi(\lambda, \varphi, t)$ is k_0 times differentiable with respect to λ as well. Also notice that $\Phi^{-1}(t) = \Phi(-t) = \overline{\Phi}(t)$, because these operators solve the same Cauchy problem. Moreover, if $a(\varphi, x)$ is odd(φ)even(x), then, recalling Section 2.5, the real operator

$$\boldsymbol{\Phi}(\varphi, t) := \begin{pmatrix} \Phi(\varphi, t) & 0 \\ 0 & \overline{\Phi}(\varphi, t) \end{pmatrix}$$

is even and reversibility preserving.

Proposition 2.40. *Assume that $\|a\|_{2s_0+\frac{3}{2}} \leq 1$ and $\|a\|_{2s_0+1} \leq \delta(s)$ for some $\delta(s) > 0$ small. Then the following tame estimates hold:*

$$\sup_{t\in[0,1]} \|\Phi(t) u_0\|_s \lesssim_s \|u_0\|_s\,, \qquad \forall s \in [0, s_0 + 1]\,, \tag{2.98}$$

$$\sup_{t\in[0,1]} \|\Phi(t) u_0\|_s \lesssim_s \|u_0\|_s + \|a\|_{s+s_0+\frac{1}{2}} \|u_0\|_{s_0}\,, \quad \forall s \geq s_0\,. \tag{2.99}$$

Proof. The proof is given in Proposition A.5 in [21]. □

The operator $\partial_\lambda^k \partial_\varphi^\beta \Phi$ loses $|D_x|^{\frac{|\beta|+|k|}{2}}$ derivatives, which, in (2.101) below, are compensated by $\langle D \rangle^{-m_1}$ on the left hand side and $\langle D \rangle^{-m_2}$ on the right hand side, with $m_1, m_2 \in \mathbb{R}$ satisfying $m_1 + m_2 = \frac{|\beta|+|k|}{2}$. The following proposition provides tame estimates in the Sobolev spaces $H^s_{\varphi,x}$.

Proposition 2.41. *Let $\beta_0, k_0 \in \mathbb{N}$. For any $\beta, k \in \mathbb{N}^\nu$ with $|\beta| \leq \beta_0$, $|k| \leq k_0$, for any $m_1, m_2 \in \mathbb{R}$ with $m_1 + m_2 = \frac{|\beta|+|k|}{2}$, for any $s \geq s_0$, there exist constants $\sigma(|\beta|, |k|, m_1, m_2) > 0$, $\delta(s, m_1) > 0$ such that if*

$$\|a\|_{2s_0+|m_1|+2} \leq \delta(s, m_1)\,, \quad \|a\|^{k_0,\gamma}_{s_0+\sigma(\beta_0,k_0,m_1,m_2)} \leq 1\,, \tag{2.100}$$

then the following estimate holds:

$$\sup_{t\in[0,1]} \|\langle D \rangle^{-m_1} \partial_\lambda^k \partial_\varphi^\beta \Phi(t) \langle D \rangle^{-m_2} h\|_s \lesssim_{s,\beta_0,k_0,m_1,m_2} \gamma^{-|k|} \Big(\|h\|_s + \|a\|^{k_0,\gamma}_{s+\sigma(|\beta|,|k|,m_1,m_2)} \|h\|_{s_0} \Big)\,. \tag{2.101}$$

Proof. We take $h \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$, so that $\partial_\lambda^k \partial_\varphi^\beta \Phi(\varphi) h$ is $\mathcal{C}^\infty$ for any $|\beta| \leq \beta_0$, $|k| \leq k_0$. We argue by induction on (k, β). We introduce the following notation:

- **Notation:** given $k', k \in \mathbb{N}^{\nu+1}$, we say that $k' \prec k$ if each component $k'_m \leq k_m$ for all $m = 1, \ldots, \nu + 1$ and $k' \neq k$. Given $(k', \beta'), (k, \beta) \in \mathbb{N}^{\nu+1} \times \mathbb{N}^\nu$, we say that $(k', \beta') \prec (k, \beta)$ if $k'_m \leq k_m$, $\beta'_n \leq \beta_n$ for all $m = 1, \ldots, \nu + 1$ and all $n = 1, \ldots, \nu$, and $(k', \beta') \neq (k, \beta)$.

Proof of (2.101) for $k = \beta = 0$. Since $m_1 + m_2 = \frac{|\beta|+|k|}{2} = 0$, we need to estimate the operator $\Phi_m(t) := \langle D\rangle^m \Phi(t) \langle D\rangle^{-m}$ where $m := -m_1 = m_2 \in \mathbb{R}$. By (2.96), the operator $\Phi_m(t)$ solves

$$\begin{cases} \partial_t \Phi_m(t) = \mathrm{i} a |D|^{\frac{1}{2}} \Phi_m(t) + A_m \Phi_m(t) \\ \Phi_m(0) = \mathrm{Id}\,, \end{cases} \qquad A_m := \big[\langle D\rangle^m, \mathrm{i} a |D|^{\frac{1}{2}}\big] \langle D\rangle^{-m}\,.$$

Then Duhamel's principle implies that

$$\Phi_m(t) = \Phi(t) + \Psi_m(t)\,, \quad \Psi_m(t) := \int_0^t \Phi(t-\tau) A_m \Phi_m(\tau)\, d\tau\,. \tag{2.102}$$

By (2.47), (2.40) and Lemma 2.11 (applied for $k_0 = 0$), we deduce that

$$|A_m|_{0,s,0} \lesssim_{s,m} \|a\|_{s+|m|+2}\,, \quad \forall s \geq s_0\,. \tag{2.103}$$

Applying (2.102), estimates (2.99), (2.103), and Lemma 2.29 (applied for $k_0 = 0$), for $\|a\|_{2s_0+\frac{1}{2}} \leq 1$, $\|a\|_{s_0+|m|+2} \leq 1$ we obtain, for all $s \geq s_0$,

$$\begin{aligned} \sup_{t\in[0,1]} \|\Phi_m(t) h\|_s \lesssim_{s,m} \|h\|_s + \|a\|_{s+s_0+\frac{1}{2}} \|h\|_{s_0} + \|a\|_{s_0+|m|+2} \sup_{t\in[0,1]} \|\Phi_m(t) h\|_s \\ + \big(\|a\|_{s+|m|+2} + \|a\|_{s+s_0+\frac{1}{2}}\big) \sup_{t\in[0,1]} \|\Phi_m(t) h\|_{s_0}\,. \end{aligned} \tag{2.104}$$

For $C(s_0, m)(\|a\|_{s_0+|m|+2} + \|a\|_{2s_0+\frac{1}{2}}) \leq \frac{1}{2}$ (which is implied by (2.100)), (2.104) at $s = s_0$ implies that $\sup_{t\in[0,1]} \|\Phi_m(t) h\|_{s_0} \lesssim_m \|h\|_{s_0}$. Plugging this bound in (2.104) gives, for $s \geq s_0$ and $C(s,m)\|a\|_{s_0+|m|+2} \leq \frac{1}{2}$ (see (2.100)),

$$\sup_{t\in[0,1]} \|\Phi_m(t) h\|_s \lesssim_{s,m} \|h\|_s + \|a\|_{s+\max\{s_0+\frac{1}{2}\,, |m|+2\}} \|h\|_{s_0}\,.$$

This proves (2.101) for $\beta = k = 0$, with $\sigma(0,0,m,-m) := \max\{s_0 + \frac{1}{2}\,, |m|+2\}$.

Proof of (2.101): induction step. Let us suppose that (2.101) holds for all $(k_1, \beta_1) \prec (k, \beta)$, $|k| \leq k_0$, $|\beta| \leq \beta_0$, $m_1, m_2 \in \mathbb{R}$ with $m_1 + m_2 = \frac{|\beta_1|+|k_1|}{2}$. We have to prove the claimed estimate for the operator $\langle D\rangle^{-m_1} \partial_\lambda^k \partial_\varphi^\beta \Phi \langle D\rangle^{-m_2}$, with $m_1, m_2 \in \mathbb{R}$, $m_1 + m_2 = \frac{|\beta|+|k|}{2}$. Differentiating (2.96) and using Duhamel's principle we get

$$\partial_\lambda^k \partial_\varphi^\beta \Phi(t) = \int_0^t \Phi(t-\tau) F_{\beta,k}(\tau)\, d\tau$$

where

$$F_{\beta,k}(\tau) := \sum_{\substack{k_1+k_2=k \\ \beta_1+\beta_2=\beta \\ (k_1,\beta_1)\prec(k,\beta)}} C(k_1,k_2,\beta_1,\beta_2) (\partial_\lambda^{k_2} \partial_\varphi^{\beta_2} a) |D|^{\frac{1}{2}} \partial_\lambda^{k_1} \partial_\varphi^{\beta_1} \Phi(\tau)\,.$$

For any $m_1, m_2 \in \mathbb{R}$ satisfying $m_1 + m_2 = \frac{|\beta|+|k|}{2}$, for any $t, \tau \in [0,1]$, we write

$$\langle D\rangle^{-m_1} \Phi(t-\tau) F_{\beta,k}(\tau) \langle D\rangle^{-m_2} = \langle D\rangle^{-m_1} \Phi(t-\tau) \langle D\rangle^{m_1} \langle D\rangle^{-m_1} F_{\beta,k}(\tau) \langle D\rangle^{-m_2}\,.$$

Then for any $k_1 + k_2 = k, \beta_1 + \beta_2 = \beta, (k_1, \beta_1) \prec (k, \beta)$ we write

$$\begin{aligned} &\langle D\rangle^{-m_1} (\partial_\lambda^{k_2} \partial_\varphi^{\beta_2} a) |D|^{\frac{1}{2}} \partial_\lambda^{k_1} \partial_\varphi^{\beta_1} \Phi(\tau) \langle D\rangle^{-m_2} \\ &= \langle D\rangle^{-m_1} (\partial_\lambda^{k_2} \partial_\varphi^{\beta_2} a) |D|^{\frac{1}{2}} \langle D\rangle^{-m_2+\frac{|\beta_1|+|k_1|}{2}} \langle D\rangle^{m_2-\frac{|\beta_1|+|k_1|}{2}} \partial_\lambda^{k_1} \partial_\varphi^{\beta_1} \Phi(\tau) \langle D\rangle^{-m_2}\,, \end{aligned}$$

and we have to estimate, uniformly in $t, \tau \in [0,1]$,

$$\Big(\langle D\rangle^{-m_1} \Phi(t-\tau) \langle D\rangle^{m_1}\Big) \Big(\langle D\rangle^{-m_1} (\partial_\lambda^{k_2} \partial_\varphi^{\beta_2} a) |D|^{\frac{1}{2}} \langle D\rangle^{-m_2+\frac{|\beta_1|+|k_1|}{2}}\Big) \Big(\langle D\rangle^{m_2-\frac{|\beta_1|+|k_1|}{2}} \partial_\lambda^{k_1} \partial_\varphi^{\beta_1} \Phi(\tau) \langle D\rangle^{-m_2}\Big)\,. \tag{2.105}$$

For any $s \geq s_0$, using the induction hyphothesis, one has that

$$\|\langle D\rangle^{-m_1}\Phi(t-\tau)\langle D\rangle^{m_1}h\|_s \lesssim_{s,m_1} \|h\|_s + \|a\|^{k_0,\gamma}_{s+\sigma(0,0,m_1,-m_1)}\|h\|_{s_0} \tag{2.106}$$

$$\|\langle D\rangle^{m_2-\frac{|\beta_1|+|k_1|}{2}}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\Phi(\tau)\langle D\rangle^{-m_2}h\|_s \lesssim_{s,\beta_0,k_0,m_2} \gamma^{-|k_1|}(\|h\|_s + \|a\|^{k_0,\gamma}_{s+\sigma(|\beta_1|,|k_1|,-m_2+\frac{|\beta_1|+|k_1|}{2},m_2)}\|h\|_{s_0})\,. \tag{2.107}$$

Using the fact that $m_1+m_2=\frac{|\beta|+|k|}{2}$, and that $(\beta_1,k_1)\prec(\beta,k)$, we obtain

$$-m_1-m_2+\frac{1}{2}+\frac{|\beta_1|+|k_1|}{2}=\frac{-|\beta|-|k|+1+|\beta_1|+|k_1|}{2}\leq 0\,.$$

Therefore the operator $\langle D\rangle^{-m_1}(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac{1}{2}}\langle D\rangle^{-m_2+\frac{|\beta_1|+|k_1|}{2}}$ belongs to OPS^0 and by Lemma 2.10, (2.40), (2.47), (2.41), we get

$$|\!|\langle D\rangle^{-m_1}(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}a)|D|^{\frac{1}{2}}\langle D\rangle^{-m_2+\frac{|\beta_1|+|k_1|}{2}}|\!|_{0,s,0} \lesssim_{s,\beta_0,k_0,m_1,m_2} \gamma^{-|k_2|}\|a\|^{k_0,\gamma}_{s+|m_1|+|\beta_2|}\,. \tag{2.108}$$

Applying (2.105)-(2.108), recalling Lemma 2.29, the smallness condition (2.100) and setting

$$\sigma(|\beta|,|k|,m_1,m_2):=\max_{(\beta_1,k_1)\prec(\beta,k)}\max\Big\{\sigma(0,0,m_1,-m_1),\,\sigma(|\beta_1|,|k_1|,-m_2+\frac{|\beta_1|+|k_1|}{2},m_2),\,|\beta_2|+|m_1|\Big\}$$

we deduce (2.101) for $h\in\mathcal{C}^\infty$. The thesis follows by density. ☐

Proposition 2.42. *Assume* (2.100). *For all $k\in\mathbb{N}^{\nu+1}$, $|k|\leq k_0$, $s\geq s_0$, the flow $\Phi(t):=\Phi(\lambda,\varphi,t)$ of* (2.95) *satisfies*

$$\sup_{t\in[0,1]}\|\partial_\lambda^k\Phi(t)h\|_s \lesssim_s \gamma^{-|k|}\big(\|h\|_{s+\frac{|k|}{2}}+\|a\|^{k_0,\gamma}_{s+s_0+|k|+1}\|h\|_{s_0+\frac{|k|}{2}}\big)\,, \tag{2.109}$$

$$\sup_{t\in[0,1]}\|\partial_\lambda^k(\Phi(t)-\mathrm{Id})h\|_s \lesssim_s \gamma^{-|k|}\big(\|a\|^{k_0,\gamma}_{s_0}\|h\|_{s+\frac{|k|+1}{2}}+\|a\|^{k_0,\gamma}_{s+s_0+k_0+\frac{3}{2}}\|h\|_{s_0+\frac{|k|+1}{2}}\big)\,. \tag{2.110}$$

Proof. See Proposition A.11 in [21]. Note that (2.101) with $m_1=0,\beta=0,m_2=|k|/2$, and h replaced by $\langle D\rangle^{m_2}h$ implies (2.109) with loss $\sigma(0,|k|,0,\frac{|k|}{2})$ instead of $s_0+|k|+1$. ☐

We consider also the dependence of the flow Φ with respect to the torus $i:=i(\varphi)$.

Proposition 2.43. *Let $s_1>s_0$, $\beta_0\in\mathbb{N}$. For any $\beta\in\mathbb{N}^\nu$, $|\beta|\leq\beta_0$, for any $m_1,m_2\in\mathbb{R}$ satisfying $m_1+m_2=\frac{|\beta|+1}{2}$ there exists a constant $\sigma(|\beta|)=\sigma(|\beta|,m_1,m_2)>0$ such that if $\|a\|_{s_1+\sigma(\beta_0)}\leq\delta(s_1)$ with $\delta(s_1)>0$ small enough, then the following estimate holds:*

$$\sup_{t\in[0,1]}\|\langle D\rangle^{-m_1}\partial_\varphi^\beta\Delta_{12}\Phi(t)\langle D\rangle^{-m_2}h\|_{s_1}\lesssim_{s_1}\|\Delta_{12}a\|_{s_1+\sigma(|\beta|)}\|h\|_{s_1}\,, \tag{2.111}$$

where $\Delta_{12}\Phi:=\Phi(i_2)-\Phi(i_1)$ and $\Delta_{12}a:=a(i_2)-a(i_1)$.

Proof. The proposition can be proved arguing as in the proof of Proposition 2.41. ☐

We also consider similar properties for the adjoint flow operator. Let $\Phi:=\Phi(1)$ denote the time-1 flow of (2.95) and Φ^* its adjoint with respect to the L^2 scalar product.

Proposition 2.44. (Adjoint) *Assume that $\|a\|^{k_0,\gamma}_{2s_0+\frac{5}{2}+k_0}\leq 1$, $\|a\|_{2s_0+1}\leq\delta(s)$ for some $\delta(s)>0$ small enough. Then for any $k\in\mathbb{N}^{\nu+1}$, $|k|\leq k_0$, for all $s\geq s_0$,*

$$\|(\partial_\lambda^k\Phi^*)h\|_s\lesssim_s\gamma^{-|k|}\big(\|h\|_{s+\frac{|k|}{2}}+\|a\|^{k_0,\gamma}_{s+s_0+|k|+\frac{3}{2}}\|h\|_{s_0+\frac{|k|}{2}}\big)$$
$$\|\partial_\lambda^k(\Phi^*-\mathrm{Id})h\|_s\lesssim_s\gamma^{-|k|}\big(\|a\|^{k_0,\gamma}_{s_0}\|h\|_{s+\frac{|k|+1}{2}}+\|a\|^{k_0,\gamma}_{s+s_0+|k|+2}\|h\|_{s_0+\frac{|k|+1}{2}}\big)\,.$$

Proof. See Proposition A.17 in [21]. □

Finally we estimate the variation of the adjoint operator Φ^* with respect to the torus $i(\varphi)$.

Proposition 2.45. *Let $s_1 > s_0$ and assume the condition $\|a\|_{s_1+s_0+3} \leq 1$, $\|a\|_{s_1+s_0+1} \leq \delta(s_1)$, for some $\delta(s_1) > 0$ small. Then, for all $s \in [s_0, s_1]$,*

$$\|\Delta_{12}\Phi^* h\|_s \lesssim_s \|\Delta_{12} a\|_{s+s_0+\frac{1}{2}} \|h\|_{s+\frac{1}{2}} .$$

Proof. It follows by Proposition A.18 in [21]. □

3 Dirichlet-Neumann operator

We collect some fundamental properties of the Dirichlet-Neumann operator $G(\eta)$, defined in (1.5), which are used in the paper.

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 is called the "shape derivative", and it is given by (see e.g. [45], [46])

$$G'(\eta)[\hat{\eta}]\psi = \lim_{\epsilon \to 0} \frac{1}{\epsilon}\{G(\eta + \epsilon\hat{\eta})\psi - G(\eta)\psi\} = -G(\eta)(B\hat{\eta}) - \partial_x(V\hat{\eta}) \tag{3.1}$$

where

$$B := B(\eta, \psi) := \frac{\eta_x \psi_x + G(\eta)\psi}{1 + \eta_x^2}, \qquad V := V(\eta, \psi) := \psi_x - B\eta_x . \tag{3.2}$$

It turns out that $(V, B) = \nabla_{x,y}\Phi$ is the velocity field evaluated at the free surface $(x, \eta(x))$. The operator $G(\eta)$ is even according to Definition 2.20.

Let $\eta \in \mathcal{C}^\infty(\mathbb{T})$. It is well-known (see e.g. [46], [5], [39]) that the Dirichlet-Neumann operator is a *pseudo-differential* operator of the form

$$G(\eta) = G(0) + \mathcal{R}_G(\eta), \qquad \text{where} \quad G(0) = |D| \tanh(\mathtt{h}|D|) \tag{3.3}$$

is the Dirichlet-Neumann operator at the flat surface $\eta(x) = 0$ and the remainder $\mathcal{R}_G(\eta)$ is in $OPS^{-\infty}$ and it is $O(\eta)$-small. Note that the profile $\eta(x) := \eta(\omega, \mathtt{h}, \varphi, x)$, as well as the velocity potential at the free surface $\psi(x) := \psi(\omega, \mathtt{h}, \varphi, x)$, may depend on the angles $\varphi \in \mathbb{T}^\nu$ and the parameters $\lambda := (\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$.

In Proposition 3.1 we prove formula (3.3) and we provide the quantitative estimate (3.5). For simplicity of notation we sometimes omit to write the dependence with respect to φ and λ. For the sequel, it is useful to introduce the following notation. Let X and Y be Banach spaces and $B \subset X$ be a bounded open set. We denote by $\mathcal{C}_b^1(B, Y)$ the space of the $\mathcal{C}^1$ functions $B \to Y$ bounded and with bounded derivatives.

Proposition 3.1. (Dirichlet-Neumann) *Assume that $\partial_\lambda^k \eta(\lambda, \cdot, \cdot)$ is $\mathcal{C}^\infty$ for all $|k| \leq k_0$. There exists $\delta(s_0, k_0) > 0$ such that, if*

$$\|\eta\|_{2s_0+2k_0+1}^{k_0,\gamma} \leq \delta(s_0, k_0) , \tag{3.4}$$

then the Dirichlet-Neumann operator $G(\eta)$ may be written as in (3.3) where $\mathcal{R}_G(\eta)$ is an integral operator with $\mathcal{C}^\infty$ kernel K_G (see (2.56)) which satisfies, for all $m, s, \alpha \in \mathbb{N}$, the estimate

$$|\mathcal{R}_G(\eta)|_{-m,s,\alpha}^{k_0,\gamma} \leq C(s, m, \alpha, k_0)\|K_G\|_{\mathcal{C}^{s+m+\alpha}}^{k_0,\gamma} \leq C(s, m, \alpha, k_0)\|\eta\|_{s+2s_0+2k_0+m+\alpha+3}^{k_0,\gamma} . \tag{3.5}$$

Let $s_1 \geq 2s_0 + 1$. 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 $\mathcal{C}_b^1$.

The rest of this section is devoted to the proof of Proposition 3.1.

In order to analyze the Dirichlet-Neumann operator $G(\eta)$ it is convenient to transform the boundary value problem (1.3) (with $h = \mathtt{h}$) defined in the closure of the free domain $\mathcal{D}_\eta = \{(x, y) : -\mathtt{h} < y < \eta(x)\}$ into an elliptic problem in a flat lower strip

$$\big\{(X, Y) : -\mathtt{h} - c \leq Y \leq 0\big\} , \tag{3.6}$$

via a conformal diffeomorphism (close to the identity for η small) of the form

$$x = U(X, Y) = X + p(X, Y), \quad y = V(X, Y) = Y + q(X, Y) . \tag{3.7}$$

Remark 3.2. *The requirement that* (3.7) *is a conformal map implies that the system obtained transforming* (1.3) *is simply* (3.45) *(the Laplace operator and the Neumann boundary conditions are transformed into itselves).*

We require that

$$X \mapsto q(X,Y)\,,\ p(X,Y) \quad \text{are } 2\pi\text{−periodic}\,, \tag{3.8}$$

so that (3.7) defines a diffeomorphism between the cylinder $\mathbb{T}\times[-\mathtt{h}-c,0]$ and $\mathcal{D}_\eta$. The bottom $\{Y=-\mathtt{h}-c\}$ is transformed in the bottom $\{y=-\mathtt{h}\}$ if

$$V(X,-\mathtt{h}-c) = -\mathtt{h} \qquad \Leftrightarrow \qquad q(X,-\mathtt{h}-c) = c\,, \quad \forall X \in \mathbb{R}\,, \tag{3.9}$$

and the boundary $\{Y=0\}$ is transformed in the free surface $\{y=\eta(x)\}$ if

$$V(X,0) = \eta(U(X,0)) \qquad \Leftrightarrow \qquad q(X,0) = \eta(X+p(X,0))\,. \tag{3.10}$$

The diffeomorphism (3.7) is conformal if and only if the map

$$U(X,Y) + \mathrm{i}V(X,Y)$$

is analytic, which amounts to the Cauchy-Riemann equations

$$U_X = V_Y\ ,\ U_Y = -V_X\ , \qquad i.e. \quad p_X = q_Y\ ,\ p_Y = -q_X\ . \tag{3.11}$$

The functions (U,V), i.e. (p,q), are harmonic conjugate. Moreover, (3.9) and (3.11) imply that

$$U_Y(X,-\mathtt{h}-c) = p_Y(X,-\mathtt{h}-c) = 0\,. \tag{3.12}$$

The most general function p which is harmonic, namely $\Delta p = 0$, and satisfies (3.8) and (3.12) is

$$p(X,Y) = \beta_0 + \sum_{k\neq 0} \beta_k \cosh(|k|(Y+\mathtt{h}+c))e^{\mathrm{i}kX}$$

where $\beta_0\in\mathbb{R}$, $\beta_k\in\mathbb{C}$, $k\in\mathbb{Z}\setminus\{0\}$, are fixed by specifying the value of p at the boundary $\{Y=0\}$, namely

$$p(X,0) = \mathtt{p}(X) = \mathtt{p}_0 + \sum_{k\neq 0} \mathtt{p}_k e^{\mathrm{i}kX}\,. \tag{3.13}$$

As a consequence, the solution $p(X,Y)$ of

$$\Delta p = 0\,, \quad p(X,0) = \mathtt{p}(X), \quad p_Y(X,-\mathtt{h}-c) = 0\,, \quad 2\pi\text{-periodic in } X, \tag{3.14}$$

is

$$p(X,Y) = \sum_{k\in\mathbb{Z}} \mathtt{p}_k \frac{\cosh(|k|(Y+\mathtt{h}+c))}{\cosh(|k|(\mathtt{h}+c))}\, e^{\mathrm{i}kX}\,. \tag{3.15}$$

The most general function q which is harmonic, namely $\Delta q = 0$, and satisfies (3.8) and (3.9) is

$$q(X,Y) = \alpha_0 + \frac{\alpha_0 - c}{\mathtt{h}+c} Y + \sum_{k\neq 0} \alpha_k \sinh(|k|(Y+\mathtt{h}+c))e^{\mathrm{i}kX}\,, \tag{3.16}$$

where $\alpha_0\in\mathbb{R}$, $\alpha_k\in\mathbb{C}$, $k\in\mathbb{Z}\setminus\{0\}$. By (3.11), (3.15), (3.16) we get

$$\alpha_0 = c\,, \quad \alpha_k = \mathrm{i}\mathtt{p}_k \frac{\mathrm{sign}(k)}{\cosh(|k|(\mathtt{h}+c))}\,,$$

so that q is uniquely determined as

$$q(X,Y) = c + \sum_{k\neq 0} \mathrm{i}\mathtt{p}_k \frac{\mathrm{sign}(k)}{\cosh(|k|(\mathtt{h}+c))} \sinh(|k|(Y+\mathtt{h}+c))e^{\mathrm{i}kX}\,. \tag{3.17}$$

We still have to impose the condition (3.10). By (3.17) we have

$$q(X,0) = c + \sum_{k\neq 0} \mathrm{i}\,\mathrm{sign}(k)\tanh(|k|(\mathtt{h}+c))\mathtt{p}_k e^{\mathrm{i}kX} = c - \mathcal{H}\tanh((\mathtt{h}+c)|D|)\mathtt{p}(X) \tag{3.18}$$

where $\mathtt{p}(X)$ is defined in (3.13) and $\mathcal{H}$ is the Hilbert transform defined as the Fourier multiplier in (2.34).

By (3.13), (3.18), the condition (3.10) amounts to solve

$$c - \mathcal{H}\tanh((\mathtt{h}+c)|D|)\mathtt{p}(X) = \eta(X+\mathtt{p}(X)). \tag{3.19}$$

Remark 3.3. *If we had required $c = 0$ (fixing the strip of the straight domain (3.6)), equation (3.19) would, in general, have no solution. For example, if $\eta(x) = \eta_0 \neq 0$, then $-\mathcal{H}\tanh((\mathtt{h}+c)|D|)\mathtt{p}(X) = \eta_0$ has no solutions because the left hand side has zero average while the right hand side has average $\eta_0 \neq 0$.*

Since the range of $\mathcal{H}$ are the functions with zero average, equation (3.19) is equivalent to

$$c = \langle \eta(X+\mathtt{p}(X))\rangle, \quad -\mathcal{H}\tanh((\mathtt{h}+c)|D|)\mathtt{p}(X) = \pi_0^{\perp}\eta(X+\mathtt{p}(X)) \tag{3.20}$$

where $\langle f\rangle = f_0 = \pi_0 f$ is the average in X of any function f, π_0 is defined in (2.35), and $\pi_0^{\perp} := \mathrm{Id} - \pi_0$. We look for a solution $(c(\varphi), \mathtt{p}(\varphi, X))$, where $\mathtt{p}$ has zero average in X, of the system

$$c = \langle \eta(X+\mathtt{p}(X))\rangle, \quad \mathtt{p}(X) = \frac{\mathcal{H}}{\tanh((\mathtt{h}+c)|D|)}[\eta(X+\mathtt{p}(X))]. \tag{3.21}$$

Since $\mathcal{H}^2 = -\pi_0^{\perp}$, if $\mathtt{p}$ solves the second equation in (3.21), then $\mathtt{p}$ is also a solution of the second equation in (3.20).

Lemma 3.4. *Let $\eta(\lambda,\varphi,x)$ satisfy $\partial_\lambda^k \eta(\lambda,\cdot,\cdot) \in \mathcal{C}^\infty(\mathbb{T}^{\nu+1})$ for all $|k| \leq k_0$. There exists $\delta(s_0,k_0) > 0$ such that, if $\|\eta\|_{2s_0+k_0+2}^{k_0,\gamma} \leq \delta(s_0,k_0)$, then there exists a unique $\mathcal{C}^\infty$ solution $(c(\eta), \mathtt{p}(\eta))$ of system (3.21) satisfying*

$$\|\mathtt{p}\|_s^{k_0,\gamma}, \|c\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|\eta\|_{s+k_0}^{k_0,\gamma}, \quad \forall s \geq s_0. \tag{3.22}$$

Moreover, let $s_1 \geq 2s_0+1$. There exists $\delta(s_1) > 0$ such that the map $\{\|\eta\|_{s_1+2} < \delta(s_1)\} \to H_\varphi^{s_1} \times H^{s_1}$, $\eta \mapsto (c(\eta), \mathtt{p}(\eta))$ is $\mathcal{C}_b^1$.

Proof. We look for a fixed point of the map

$$\Phi(\mathtt{p}) := \mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)[\eta(\cdot + \mathtt{p}(\cdot))], \qquad \text{where} \quad \mathtt{f}(\xi) := \frac{1}{\tanh(\xi)}, \quad \xi \neq 0, \tag{3.23}$$

and $c := \langle \eta(X+\mathtt{p}(X))\rangle$. We are going to prove that Φ is a contraction in a ball $\mathcal{B}_{2s_0+1}(r) := \{\|\mathtt{p}\|_{2s_0+1}^{k_0,\gamma} \leq r, \langle \mathtt{p}\rangle = 0\}$ with radius r small enough. We begin by proving some preliminary estimates.

The operator $\mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)$ is the Fourier multiplier, acting on the periodic functions, with symbol

$$-\mathrm{i}\,\mathrm{sign}(\xi)\chi(\xi)\mathtt{f}\big((\mathtt{h}+c(\lambda,\varphi))|\xi|\big) =: g(\mathtt{h}+c(\lambda,\varphi),\xi), \quad \text{where} \quad g(y,\xi) := -\mathrm{i}\,\mathrm{sign}(\xi)\chi(\xi)\mathtt{f}(y|\xi|) \quad \forall y > 0,$$

where the cut-off $\chi(\xi)$ is defined in (2.16). For all $n \in \mathbb{N}$, one has

$$\partial_y^n g(y,\xi) = -\mathrm{i}\,\mathrm{sign}(\xi)\chi(\xi)\mathtt{f}^{(n)}(y|\xi|)|\xi|^n. \tag{3.24}$$

For all $x \in \mathbb{R}\setminus\{0\}$, denoting, in short, $T := \tanh(x)$, one has $\mathtt{f}'(x) = -T^{-2}(1-T^2)$, $\mathtt{f}''(x) = 2T^{-3}(1-T^2)$, and, by induction, $\mathtt{f}^{(n)}(x) = P_n(T^2)T^{-n-1}(1-T^2)$ for all $n \geq 2$, where P_n is a polynomial of degree $n-2$. Since $1-\tanh^2(x)$ vanishes exponentially as $x \to +\infty$, for every $\rho > 0$, $n \in \mathbb{N}$, there exists a constant $C(n,\rho) > 0$ such that

$$|\mathtt{f}^{(n)}(x)|x^n \leq C(n,\rho), \quad \forall x \geq \rho. \tag{3.25}$$

Since $\chi(\xi) = 0$ for $|\xi| \leq 1/3$, by (3.24) and (3.25) (with $\rho = \mathtt{h}_1/6$) we deduce that for every $n \in \mathbb{N}$ there exists a constant $C_n(\mathtt{h}_1) > 0$ such that

$$|\partial_y^n g(y,\xi)| \leq C_n(\mathtt{h}_1), \quad \forall y \geq \mathtt{h}_1/2, \ \forall \xi \in \mathbb{R}. \tag{3.26}$$

We consider a smooth extension $\tilde{g}(y,\xi)$ of $g(y,\xi)$, defined for any $(y,\xi)\in\mathbb{R}\times\mathbb{R}$, satisfying the same bound (3.26). Now $|c(\lambda,\varphi)|\le\|\eta\|_{L^\infty}\le C\|\eta\|_{s_0}$, and therefore $\mathtt{h}+c(\lambda,\varphi)\ge\mathtt{h}_1/2$ for all λ,φ if $\|\eta\|_{s_0}$ is sufficiently small. Then, by Lemma 2.6, the composition $\tilde{g}(\mathtt{h}+c(\lambda,\varphi),\xi)$ satisfies

$$\|\tilde{g}(\mathtt{h}+c,\xi)\|_s^{k_0,\gamma}\lesssim_{s,k_0,\mathtt{h}_1,\mathtt{h}_2}1+\|c\|_s^{k_0,\gamma}$$

uniformly in $\xi\in\mathbb{R}$ (the dependence on $\mathtt{h}_1,\mathtt{h}_2$ is omitted in the sequel, as usual). As a consequence, we have the following estimates for pseudo-differential norms (recall Definition 2.9) of the Fourier multiplier in (3.23): for all $s\ge s_0$,

$$|\!|\!|\mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)|\!|\!|_{0,s,0}^{k_0,\gamma}\,,\ |\!|\!|\mathcal{H}|D|\mathtt{f}'\big((\mathtt{h}+c)|D|\big)|\!|\!|_{0,s,0}^{k_0,\gamma}\lesssim_{s,k_0}1+\|c\|_s^{k_0,\gamma}\,. \tag{3.27}$$

Estimate (2.11) with $k+1=k_0$ implies that, for $\|\mathtt{p}\|_{2s_0+1}^{k_0,\gamma}\le\delta(s_0,k_0)$, the function $c\equiv c(\eta,\mathtt{p})=\langle\eta(X+\mathtt{p}(X))\rangle$ satisfies, for all $s\ge s_0$,

$$\|c\|_s^{k_0,\gamma}\lesssim_{s,k_0}\|\eta\|_{s+k_0}^{k_0,\gamma}+\|\mathtt{p}\|_s^{k_0,\gamma}\|\eta\|_{s_0+k_0+1}^{k_0,\gamma}\,. \tag{3.28}$$

Therefore by (3.27), (3.28) we get, for all $s\ge s_0$,

$$|\!|\!|\mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)|\!|\!|_{0,s,0}^{k_0,\gamma},\ |\!|\!|\mathcal{H}|D|\mathtt{f}'\big((\mathtt{h}+c)|D|\big)|\!|\!|_{0,s,0}^{k_0,\gamma}\lesssim_{s,k_0}1+\|\eta\|_{s+k_0}^{k_0,\gamma}+\|\mathtt{p}\|_s^{k_0,\gamma}\|\eta\|_{s_0+k_0+1}^{k_0,\gamma}\,. \tag{3.29}$$

Now we prove that Φ is a contraction in the ball $\mathcal{B}_{2s_0+1}(r):=\{\|\mathtt{p}\|_{2s_0+1}^{k_0,\gamma}\le r,\ \langle\mathtt{p}\rangle=0\}$.

STEP 1: CONTRACTION IN LOW NORM. For any $\|\mathtt{p}\|_{2s_0+1}^{k_0,\gamma}\le r\le\delta(s_0,k_0)$, by (2.77), (3.29), (2.11), and using the bound $\|\eta\|_{s_0+k_0+1}^{k_0,\gamma}\le1$, we have, $\forall s\ge s_0$,

$$\|\Phi(\mathtt{p})\|_s^{k_0,\gamma}\lesssim_{s,k_0}\|\eta\|_{s+k_0}^{k_0,\gamma}+\|\eta\|_{s_0+k_0+1}^{k_0,\gamma}\|\mathtt{p}\|_s^{k_0,\gamma}\,. \tag{3.30}$$

In particular, (3.30) at $s=2s_0+1$ gives

$$\|\Phi(\mathtt{p})\|_{2s_0+1}^{k_0,\gamma}\le C(s_0,k_0)\big(\|\eta\|_{2s_0+k_0+1}^{k_0,\gamma}+\|\eta\|_{s_0+k_0+1}^{k_0,\gamma}\|\mathtt{p}\|_{2s_0+1}^{k_0,\gamma}\big). \tag{3.31}$$

We fix $r:=2C(s_0,k_0)\|\eta\|_{2s_0+k_0+1}^{k_0,\gamma}$ and we assume that $r\le1$. Then, by (3.31), Φ maps the ball $\mathcal{B}_{2s_0+1}(r)$ into itself. To prove that Φ is a contraction in this ball, we estimate its differential at any $\mathtt{p}\in\mathcal{B}_{2s_0+1}(r)$ in the direction $\tilde{\mathtt{p}}$, which is

$$\Phi'(\mathtt{p})[\tilde{\mathtt{p}}]=\mathcal{A}(\mathtt{m}\,\tilde{\mathtt{p}})\,, \tag{3.32}$$

where the operator $\mathcal{A}$ and the function $\mathtt{m}$ are

$$\mathcal{A}(h):=\langle h\rangle\mathcal{H}\mathtt{f}'((\mathtt{h}+c)|D|)|D|[\eta(X+\mathtt{p}(X))]+\mathcal{H}\mathtt{f}((\mathtt{h}+c)|D|)[h],\quad \mathtt{m}:=\eta_x(X+\mathtt{p}(X))\,. \tag{3.33}$$

To obtain (3.32)-(3.33), note that $\partial_{\mathtt{p}}c[\tilde{\mathtt{p}}]=\langle\mathtt{m}\tilde{\mathtt{p}}\rangle$. By (2.11), for all $s\ge s_0$,

$$\|\mathtt{m}\|_s^{k_0,\gamma}\lesssim_{s,k_0}\|\eta\|_{s+k_0+1}^{k_0,\gamma}+\|\mathtt{p}\|_s^{k_0,\gamma}\|\eta\|_{s_0+k_0+2}^{k_0,\gamma}\,. \tag{3.34}$$

By (2.77), (3.29), (2.11), using the bounds $\|\eta\|_{s_0+k_0+1}^{k_0,\gamma}\le1$ and $\|\mathtt{p}\|_{s_0}^{k_0,\gamma}\le1$, we get, for all $s\ge s_0$,

$$|\!|\!|\mathcal{A}|\!|\!|_{0,s,0}^{k_0,\gamma}\lesssim_{s,k_0}1+\|\eta\|_{s+k_0}^{k_0,\gamma}+\|\mathtt{p}\|_s^{k_0,\gamma}\|\eta\|_{s_0+k_0+1}^{k_0,\gamma}\,. \tag{3.35}$$

By (3.32), (2.45), (3.34), (3.35) we deduce that, for all $s\ge s_0$,

$$|\!|\!|\Phi'(\mathtt{p})|\!|\!|_{0,s,0}^{k_0,\gamma}\lesssim_{s,k_0}\|\eta\|_{s+k_0+1}^{k_0,\gamma}+\|\mathtt{p}\|_s^{k_0,\gamma}\|\eta\|_{s_0+k_0+2}^{k_0,\gamma}\,. \tag{3.36}$$

In particular, by (3.36) at $s=2s_0+1$, and (2.77), we get

$$\|\Phi'(\mathtt{p})[\tilde{\mathtt{p}}]\,\|_{2s_0+1}^{k_0,\gamma}\le C(s_0,k_0)\|\eta\|_{2s_0+k_0+2}^{k_0,\gamma}\|\tilde{\mathtt{p}}\|_{2s_0+1}^{k_0,\gamma}\le\frac{1}{2}\|\tilde{\mathtt{p}}\|_{2s_0+1}^{k_0,\gamma} \tag{3.37}$$

provided $C(s_0,k_0)\|\eta\|^{k_0,\gamma}_{2s_0+k_0+2} \le 1/2$. Thus Φ is a contraction in the ball $\mathcal{B}_{2s_0+1}(r)$ and, by the contraction mapping theorem, there exists a unique fixed point $\mathtt{p} = \Phi(\mathtt{p})$ in $\mathcal{B}_{2s_0+1}(r)$. Moreover, by (3.30), there is $C(s_0,k_0) > 0$ such that, for all $s \in [s_0, 2s_0+1]$,

$$\|\mathtt{p}\|^{k_0,\gamma}_s = \|\Phi(\mathtt{p})\|^{k_0,\gamma}_s \le C(s_0,k_0)\|\eta\|^{k_0,\gamma}_{s+k_0} + C(s_0,k_0)\|\eta\|^{k_0,\gamma}_{s_0+k_0+1}\|\mathtt{p}\|^{k_0,\gamma}_s$$

and, for $C(s_0,k_0)\|\eta\|^{k_0,\gamma}_{s_0+k_0+1} \le 1/2$, we deduce the estimate $\|\mathtt{p}\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0}$ for all $s \in [s_0, 2s_0+1]$. By (3.28), using that $\|\eta\|^{k_0,\gamma}_{s_0+k_0+1} \le 1$, we obtain $\|c\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0}$ for all $s \in [s_0, 2s_0+1]$. Thus we have proved (3.22) for all $s \in [s_0, 2s_0+1]$.

STEP 2: REGULARITY. Now we prove that $\mathtt{p}$ is $\mathcal{C}^\infty$ in (φ, x) and we estimate the norm $\|\mathtt{p}\|^{k_0,\gamma}_s$ as in (3.22) arguing by induction on s. Assume that, for a given $s \ge 2s_0+1$, we have already proved that

$$\|\mathtt{p}\|^{k_0,\gamma}_s\,,\ \|c\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0}\,. \tag{3.38}$$

We want to prove that (3.38) holds for $s+1$. We have to estimate $\|\mathtt{p}\|^{k_0,\gamma}_{s+1} \simeq \max\{\|\mathtt{p}\|^{k_0,\gamma}_s,\ \|\partial_X \mathtt{p}\|^{k_0,\gamma}_s,\ \|\partial_{\varphi_i}\mathtt{p}\|^{k_0,\gamma}_s,\ i=1,\dots,\nu\}$. Using the definition (3.23) of Φ, we derive explicit formulas for the derivatives $\partial_X \mathtt{p}, \partial_{\varphi_i}\mathtt{p}$ in terms of $\mathtt{p}, \eta, \partial_x \eta, \partial_{\varphi_i}\eta$. Differentiating the identity $\mathtt{p} = \Phi(\mathtt{p})$ with respect to X we get

$$\mathtt{p}_X = \mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)[\eta_x(X+\mathtt{p}(X))(1+\mathtt{p}_X)] = \Phi'(\mathtt{p})[\mathtt{p}_X] + \mathcal{A}(\mathtt{m}) \tag{3.39}$$

where the operator $\Phi'(\mathtt{p})$ is given by (3.32) and $\mathcal{A}, \mathtt{m}$ are defined in (3.33) (note that $\langle \eta_x(X+\mathtt{p}(X))(1+\mathtt{p}_X(X))\rangle = 0$). By (3.36) at $s=s_0$, for $\|\eta\|^{k_0,\gamma}_{s_0+k_0+2} \le \delta(s_0,k_0)$ small enough, condition (2.54) for $A = -\Phi'(\mathtt{p})$ (with $\alpha = 0$) holds. Therefore the operator $\mathrm{Id} - \Phi'(\mathtt{p})$ is invertible and, by (2.55) (with $\alpha=0$), (3.38) and (2.77), its inverse satisfies, for all $s \ge s_0$,

$$\|(\mathrm{Id}-\Phi'(\mathtt{p}))^{-1}h\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|h\|^{k_0,\gamma}_s + \|\eta\|^{k_0,\gamma}_{s+k_0+1}\|h\|^{k_0,\gamma}_{s_0}\,. \tag{3.40}$$

By (3.39), we deduce that $\mathtt{p}_X = (\mathrm{Id}-\Phi'(\mathtt{p}))^{-1}\mathcal{A}(\mathtt{m})$. By (2.77), (3.34)-(3.35) and (3.38), we get $\|\mathcal{A}(\mathtt{m})\|^{k_0,\gamma}_s \lesssim_s \|\eta\|^{k_0,\gamma}_{s+k_0+1}$. Hence, by (3.40), using $\|\eta\|^{k_0,\gamma}_{s_0+k_0+2} \le 1$, we get

$$\|\mathtt{p}_X\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0+1}\,. \tag{3.41}$$

Differentiating the identity $\mathtt{p} = \Phi(\mathtt{p})$ with respect to φ_i, $i=1,\dots,\nu$, by (3.23) we get

$$\begin{aligned}\partial_{\varphi_i}\mathtt{p} &= (\partial_{\varphi_i}c)\mathcal{H}\mathtt{f}'\big((\mathtt{h}+c)|D|\big)|D|[\eta(X+\mathtt{p}(X))] + \mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)[(\partial_{\varphi_i}\eta)(X+\mathtt{p}(X))] \\ &\quad + \mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)[\eta_x(X+\mathtt{p}(X))\partial_{\varphi_i}\mathtt{p}] \\ &= \Phi'(\mathtt{p})[\partial_{\varphi_i}\mathtt{p}] + \mathcal{A}[(\partial_{\varphi_i}\eta)(X+\mathtt{p}(X))]\end{aligned} \tag{3.42}$$

where $\mathcal{A}$ is defined in (3.33). To get (3.42) we have used that $\partial_{\varphi_i}c = \langle(\partial_{\varphi_i}\eta)(\cdot+\mathtt{p}(\cdot))\rangle + \langle \eta_x(\cdot+\mathtt{p}(\cdot))\partial_{\varphi_i}\mathtt{p}\rangle$. Therefore $\partial_{\varphi_i}\mathtt{p} = (\mathrm{Id}-\Phi'(\mathtt{p}))^{-1}\mathcal{A}[(\partial_{\varphi_i}\eta)(X+\mathtt{p}(X))]$ and, by (3.40), (3.35), (2.11), (2.77), (3.38), we get

$$\|\partial_{\varphi_i}\mathtt{p}\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0+1}, \quad i=1,\dots,\nu\,. \tag{3.43}$$

Thus (3.38), (3.41) and (3.43) imply (3.38) at $s+1$ for $\mathtt{p}$. By (3.28), the same estimate holds for c, and the induction step is proved. This completes the proof of (3.22).

The fact that the map $\{\|\eta\|_{s_1+2} < \delta(s_1)\} \to H^{s_1}_\varphi \times H^{s_1}$ defined by $\eta \mapsto (c(\eta), \mathtt{p}(\eta))$ is $\mathcal{C}^1$ follows by the implicit function theorem using the $\mathcal{C}^1$ map

$$\begin{aligned} &F : H^{s_1+2}(\mathbb{T}^{\nu+1}) \times H^{s_1}_\varphi(\mathbb{T}^\nu) \times H^{s_1}(\mathbb{T}^{\nu+1}) \to H^{s_1}_\varphi(\mathbb{T}^\nu) \times H^{s_1}(\mathbb{T}^{\nu+1})\,, \\ &F(\eta, c, \mathtt{p})(\varphi, X) := \begin{pmatrix} c - \langle \eta(X+\mathtt{p}(X))\rangle \\ \mathtt{p}(\varphi,X) - \dfrac{\mathcal{H}}{\tanh((\mathtt{h}+c)|D|)}[\eta(\varphi, X+\mathtt{p}(\varphi,X))] \end{pmatrix}. \end{aligned}$$

Since $F(0,0,0) = 0$ and $\partial_{(c,\mathtt{p})}F(0,0,0) = \mathrm{Id}$, by the implicit function theorem there exists $\delta(s_1) > 0$ and a $\mathcal{C}^1$ map $\{\|\eta\|_{s_1+2} \le \delta(s_1)\} \to H^{s_1}_\varphi(\mathbb{T}^\nu) \times H^{s_1}(\mathbb{T}^{\nu+1})$, $\eta \mapsto (c(\eta), \mathtt{p}(\eta))$, such that $F(\eta, c(\eta), \mathtt{p}(\eta)) = 0$. Moreover it can be proved that the map $\eta \mapsto (c(\eta), \mathtt{p}(\eta))$ is $\mathcal{C}^1_b$ using that F is bounded with all its derivatives on any bounded subset of the space $H^{s_1+2}(\mathbb{T}^{\nu+1}) \times H^{s_1}_\varphi(\mathbb{T}^\nu) \times H^{s_1}(\mathbb{T}^{\nu+1})$. $\square$

Notice that (3.4) implies the smallness condition of Lemma 3.4. We have proved the following:

Lemma 3.5. (Conformal diffeomorphism) *Assume* (3.4). *Then the transformation*

$$\begin{aligned} U(X,Y) &:= X + \sum_{k\neq 0} \mathtt{p}_k \frac{\cosh(|k|(Y+\mathtt{h}+c))}{\cosh(|k|(\mathtt{h}+c))} e^{\mathrm{i}kX} \\ V(X,Y) &:= Y + c + \sum_{k\neq 0} \mathrm{i}\mathtt{p}_k \frac{\mathrm{sign}(k)}{\cosh(|k|(\mathtt{h}+c))} \sinh(|k|(Y+\mathtt{h}+c)) e^{\mathrm{i}kX} \end{aligned} \tag{3.44}$$

where c and $\mathtt{p}$ are the solutions of (3.21) *provided by Lemma 3.4, is a conformal diffeomorphism between the cylinder $\mathbb{T}\times[-\mathtt{h}-c,0]$ and $\mathcal{D}_\eta$. The conditions* (3.9), (3.10) *hold: the bottom $\{Y=-\mathtt{h}-c\}$ is transformed into the bottom $\{y=-\mathtt{h}\}$ and the boundary $\{Y=0\}$ is transformed into the free surface $\{y=\eta(x)\}$.*

We transform (1.3) via the conformal diffeomorphism (3.44). Denote

$$(Pu)(X) := u(X+\mathtt{p}(X)) .$$

The velocity potential

$$\phi(X,Y) := \Phi(U(X,Y),V(X,Y))$$

satisfies, using the Cauchy-Riemann equations (3.11), and (3.9)-(3.12),

$$\Delta\phi = 0 \text{ in } \{-\mathtt{h}-c<Y<0\}\,, \quad \phi(X,0) = (P\psi)(X)\,, \quad \phi_Y(X,-\mathtt{h}-c) = 0\,. \tag{3.45}$$

We calculate explicitly the solution ϕ of (3.45), which is (see (3.15))

$$\phi(X,Y) = \sum_{k\in\mathbb{Z}} \widehat{(P\psi)}_k \frac{\cosh(|k|(Y+\mathtt{h}+c))}{\cosh(|k|(\mathtt{h}+c))} e^{\mathrm{i}kX}\,,$$

where $\widehat{(P\psi)}_k$ denotes the k-th Fourier coefficient of the periodic function $P\psi$. Therefore the Dirichlet-Neumann operator in the domain $\{-\mathtt{h}-c\le Y\le 0\}$ at the flat surface $Y=0$ is given by

$$\phi_Y(X,0) = \sum_{k\neq 0} \widehat{(P\psi)}_k \tanh(|k|(\mathtt{h}+c))|k| e^{\mathrm{i}kX} = |D|\tanh((\mathtt{h}+c)|D|)(P\psi)(X)\ . \tag{3.46}$$

Lemma 3.6. $G(\eta) = \partial_x P^{-1}\mathcal{H}\tanh((\mathtt{h}+c)|D|)\,P$.

Proof. Using (3.11), we have

$$\Phi_x = \frac{\phi_X U_X + \phi_Y U_Y}{U_X^2+U_Y^2}\,, \quad \Phi_y = \frac{\phi_Y U_X - \phi_X U_Y}{U_X^2+U_Y^2}\,. \tag{3.47}$$

Moreover, since $V(X,0)=\eta(U(X,0))$ (see (3.10)) we derive that

$$-U_Y(X,0) = V_X(X,0) = \eta_x(U(X,0))U_X(X,0)\,. \tag{3.48}$$

By the definition (1.5) of the Dirichlet-Neumann operator we get, at $x=U(X,0)$,

$$\begin{aligned} G(\eta)\psi(x) &\overset{(1.5),(3.47)}{=} \frac{1}{U_X^2+U_Y^2}\Big(\phi_X(-U_Y-\eta_x U_X)+\phi_Y(U_X-\eta_x U_Y)\Big)_{|Y=0} \\ &\overset{(3.48)}{=} \frac{1}{U_X(X,0)}\,\phi_Y(X,0) \overset{(3.13),(3.46)}{=} \frac{1}{1+\mathtt{p}_X(X)}\,|D|\tanh((\mathtt{h}+c)|D|)(P\psi)(X) \\ &= \Big\{\frac{1}{1+\mathtt{p}_X}\,|D|\tanh((\mathtt{h}+c)|D|)P\psi\Big\}(x+\breve{\mathtt{p}}(x)) \end{aligned}$$

where $X = x + \breve{\mathtt{p}}(x)$ is the inverse diffeomorphism of $x = X + \mathtt{p}(X)$. In operatorial notation, and writing $|D| = \mathcal{H}\partial_X$ we have

$$
\begin{aligned}
G(\eta) &= P^{-1}\frac{1}{1+\mathtt{p}_X}\,\partial_X\mathcal{H}\tanh((\mathtt{h}+c)|D|)P = \frac{1}{1+P^{-1}\mathtt{p}_X}\,P^{-1}\partial_X P\;P^{-1}\mathcal{H}\tanh((\mathtt{h}+c)|D|)P \\
&= \frac{1}{1+P^{-1}\mathtt{p}_X}\,(1+P^{-1}\mathtt{p}_X)\;\partial_x\;P^{-1}\mathcal{H}\tanh((\mathtt{h}+c)|D|)P = \partial_x P^{-1}\mathcal{H}\tanh((\mathtt{h}+c)|D|)P
\end{aligned}
$$

by the rule $P^{-1}\partial_X P = (1+P^{-1}\mathtt{p}_X)\,\partial_x$ for the changes of coordinates. □

Proof of Proposition 3.1 concluded. By Lemma 3.6 we write the Dirichlet-Neumann operator as

$$
G(\eta) = \partial_x P^{-1}\mathcal{H}\tanh((\mathtt{h}+c)|D|)P = |D|\tanh(\mathtt{h}|D|) + \mathcal{R}_G(\eta)\,, \quad \mathcal{R}_G(\eta) := \mathcal{R}_G^{(1)}(\eta) + \mathcal{R}_G^{(2)}(\eta)\,,
$$

where

$$
\mathcal{R}_G^{(1)}(\eta) := \partial_x\big(P^{-1}\mathcal{H}\tanh((\mathtt{h}+c)|D|)P - \mathcal{H}\tanh((\mathtt{h}+c)|D|)\big) \tag{3.49}
$$

$$
\mathcal{R}_G^{(2)}(\eta) := \partial_x\mathcal{H}\big(\tanh((\mathtt{h}+c)|D|) - \tanh(\mathtt{h}|D|)\big)\,. \tag{3.50}
$$

We analyze separately the two operators $\mathcal{R}_G^{(1)}(\eta)$, $\mathcal{R}_G^{(2)}(\eta)$ in (3.49), (3.50).

Analysis of $\mathcal{R}_G^{(1)}(\eta)$. Since $\tanh(x) = 1 - \frac{2}{1+e^{2x}}$, recalling the notation (2.33), for any $\mathtt{h} > 0$ one has

$$
\tanh(\mathtt{h}|D|) = \mathrm{Id} + \mathrm{Op}(r_{\mathtt{h}}), \quad r_{\mathtt{h}}(\xi) := -\frac{2}{1+e^{2\mathtt{h}|\xi|\chi(\xi)}} \in S^{-\infty} \tag{3.51}
$$

where the cut-off function χ is defined in (2.16). According to (3.51) and recalling (3.49) one gets

$$
\mathcal{R}_G^{(1)}(\eta) = \partial_x(P^{-1}\mathcal{H}P - \mathcal{H}) + \partial_x(P^{-1}\mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c})P - \mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c}))\,.
$$

Since, by (3.22) and (3.4), $\|\mathtt{p}\|_{2s_0+k_0+1}^{k_0,\gamma} \lesssim_{s_0,k_0} \|\eta\|_{2s_0+2k_0+1}^{k_0,\gamma} \le \delta(s_0,k_0)$, we can apply Lemma 2.18, obtaining that the operator $\partial_x(P^{-1}\mathcal{H}P - \mathcal{H})$ is an integral operator with kernel K_1 satisfying

$$
\|K_1\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|\mathtt{p}\|_{s+k_0+3}^{k_0,\gamma} \overset{(3.22)}{\lesssim_{s,k_0}} \|\eta\|_{s+2k_0+3}^{k_0,\gamma}\,. \tag{3.52}
$$

By (3.51), the operator $\Lambda_{\mathtt{h},c} := \mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c})$ is a Fourier multiplier in $OPS^{-\infty}$ and, reasoning similarly as in (3.24)-(3.27), we get

$$
|\!|\!|\Lambda_{\mathtt{h},c}|\!|\!|_{-m,s,\alpha}^{k_0,\gamma} \lesssim_{m,s,\alpha,k_0} 1 + \|c\|_s^{k_0,\gamma}\,, \quad \forall m \ge 0,\ s \ge s_0, \quad \forall \alpha \in \mathbb{N}\,.
$$

Hence, by Lemma 2.19, the operator $\Lambda_{\mathtt{h},c}$ is an integral operator with a $\mathcal{C}^\infty$ kernel $K_{\mathtt{h},c}$ satisfying

$$
\|K_{\mathtt{h},c}\|_s^{k_0,\gamma} \lesssim_{s,k_0} 1 + \|c\|_{s+s_0}^{k_0,\gamma} \overset{(3.22)}{\lesssim_{s,k_0}} 1 + \|\eta\|_{s+s_0+k_0}^{k_0,\gamma}\,, \quad \forall s \ge s_0\,. \tag{3.53}
$$

By formula (2.59) one has that the operator $\partial_x(P^{-1}\mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c})P - \mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c}))$ is an integral operator with a $\mathcal{C}^\infty$ kernel K_2 defined as

$$
\begin{aligned}
K_2(\lambda,\varphi,x,z) := \partial_x\Big(&K_{\mathtt{h},c}(\lambda,\varphi,x+\breve{\mathtt{p}}(\lambda,\varphi,x),z+\breve{\mathtt{p}}(\lambda,\varphi,z)) - K_{\mathtt{h},c}(\lambda,\varphi,x,z)\Big) \\
&+ \partial_x\Big(\partial_z\breve{\mathtt{p}}(\lambda,\varphi,z)K_{\mathtt{h},c}(\lambda,\varphi,x+\breve{\mathtt{p}}(\lambda,\varphi,x),z+\breve{\mathtt{p}}(\lambda,\varphi,z))\Big)
\end{aligned} \tag{3.54}
$$

where $z \mapsto \breve{\mathtt{p}}(\lambda,\varphi,z)$ is the inverse diffeomorphism of $x \mapsto \mathtt{p}(\lambda,\varphi,x)$. By Lemma 2.4 one gets

$$
\|\breve{\mathtt{p}}\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|\mathtt{p}\|_{s+k_0}^{k_0,\gamma} \overset{(3.22)}{\lesssim_{s,k_0}} \|\eta\|_{s+2k_0}^{k_0,\gamma}\,, \quad \forall s \ge s_0\,, \tag{3.55}
$$

and therefore, using also the mean value theorem to estimate the first term in (3.54), and (3.53), (3.55), (2.9), (2.10), the kernel K_2 satisfies

$$\|K_2\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|\eta\|_{s+s_0+2k_0+3}^{k_0,\gamma}, \quad \forall s \geq s_0 \,. \tag{3.56}$$

Hence $\mathcal{R}_G^{(1)}(\eta)$ is an integral operator with kernel $K_G^{(1)} := K_1 + K_2$ and by (3.52), (3.56) it satisfies

$$\|K_G^{(1)}\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|\eta\|_{s+s_0+2k_0+3}^{k_0,\gamma}, \quad \forall s \geq s_0 \,. \tag{3.57}$$

ANALYSIS OF $\mathcal{R}_G^{(2)}(\eta)$. By (3.50), (3.51) we write the Fourier multiplier

$$\mathcal{R}_G^{(2)}(\eta) = \partial_x \mathcal{H} \mathrm{Op}(r_{\mathtt{h}+c} - r_{\mathtt{h}}) = c\, \partial_x \mathcal{H} \mathrm{Op}(\breve{r}_{\mathtt{h},c}) \in OPS^{-\infty} \tag{3.58}$$

where

$$r_{\mathtt{h}+c}(\xi) - r_{\mathtt{h}}(\xi) = \breve{r}_{\mathtt{h},c}(\xi)\, c\,, \qquad \breve{r}_{\mathtt{h},c}(\xi) := 2|\xi|\chi(\xi) \int_0^1 \frac{2\exp\{2(\mathtt{h}+tc)|\xi|\chi(\xi)\}}{(1+\exp\{2(\mathtt{h}+tc)|\xi|\chi(\xi)\})^2}\, dt \in S^{-\infty} \,. \tag{3.59}$$

By a direct verification we have that

$$|\mathrm{Op}\big(r_{\mathtt{h}+c} - r_{\mathtt{h}}\big)|_{-m,s,\alpha}^{k_0,\gamma} \lesssim_{m,s,\alpha,k_0} \|c\|_s^{k_0,\gamma}, \quad \forall m \geq 0,\, s \geq s_0\,, \quad \forall \alpha \in \mathbb{N}\,. \tag{3.60}$$

Applying Lemma 2.19, we get that $\mathcal{R}_G^{(2)}(\eta)$ is an integral operator with $\mathcal{C}^\infty$ kernel $K_G^{(2)}$ and, using (3.22), (3.60),

$$\|K_G^{(2)}\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|\eta\|_{s+s_0+k_0}^{k_0,\gamma}, \qquad \forall s \geq s_0 \,. \tag{3.61}$$

Finally, defining $K_G := K_G^{(1)} + K_G^{(2)}$, the claimed estimate (3.5) follows by (2.57), (3.57), (3.61).

DIFFERENTIABILITY OF $\eta \mapsto K_G(\eta)$. Let $s_1 \geq 2s_0 + 1$. By applying Lemma 3.4 (with $s_1 + 4$ instead of s_1), the map

$$\{\|\eta\|_{s_1+6} < \delta(s_1)\} \mapsto H_\varphi^{s_1+4} \times H^{s_1+4}, \quad \eta \mapsto (c(\eta), \mathtt{p}(\eta)) \quad \text{is} \quad \mathcal{C}_b^1 \,. \tag{3.62}$$

Then, since $\breve{\mathtt{p}}(\varphi, x) = -\mathtt{p}(\varphi, x + \breve{\mathtt{p}}(\varphi, x))$, by the implicit function theorem, for $\mathtt{p}$ small in $\|\cdot\|_{s_1+4}$ norm, also the map $\mathtt{p} \mapsto \breve{\mathtt{p}}(\mathtt{p}) \in H^{s_1+2}$ is $\mathcal{C}_b^1$ implying that

$$\{\|\eta\|_{s_1+6} < \delta(s_1)\} \mapsto H^{s_1+2}, \quad \eta \mapsto \breve{\mathtt{p}}(\eta) \quad \text{is} \quad \mathcal{C}_b^1 \,. \tag{3.63}$$

By composition, using (2.61)-(2.62), (3.62), (3.63) the map $\big\{\|\eta\|_{s_1+6} \leq \delta(s_1)\big\} \to H^{s_1}$, $\eta \mapsto K_1(\eta)$ is $\mathcal{C}_b^1$, where K_1 is the kernel of the integral operator $\partial_x(P^{-1}\mathcal{H}P - \mathcal{H})$. Let us analyze the kernel K_2 in (3.54) of the operator $\partial_x(P^{-1}\mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c})P - \mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c}))$. Recalling (3.54) and using (3.62), (3.63), one gets that $\{\|\eta\|_{s_1+6} < \delta(s_1)\} \mapsto H^{s_1}$, $\eta \mapsto K_2(\eta)$ is $\mathcal{C}_b^1$. Therefore, recalling that $K_G^{(1)} = K_1 + K_2$ we get that

$$\{\|\eta\|_{s_1+6} < \delta(s_1)\} \mapsto H^{s_1}\,, \quad \eta \mapsto K_G^{(1)}(\eta) \quad \text{is} \quad \mathcal{C}_b^1.$$

The fact that the map $\{\|\eta\|_{s_1+6} < \delta(s_1)\} \mapsto H^{s_1}$, $\eta \mapsto K_G^{(2)}(\eta)$ is $\mathcal{C}_b^1$ follows by recalling (3.59), (3.58), (2.63) and (3.62). Then the proposition follows since $K_G = K_G^{(1)} + K_G^{(2)}$.

4 Degenerate KAM theory

In this section we verify that it is possible to develop degenerate KAM theory as in [11] and [21].

Definition 4.1. *A function $f := (f_1, \ldots, f_N) : [\mathtt{h}_1, \mathtt{h}_2] \to \mathbb{R}^N$ is called non-degenerate if, for any vector $c := (c_1, \ldots, c_N) \in \mathbb{R}^N \setminus \{0\}$, the function $f \cdot c = f_1 c_1 + \ldots + f_N c_N$ is not identically zero on the whole interval $[\mathtt{h}_1, \mathtt{h}_2]$.*

From a geometric point of view, f non-degenerate means that the image of the curve $f([\mathtt{h}_1,\mathtt{h}_2]) \subset \mathbb{R}^N$ is not contained in any hyperplane of $\mathbb{R}^N$. For such a reason a curve f which satisfies the non-degeneracy property of Definition 4.1 is also referred to as an *essentially non-planar* curve, or a curve with *full torsion.* Given $\mathbb{S}^+ \subset \mathbb{N}^+$ we denote the unperturbed tangential and normal frequency vectors by

$$\vec{\omega}(\mathtt{h}) := (\omega_j(\mathtt{h}))_{j\in\mathbb{S}^+} , \quad \vec{\Omega}(\mathtt{h}) := (\Omega_j(\mathtt{h}))_{j\in\mathbb{N}^+\setminus\mathbb{S}^+} := (\omega_j(\mathtt{h}))_{j\in\mathbb{N}^+\setminus\mathbb{S}^+} , \tag{4.1}$$

where $\omega_j(\mathtt{h}) = \sqrt{j\tanh(\mathtt{h}j)}$ are defined in (1.19).

Lemma 4.2. (Non-degeneracy) *The frequency vectors* $\vec{\omega}(\mathtt{h}) \in \mathbb{R}^\nu$, $(\vec{\omega}(\mathtt{h}),1) \in \mathbb{R}^{\nu+1}$ *and*

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

are non-degenerate.

Proof. We first prove that for any N, for any $\omega_{j_1}(\mathtt{h}),\dots,\omega_{j_N}(\mathtt{h})$ with $1 \le j_1 < j_2 < \dots < j_N$ the function $[\mathtt{h}_1,\mathtt{h}_2] \ni \mathtt{h} \mapsto (\omega_{j_1}(\mathtt{h}),\dots,\omega_{j_N}(\mathtt{h})) \in \mathbb{R}^N$ is non-degenerate according to Definition 4.1, namely that, for all $c \in \mathbb{R}^N \setminus \{0\}$, the function $\mathtt{h} \mapsto c_1\omega_{j_1}(\mathtt{h}) + \dots + c_N\omega_{j_N}(\mathtt{h})$ is not identically zero on the interval $[\mathtt{h}_1,\mathtt{h}_2]$. We shall prove, equivalently, that the function

$$\mathtt{h} \mapsto c_1\omega_{j_1}(\mathtt{h}^4) + \dots + c_N\omega_{j_N}(\mathtt{h}^4)$$

is not identically zero on the interval $[\mathtt{h}_1^4,\mathtt{h}_2^4]$. The advantage of replacing $\mathtt{h}$ with $\mathtt{h}^4$ is that each function

$$\mathtt{h} \mapsto \omega_j(\mathtt{h}^4) = \sqrt{j\tanh(\mathtt{h}^4 j)}$$

is *analytic also in a neighborhood of* $\mathtt{h} = 0$, unlike the function $\omega_j(\mathtt{h}) = \sqrt{j\tanh(\mathtt{h}j)}$. Clearly, the function $g_1(\mathtt{h}) := \sqrt{\tanh(\mathtt{h}^4)}$ is analytic in a neighborhood of any $\mathtt{h} \in \mathbb{R}\setminus\{0\}$, because g_1 is the composition of analytic functions. Let us prove that it has an analytic continuation at $\mathtt{h} = 0$. The Taylor series at $z = 0$ of the hyperbolic tangent has the form

$$\tanh(z) = \sum_{n=0}^{\infty} T_n z^{2n+1} = z - \frac{z^3}{3} + \frac{2}{15}z^5 + \dots ,$$

and it is convergent for $|z| < \pi/2$ (the poles of $\tanh z$ closest to $z = 0$ are $\pm \mathrm{i}\pi/2$). Then the power series

$$\tanh(z^4) = \sum_{n=0}^{\infty} T_n z^{4(2n+1)} = z^4\Big(1 + \sum_{n\ge1} T_n z^{8n}\Big) = z^4 - \frac{z^{12}}{3} + \frac{2}{15}z^{20} + \dots$$

is convergent in $|z| < (\pi/2)^{1/4}$. Moreover $|\sum_{n\ge1} T_n z^{8n}| < 1$ in a ball $|z| < \delta$, for some positive δ sufficiently small. As a consequence, also the real function

$$g_1(\mathtt{h}) := \omega_1(\mathtt{h}^4) = \sqrt{\tanh(\mathtt{h}^4)} = \mathtt{h}^2\Big(1 + \sum_{n\ge1} T_n \mathtt{h}^{8n}\Big)^{1/2} = \sum_{n=0}^{+\infty} b_n \frac{\mathtt{h}^{8n+2}}{(8n+2)!} = \mathtt{h}^2 - \frac{\mathtt{h}^{10}}{6} + \dots \tag{4.2}$$

is analytic in the ball $|z| < \delta$. Thus g_1 is analytic on the whole real axis. The Taylor coefficients b_n are computable. We expand in Taylor series at $\mathtt{h} = 0$ also each function, for $j \ge 1$,

$$g_j(\mathtt{h}) := \omega_j(\mathtt{h}^4) = \sqrt{j}\sqrt{\tanh(\mathtt{h}^4 j)} = \sqrt{j}\, g_1(j^{1/4}\mathtt{h}) = \sum_{n=0}^{+\infty} b_n j^{2n+1} \frac{\mathtt{h}^{8n+2}}{(8n+2)!} , \tag{4.3}$$

which is analytic on the whole $\mathbb{R}$, similarly as g_1.

Now fix N integers $1 \le j_1 < j_2 < \dots < j_N$. We prove that for all $c \in \mathbb{R}^N \setminus \{0\}$, the analytic function $c_1 g_{j_1}(\mathtt{h}) + \dots + c_N g_{j_N}(\mathtt{h})$ is not identically zero. Suppose, by contradiction, that there exists $c \in \mathbb{R}^N \setminus \{0\}$ such that

$$c_1 g_{j_1}(\mathtt{h}) + \dots + c_N g_{j_N}(\mathtt{h}) = 0 \quad \forall \mathtt{h} \in \mathbb{R}. \tag{4.4}$$

The real analytic function $g_1(\mathtt{h})$ defined in (4.2) is not a polynomial (to see this, observe its limit as $\mathtt{h} \to \infty$). Hence there exist N Taylor coefficients $b_n \neq 0$ of g_1, say $b_{n_1}, \dots, b_{n_N}$ with $n_1 < n_2 < \dots < n_N$. We differentiate with respect to $\mathtt{h}$ the identity in (4.4) and we find

$$\begin{cases} c_1\big(D_{\mathtt{h}}^{(8n_1+2)} g_{j_1}\big)(\mathtt{h}) + \dots + c_N\big(D_{\mathtt{h}}^{(8n_1+2)} g_{j_N}\big)(\mathtt{h}) = 0 \\ c_1\big(D_{\mathtt{h}}^{(8n_2+2)} g_{j_1}\big)(\mathtt{h}) + \dots + c_N\big(D_{\mathtt{h}}^{(8n_2+2)} g_{j_N}\big)(\mathtt{h}) = 0 \\ \dots\dots\dots \\ c_1\big(D_{\mathtt{h}}^{(8n_N+2)} g_{j_1}\big)(\mathtt{h}) + \dots + c_N\big(D_{\mathtt{h}}^{(8n_N+2)} g_{j_N}\big)(\mathtt{h}) = 0\,. \end{cases}$$

As a consequence the $N \times N$-matrix

$$\mathcal{A}(\mathtt{h}) := \begin{pmatrix} \big(D_{\mathtt{h}}^{(8n_1+2)} g_{j_1}\big)(\mathtt{h}) & \dots & \big(D_{h}^{(8n_1+2)} g_{j_N}\big)(\mathtt{h}) \\ \big(D_{\mathtt{h}}^{(8n_2+2)} g_{j_1}\big)(\mathtt{h}) & \dots & \big(D_{h}^{(8n_2+2)} g_{j_N}\big)(\mathtt{h}) \\ \vdots & \ddots & \vdots \\ \big(D_{\mathtt{h}}^{(8n_N+2)} g_{j_1}\big)(\mathtt{h}) & \dots & \big(D_{\mathtt{h}}^{(8n_N+2)} g_{j_N}\big)(\mathtt{h}) \end{pmatrix} \tag{4.5}$$

is singular for all $\mathtt{h} \in \mathbb{R}$, and so the analytic function

$$\det \mathcal{A}(\mathtt{h}) = 0 \quad \forall \mathtt{h} \in \mathbb{R} \tag{4.6}$$

is identically zero. In particular at $\mathtt{h} = 0$ we have $\det \mathcal{A}(0) = 0$. On the other hand, by (4.3) and the multi-linearity of the determinant we compute

$$\det \mathcal{A}(0) := \det \begin{pmatrix} b_{n_1} j_1^{2n_1+1} & \dots & b_{n_1} j_N^{2n_1+1} \\ b_{n_2} j_1^{2n_2+1} & \dots & b_{n_2} j_N^{2n_2+1} \\ \vdots & \ddots & \vdots \\ b_{n_N} j_1^{2n_N+1} & \dots & b_{n_N} j_N^{2n_N+1} \end{pmatrix} = b_{n_1} \dots b_{n_N} \det \begin{pmatrix} j_1^{2n_1+1} & \dots & j_N^{2n_1+1} \\ j_1^{2n_2+1} & \dots & j_N^{2n_2+1} \\ \vdots & \ddots & \vdots \\ j_1^{2n_N+1} & \dots & j_N^{2n_N+1} \end{pmatrix}.$$

This is a generalized Van der Monde determinant. We use the following result.

Lemma 4.3. *Let $x_1, \dots, x_N, \alpha_1, \dots, \alpha_N$ be real numbers, with $0 < x_1 < \dots < x_N$ and $\alpha_1 < \dots < \alpha_N$. Then*

$$\det \begin{pmatrix} x_1^{\alpha_1} & \dots & x_N^{\alpha_1} \\ \vdots & \ddots & \vdots \\ x_1^{\alpha_N} & \dots & x_N^{\alpha_N} \end{pmatrix} > 0\,.$$

Proof. The lemma is proved in [56]. □

Since $1 \leq j_1 < j_2 < \dots < j_N$ and the exponents $\alpha_j := 2n_j + 1$ are increasing $\alpha_1 < \dots < \alpha_N$, Lemma 4.3 implies that $\det \mathcal{A}(0) \neq 0$ (recall that $b_{n_1}, \dots, b_{n_N} \neq 0$). This is a contradiction with (4.6).

In order to conclude the proof of Lemma 4.2 we have to prove that, for any N, for any $1 \leq j_1 < j_2 < \dots < j_N$, the function $[\mathtt{h}_1, \mathtt{h}_2] \ni \mathtt{h} \mapsto (1, \omega_{j_1}(\mathtt{h}), \dots, \omega_{j_N}(\mathtt{h})) \in \mathbb{R}^{N+1}$ is non-degenerate according to Definition 4.1, namely that, for all $c = (c_0, c_1, \dots, c_N) \in \mathbb{R}^{N+1} \setminus \{0\}$, the function $\mathtt{h} \mapsto c_0 + c_1\omega_{j_1}(\mathtt{h}) + \dots + c_N\omega_{j_N}(\mathtt{h})$ is not identically zero on the interval $[\mathtt{h}_1, \mathtt{h}_2]$. We shall prove, equivalently, that the real analytic function $\mathtt{h} \mapsto c_0 + c_1\omega_{j_1}(\mathtt{h}^4) + \dots + c_N\omega_{j_N}(\mathtt{h}^4)$ is not identically zero on $\mathbb{R}$.

Suppose, by contradiction, that there exists $c = (c_0, c_1, \dots, c_N) \in \mathbb{R}^{N+1} \setminus \{0\}$ such that

$$c_0 + c_1 g_{j_1}(\mathtt{h}) + \dots + c_N g_{j_N}(\mathtt{h}) = 0 \quad \forall \mathtt{h} \in \mathbb{R}. \tag{4.7}$$

As above, we differentiate with respect to $\mathtt{h}$ the identity (4.7), and we find that the $(N+1) \times (N+1)$-matrix

$$\mathcal{B}(\mathtt{h}) := \begin{pmatrix} 1 & g_{j_1}(\mathtt{h}) & \dots & g_{j_N}(\mathtt{h}) \\ 0 & (D_{\mathtt{h}}^{(8n_1+2)} g_{j_1})(\mathtt{h}) & \dots & (D_{\mathtt{h}}^{(8n_1+2)} g_{j_N})(\mathtt{h}) \\ 0 & \vdots & \ddots & \vdots \\ 0 & (D_{\mathtt{h}}^{(8n_N+2)} g_{j_1})(\mathtt{h}) & \dots & (D_{\mathtt{h}}^{(8n_N+2)} g_{j_N})(\mathtt{h}) \end{pmatrix} \tag{4.8}$$

is singular for all $\mathtt{h} \in \mathbb{R}$, and so the analytic function $\det \mathcal{B}(\mathtt{h}) = 0$ for all $\mathtt{h} \in \mathbb{R}$. By expanding the determinant of the matrix in (4.8) along the first column by Laplace we get $\det \mathcal{B}(\mathtt{h}) = \det \mathcal{A}(\mathtt{h})$, where the matrix $\mathcal{A}(\mathtt{h})$ is defined in (4.5). We have already proved that $\det \mathcal{A}(0) \neq 0$, and this gives a contradiction. $\square$

In the next proposition we deduce the quantitative bounds (4.9)-(4.12) from the qualitative non-degeneracy condition of Lemma 4.2, the analyticity of the linear frequencies ω_j in (1.19), and their asymptotics (1.24).

Proposition 4.4. (Transversality) *There exist $k_0^* \in \mathbb{N}$, $\rho_0 > 0$ such that, for any $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$,*

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell\}| \geq \rho_0 \langle \ell \rangle , \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \tag{4.9}$$

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h})\}| \geq \rho_0 \langle \ell \rangle , \quad \forall \ell \in \mathbb{Z}^\nu, \ j \in \mathbb{N}^+ \setminus \mathbb{S}^+, \tag{4.10}$$

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})\}| \geq \rho_0 \langle \ell \rangle , \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+, \tag{4.11}$$

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) + \Omega_{j'}(\mathtt{h})\}| \geq \rho_0 \langle \ell \rangle , \quad \forall \ell \in \mathbb{Z}^\nu, \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \tag{4.12}$$

where $\vec{\omega}(\mathtt{h})$ and $\Omega_j(\mathtt{h})$ are defined in (4.1). *We recall the notation* $\langle \ell \rangle := \max\{1, |\ell|\}$. *We call (following [57]) ρ_0 the "*amount of non-degeneracy*" and k_0^* the "*index of non-degeneracy*".*

Note that in (4.11) we exclude the index $\ell = 0$. In this case we directly have that, for all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$

$$|\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})| \geq c_1 |\sqrt{j} - \sqrt{j'}| = c_1 \frac{|j - j'|}{\sqrt{j} + \sqrt{j'}} \quad \forall j, j' \in \mathbb{N}^+, \qquad \text{where } c_1 := \sqrt{\tanh(\mathtt{h}_1)} \, . \tag{4.13}$$

Proof. All the inequalities (4.9)-(4.12) are proved by contradiction.

Proof of (4.9). Suppose that for all $k_0^* \in \mathbb{N}$, for all $\rho_0 > 0$ there exist $\ell \in \mathbb{Z}^\nu \setminus \{0\}$, $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ such that $\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell\}| < \rho_0 \langle \ell \rangle$. This implies that for all $m \in \mathbb{N}$, taking $k_0^* = m$, $\rho_0 = \frac{1}{1+m}$, there exist $\ell_m \in \mathbb{Z}^\nu \setminus \{0\}$, $\mathtt{h}_m \in [\mathtt{h}_1, \mathtt{h}_2]$ such that

$$\max_{k \leq m} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}_m) \cdot \ell_m\}| < \frac{1}{1+m} \langle \ell_m \rangle$$

and therefore

$$\forall k \in \mathbb{N}, \quad \forall m \geq k \, , \quad \Big| \partial_{\mathtt{h}}^k \vec{\omega}(\mathtt{h}_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle} \Big| < \frac{1}{1+m} \, . \tag{4.14}$$

The sequences $(\mathtt{h}_m)_{m \in \mathbb{N}} \subset [\mathtt{h}_1, \mathtt{h}_2]$ 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 $m_n \to +\infty$ such that $\mathtt{h}_{m_n} \to \bar{\mathtt{h}} \in [\mathtt{h}_1, \mathtt{h}_2]$, $\ell_{m_n} / \langle \ell_{m_n} \rangle \to \bar{c} \neq 0$. Passing to the limit in (4.14) for $m_n \to +\infty$ we deduce that $\partial_{\mathtt{h}}^k \vec{\omega}(\bar{\mathtt{h}}) \cdot \bar{c} = 0$ for all $k \in \mathbb{N}$. We conclude that the analytic function $\mathtt{h} \mapsto \vec{\omega}(\mathtt{h}) \cdot \bar{c}$ is identically zero. Since $\bar{c} \neq 0$, this is in contradiction with Lemma 4.2.

Proof of (4.10). First of all note that for all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, we have $|\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h})| \geq \Omega_j(\mathtt{h}) - |\vec{\omega}(\mathtt{h}) \cdot \ell| \geq c_1 j^{1/2} - C|\ell| \geq |\ell|$ if $j^{1/2} \geq C_0 |\ell|$ for some $C_0 > 0$. Therefore in (4.10) we can restrict to the indices $(\ell, j) \in \mathbb{Z}^\nu \times (\mathbb{N}^+ \setminus \mathbb{S}^+)$ satisfying

$$j^{\frac{1}{2}} < C_0 |\ell| \, . \tag{4.15}$$

Arguing by contradiction (as for proving (4.9)), 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 $\mathtt{h}_m \in [\mathtt{h}_1, \mathtt{h}_2]$, such that

$$\max_{k \leq m} \Big| \partial_{\mathtt{h}}^k \Big\{ \vec{\omega}(\mathtt{h}_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle} + \frac{\Omega_{j_m}(\mathtt{h}_m)}{\langle \ell_m \rangle} \Big\} \Big| < \frac{1}{1+m}$$

and therefore

$$\forall k \in \mathbb{N}, \quad \forall m \geq k \, , \quad \Big| \partial_{\mathtt{h}}^k \Big\{ \vec{\omega}(\mathtt{h}_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle} + \frac{\Omega_{j_m}(\mathtt{h}_m)}{\langle \ell_m \rangle} \Big\} \Big| < \frac{1}{1+m} \, . \tag{4.16}$$

Since the sequences $(\mathtt{h}_m)_{m \in \mathbb{N}} \subset [\mathtt{h}_1, \mathtt{h}_2]$ and $(\ell_m / \langle \ell_m \rangle)_{m \in \mathbb{N}} \in \mathbb{R}^\nu$ are bounded, there exists a sequence $m_n \to +\infty$ such that

$$\mathtt{h}_{m_n} \to \bar{\mathtt{h}} \in [\mathtt{h}_1, \mathtt{h}_2] \, , \quad \frac{\ell_{m_n}}{\langle \ell_{m_n} \rangle} \to \bar{c} \in \mathbb{R}^\nu \, . \tag{4.17}$$

We now distinguish two cases.

Case 1: $(\ell_{m_n}) \subset \mathbb{Z}^\nu$ *is bounded.* In this case, up to a subsequence, $\ell_{m_n} \to \bar\ell \in \mathbb{Z}^\nu$, and since $|j_m| \le C|\ell_m|^2$ for all m (see (4.15)), we have $j_{m_n} \to \bar\jmath$. Passing to the limit for $m_n \to +\infty$ in (4.16) we deduce, by (4.17), that

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

Therefore the analytic function $\mathtt{h} \mapsto \vec\omega(\mathtt{h}) \cdot \bar c + \langle \bar\ell \rangle^{-1} \Omega_{\bar\jmath}(\mathtt{h})$ is identically zero. Since $(\bar c, \langle \bar\ell \rangle^{-1}) \neq 0$ this is in contradiction with Lemma 4.2.

Case 2: (ℓ_{m_n}) *is unbounded.* Up to a subsequence, $|\ell_{m_n}| \to +\infty$. In this case the constant $\bar c$ in (4.17) is nonzero. Moreover, by (4.15), we also have that, up to a subsequence,

$$j_{m_n}^{\frac12} \langle \ell_{m_n} \rangle^{-1} \to \bar d \in \mathbb{R}. \tag{4.18}$$

By (1.24), (4.17), (4.18), we get

$$\frac{\Omega_{j_{m_n}}(\mathtt{h}_{m_n})}{\langle \ell_{m_n} \rangle} = \frac{j_{m_n}^{\frac12}}{\langle \ell_{m_n} \rangle} + \frac{r(j_{m_n}, \mathtt{h}_{m_n})}{\langle \ell_{m_n} \rangle} \to \bar d \,, \quad \partial_{\mathtt{h}}^k \frac{\Omega_{j_{m_n}}(\mathtt{h}_{m_n})}{\langle \ell_{m_n} \rangle} = \partial_{\mathtt{h}}^k \frac{r(j_{m_n}, \mathtt{h}_{m_n})}{\langle \ell_{m_n} \rangle} \to 0 \quad \forall k \ge 1 \tag{4.19}$$

as $m_n \to +\infty$. Passing to the limit in (4.16), by (4.19), (4.17) we deduce that $\partial_{\mathtt{h}}^k \big\{ \vec\omega(\bar{\mathtt{h}}) \cdot \bar c + \bar d \big\} = 0$, for all $k \in \mathbb{N}$. Therefore the analytic function $\mathtt{h} \mapsto \vec\omega(\mathtt{h}) \cdot \bar c + \bar d = 0$ is identically zero. Since $(\bar c, \bar d) \neq 0$ this is in contradiction with Lemma 4.2.

Proof of (4.11). For all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, by (4.13) and (1.19), we have

$$|\vec\omega(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})| \ge |\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})| - |\vec\omega(\mathtt{h})||\ell| \ge c_1 |j^{\frac12} - j'^{\frac12}| - C|\ell| \ge \langle \ell \rangle$$

provided $|j^{\frac12} - j'^{\frac12}| \ge C_1 \langle \ell \rangle$, for some $C_1 > 0$. Therefore in (4.11) we can restrict to the indices such that

$$|j^{\frac12} - j'^{\frac12}| < C_1 \langle \ell \rangle \,. \tag{4.20}$$

Moreover in (4.11) we can also assume that $j \neq j'$, otherwise (4.11) reduces to (4.9), which is already proved. If, by contradiction, (4.11) is false, we deduce, arguing as in the previous cases, that, for all $m \in \mathbb{N}$, there exist $\ell_m \in \mathbb{Z}^\nu \setminus \{0\}$, $j_m, j'_m \in \mathbb{N}^+ \setminus \mathbb{S}^+$, $j_m \neq j'_m$, $\mathtt{h}_m \in [\mathtt{h}_1, \mathtt{h}_2]$, such that

$$\forall k \in \mathbb{N} \,, \quad \forall m \ge k \,, \quad \Big| \partial_{\mathtt{h}}^k \Big\{ \vec\omega(\mathtt{h}_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle} + \frac{\Omega_{j_m}(\mathtt{h}_m)}{\langle \ell_m \rangle} - \frac{\Omega_{j'_m}(\mathtt{h}_m)}{\langle \ell_m \rangle} \Big\} \Big| < \frac{1}{1+m} \,. \tag{4.21}$$

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

$$\mathtt{h}_{m_n} \to \bar{\mathtt{h}} \in [\mathtt{h}_1, \mathtt{h}_2] \,, \quad \ell_{m_n} / \langle \ell_{m_n} \rangle \to \bar c \in \mathbb{R}^\nu \setminus \{0\} \,. \tag{4.22}$$

We distinguish again two cases.

Case 1 : (ℓ_{m_n}) *is unbounded.* Using (4.20) we deduce that, up to a subsequence,

$$|j_m^{\frac12} - j_m'^{\frac12}| \langle \ell_m \rangle^{-1} \to \bar d \in \mathbb{R} \,. \tag{4.23}$$

Hence passing to the limit in (4.21) for $m_n \to +\infty$, we deduce by (4.22), (4.23), (1.24) that

$$\partial_{\mathtt{h}}^k \{ \vec\omega(\bar{\mathtt{h}}) \cdot \bar c + \bar d \} = 0 \quad \forall k \in \mathbb{N}.$$

Therefore the analytic function $\mathtt{h} \mapsto \vec\omega(\mathtt{h}) \cdot \bar c + \bar d$ is identically zero. This in contradiction with Lemma 4.2.

Case 2 : (ℓ_{m_n}) *is bounded.* By (4.20), we have that $|\sqrt{j_m} - \sqrt{j'_m}| \le C$ and so, up to a subsequence, only the following two subcases are possible:

(*i*) $j_m, j'_m \leq C$. Up to a subsequence, $j_{m_n} \to \bar{\jmath}$, $j'_{m_n} \to \bar{\jmath}'$, $\ell_{m_n} \to \bar{\ell} \neq 0$ and $\mathtt{h}_{m_n} \to \bar{\mathtt{h}}$. Hence passing to the limit in (4.21) we deduce that

$$\partial_{\mathtt{h}}^k \Big\{ \vec{\omega}(\bar{\mathtt{h}}) \cdot \bar{c} + \frac{\Omega_{\bar{\jmath}}(\bar{\mathtt{h}}) - \Omega_{\bar{\jmath}'}(\bar{\mathtt{h}})}{\langle \bar{\ell} \rangle} \Big\} = 0 \quad \forall k \in \mathbb{N}.$$

Hence the analytic function $\mathtt{h} \mapsto \vec{\omega}(\bar{\mathtt{h}}) \cdot \bar{c} + (\Omega_{\bar{\jmath}}(\bar{\mathtt{h}}) - \Omega_{\bar{\jmath}'}(\bar{\mathtt{h}})) \langle \bar{\ell} \rangle^{-1}$ is identically zero, which is a contradiction with Lemma 4.2.

(*ii*) $j_m, j'_m \to +\infty$. By (4.23) and (1.24), we deduce, passing to the limit in (4.21), that

$$\partial_{\mathtt{h}}^k \{ \vec{\omega}(\mathtt{h}) \cdot \bar{c} + \bar{d} \} = 0 \quad \forall k \in \mathbb{N}.$$

Hence the analytic function $\mathtt{h} \mapsto \vec{\omega}(\mathtt{h}) \cdot \bar{c} + \bar{d}$ is identically zero, which contradicts Lemma 4.2.

Proof of (4.12). The proof is similar to (4.10). First of all note that for all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, we have

$$|\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) + \Omega_{j'}(\mathtt{h})| \geq \Omega_j(\mathtt{h}) + \Omega_{j'}(\mathtt{h}) - |\vec{\omega}(\mathtt{h}) \cdot \ell| \geq c_1 \sqrt{j} + c_1 \sqrt{j'} - C|\ell| \geq |\ell|$$

if $\sqrt{j} + \sqrt{j'} \geq C_0 |\ell|$ for some $C_0 > 0$. Therefore in (4.10) we can restrict the analysis to the indices $(\ell, j, j') \in \mathbb{Z}^\nu \times (\mathbb{N}^+ \setminus \mathbb{S}^+)^2$ satisfying

$$\sqrt{j} + \sqrt{j'} < C_0 |\ell|. \tag{4.24}$$

Arguing by contradiction as above, 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 $\mathtt{h}_m \in [\mathtt{h}_1, \mathtt{h}_2]$ such that

$$\forall k \in \mathbb{N}, \quad \forall m \geq k, \quad \Big| \partial_{\mathtt{h}}^k \Big\{ \vec{\omega}(\mathtt{h}_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle} + \frac{\Omega_{j_m}(\mathtt{h}_m)}{\langle \ell_m \rangle} + \frac{\Omega_{j'_m}(\mathtt{h}_m)}{\langle \ell_m \rangle} \Big\} \Big| < \frac{1}{1+m}. \tag{4.25}$$

Since the sequences $(\mathtt{h}_m)_{m \in \mathbb{N}} \subset [\mathtt{h}_1, \mathtt{h}_2]$ and $(\ell_m / \langle \ell_m \rangle)_{m \in \mathbb{N}} \in \mathbb{R}^\nu$ are bounded, there exist $m_n \to +\infty$ such that

$$\mathtt{h}_{m_n} \to \bar{\mathtt{h}} \in [\mathtt{h}_1, \mathtt{h}_2], \quad \frac{\ell_{m_n}}{\langle \ell_{m_n} \rangle} \to \bar{c} \in \mathbb{R}^\nu. \tag{4.26}$$

We now distinguish two cases.

Case 1: $(\ell_{m_n}) \subset \mathbb{Z}^\nu$ is bounded. Up to a subsequence, $\ell_{m_n} \to \bar{\ell} \in \mathbb{Z}^\nu$, and since, by (4.24), also $j_m, j'_m \leq C$ for all m, we have $j_{m_n} \to \bar{\jmath}$, $j'_{m_n} \to \bar{\jmath}'$. Passing to the limit for $m_n \to +\infty$ in (4.25) we deduce, by (4.26), that

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

Therefore the analytic function $\mathtt{h} \mapsto \vec{\omega}(\mathtt{h}) \cdot \bar{c} + \langle \bar{\ell} \rangle^{-1} \Omega_{\bar{\jmath}}(\mathtt{h}) + \langle \bar{\ell} \rangle^{-1} \Omega_{\bar{\jmath}'}(\mathtt{h})$ is identically zero. This is in contradiction with Lemma 4.2.

Case 2: (ℓ_{m_n}) is unbounded. Up to a subsequence, $|\ell_{m_n}| \to +\infty$. In this case the constant $\bar{c}$ in (4.26) is nonzero. Moreover, by (4.24), we also have that, up to a subsequence,

$$(j_{m_n}^{\frac{1}{2}} + j'^{\frac{1}{2}}_{m_n}) \langle \ell_{m_n} \rangle^{-1} \to \bar{d} \in \mathbb{R}. \tag{4.27}$$

By (1.24), (4.26), (4.27), passing to the limit as $m_n \to +\infty$ in (4.25) we deduce that $\partial_{\mathtt{h}}^k \{ \vec{\omega}(\bar{\mathtt{h}}) \cdot \bar{c} + \bar{d} \} = 0$ for all $k \in \mathbb{N}$. Therefore the analytic function $\mathtt{h} \mapsto \vec{\omega}(\mathtt{h}) \cdot \bar{c} + \bar{d} = 0$ is identically zero. Since $(\bar{c}, \bar{d}) \neq 0$, this is in contradiction with Lemma 4.2. $\square$

5 Nash-Moser theorem and measure estimates

We rescale the variable $u = \varepsilon \tilde{u}$ with $\tilde{u} = O(1)$, writing (1.14) (after dropping the tilde) as

$$\partial_t u = J \Omega u + \varepsilon X_{P_\varepsilon}(u) \tag{5.1}$$

where $J\Omega$ is the linearized Hamiltonian vector field in (1.16) and

$$X_{P_\varepsilon}(u,\mathtt{h}) := X_{P_\varepsilon}(u) := \begin{pmatrix} \varepsilon^{-1}(G(\varepsilon\eta,\mathtt{h}) - G(0,\mathtt{h}))\psi \\ -\frac{1}{2}\psi_x^2 + \frac{1}{2}\frac{\left(G(\varepsilon\eta,\mathtt{h})\psi + \varepsilon\eta_x\psi_x\right)^2}{1+(\varepsilon\eta_x)^2} \end{pmatrix} . \tag{5.2}$$

System (5.1) is the Hamiltonian system generated by the Hamiltonian

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

where H is the water waves Hamiltonian (1.7) (with $g = 1$ and depth $\mathtt{h}$), H_L is defined in (1.17) and

$$P_\varepsilon(u,\mathtt{h}) := P_\varepsilon(u) := \frac{\varepsilon^{-1}}{2}\big(\psi, \big(G(\varepsilon\eta,\mathtt{h}) - G(0,\mathtt{h})\big)\psi\big)_{L^2(\mathbb{T}_x)} . \tag{5.3}$$

We decompose the phase space

$$H^1_{0,\mathrm{even}} := \Big\{ u := (\eta,\psi) \in H^1_0(\mathbb{T}_x) \times \dot{H}^1(\mathbb{T}_x) \,, \quad u(x) = u(-x) \Big\} \tag{5.4}$$

as the direct sum of the symplectic subspaces

$$H^1_{0,\mathrm{even}} = H_{\mathbb{S}^+} \oplus H^\perp_{\mathbb{S}^+} \tag{5.5}$$

as

$$H_{\mathbb{S}^+} := \Big\{ v := \sum_{j\in\mathbb{S}^+} \begin{pmatrix} \eta_j \\ \psi_j \end{pmatrix} \cos(jx) \Big\} , \quad H^\perp_{\mathbb{S}^+} := \Big\{ z := \begin{pmatrix} \eta \\ \psi \end{pmatrix} = \sum_{j\in\mathbb{N}^+\setminus\mathbb{S}^+} \begin{pmatrix} \eta_j \\ \psi_j \end{pmatrix} \cos(jx) \Big\} .$$

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

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

where $\xi_j > 0$, $j = 1, \ldots, \nu$, the variables I_j satisfy $|I_j| < \xi_j$, and we leave unchanged the normal component z. The symplectic 2-form in (1.8) reads

$$\mathcal{W} := \Big(\sum\nolimits_{j\in\mathbb{S}^+} d\theta_j \wedge dI_j\Big) \oplus \mathcal{W}_{|H^\perp_{\mathbb{S}^+}} = d\Lambda, \tag{5.6}$$

where Λ is the Liouville 1-form

$$\Lambda_{(\theta,I,z)}[\widehat{\theta}, \widehat{I}, \widehat{z}] := -\sum_{j\in\mathbb{S}^+} I_j \widehat{\theta}_j - \frac{1}{2}\big(Jz \,,\, \widehat{z}\big)_{L^2} . \tag{5.7}$$

Hence the Hamiltonian system (5.1) transforms into the new Hamiltonian system

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

generated by the Hamiltonian

$$H_\varepsilon := \mathcal{H}_\varepsilon \circ A = \varepsilon^{-2} H \circ \varepsilon A \tag{5.8}$$

where

$$A(\theta,I,z) := v(\theta,I) + z := \sum_{j\in\mathbb{S}^+} \sqrt{\frac{2}{\pi}} \begin{pmatrix} \omega_j^{1/2}\sqrt{\xi_j + I_j}\, \cos(\theta_j) \\ -\omega_j^{-1/2}\sqrt{\xi_j + I_j}\, \sin(\theta_j) \end{pmatrix} \cos(jx) + z \,. \tag{5.9}$$

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

$$\tilde\rho : (\theta, I, z) \mapsto (-\theta, I, \rho z) \,. \tag{5.10}$$

By (1.7) and (5.8) the Hamiltonian H_ε reads (up to a constant)

$$H_\varepsilon = \mathcal{N} + \varepsilon P\,, \quad \mathcal{N} := H_L \circ A = \vec{\omega}(\mathtt{h}) \cdot I + \frac{1}{2}(z, \Omega z)_{L^2}\,, \quad P := P_\varepsilon \circ A\,, \tag{5.11}$$

where $\vec{\omega}(\mathtt{h})$ is defined in (4.1) and Ω in (1.16). We look for an embedded invariant torus

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

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 (5.23)).

5.1 Nash-Moser theorem of hypothetical conjugation

For $\alpha \in \mathbb{R}^\nu$, we consider the modified Hamiltonian

$$H_\alpha := \mathcal{N}_\alpha + \varepsilon P\,, \quad \mathcal{N}_\alpha := \alpha \cdot I + \frac{1}{2}(z, \Omega z)_{L^2}\,. \tag{5.12}$$

We look for zeros of the nonlinear operator

$$\begin{aligned} \mathcal{F}(i,\alpha) := \mathcal{F}(i,\alpha,\omega,\mathtt{h},\varepsilon) &:= \omega \cdot \partial_\varphi i(\varphi) - X_{H_\alpha}(i(\varphi)) = \omega \cdot \partial_\varphi i(\varphi) - (X_{\mathcal{N}_\alpha} + \varepsilon X_P)(i(\varphi)) \\ &:= \begin{pmatrix} \omega \cdot \partial_\varphi \theta(\varphi) - \alpha - \varepsilon \partial_I P(i(\varphi)) \\ \omega \cdot \partial_\varphi I(\varphi) + \varepsilon \partial_\theta P(i(\varphi)) \\ \omega \cdot \partial_\varphi z(\varphi) - J(\Omega z(\varphi) + \varepsilon \nabla_z P(i(\varphi))) \end{pmatrix} \end{aligned} \tag{5.13}$$

where $\Theta(\varphi) := \theta(\varphi) - \varphi$ is $(2\pi)^\nu$-periodic. Thus $\varphi \mapsto i(\varphi)$ is an embedded torus, invariant for the Hamiltonian vector field X_{H_α} and filled by quasi-periodic solutions with frequency ω.

Each Hamiltonian H_α in (5.12) is reversible, i.e. $H_\alpha \circ \tilde{\rho} = H_\alpha$ where the involution $\tilde{\rho}$ is defined in (5.10). We look for reversible solutions of $\mathcal{F}(i,\alpha) = 0$, namely satisfying $\tilde{\rho} i(\varphi) = i(-\varphi)$ (see (5.10)), i.e.

$$\theta(-\varphi) = -\theta(\varphi)\,, \quad I(-\varphi) = I(\varphi)\,, \quad z(-\varphi) = (\rho z)(\varphi)\,. \tag{5.14}$$

The norm of the periodic component of the embedded torus

$$\mathfrak{I}(\varphi) := i(\varphi) - (\varphi, 0, 0) := (\Theta(\varphi), I(\varphi), z(\varphi))\,, \quad \Theta(\varphi) := \theta(\varphi) - \varphi\,, \tag{5.15}$$

is

$$\|\mathfrak{I}\|_s^{k_0,\gamma} := \|\Theta\|_{H^s_\varphi}^{k_0,\gamma} + \|I\|_{H^s_\varphi}^{k_0,\gamma} + \|z\|_s^{k_0,\gamma}\,, \tag{5.16}$$

where $\|z\|_s^{k_0,\gamma} = \|\eta\|_s^{k_0,\gamma} + \|\psi\|_s^{k_0,\gamma}$. We define

$$k_0 := k_0^* + 2, \tag{5.17}$$

where k_0^* is the index of non-degeneracy provided by Proposition 4.4, which only depends on the linear unperturbed frequencies. Thus k_0 is considered as an absolute constant, and we will often omit to explicitly write the dependence of the various constants with respect to k_0. We look for quasi-periodic solutions with frequency ω belonging to a δ-neighborhood (independent of ε)

$$\mathtt{\Omega} := \Big\{\omega \in \mathbb{R}^\nu \,:\, \mathrm{dist}\big(\omega, \vec{\omega}[\mathtt{h}_1, \mathtt{h}_2]\big) < \delta\Big\}, \quad \delta > 0 \tag{5.18}$$

of the unperturbed linear frequencies $\vec{\omega}[\mathtt{h}_1, \mathtt{h}_2]$ defined in (4.1).

Theorem 5.1. (Nash-Moser theorem) *Fix finitely many tangential sites $\mathbb{S}^+ \subset \mathbb{N}^+$ and let $\nu := |\mathbb{S}^+|$. Let $\tau \geq 1$. There exist positive constants $a_0, \varepsilon_0, \kappa_1, C$ depending on $\mathbb{S}^+, \nu, k_0, \tau$ such that, for all $\gamma = \varepsilon^a$, $0 < a < a_0$, for all $\varepsilon \in (0, \varepsilon_0)$, there exist a k_0 times differentiable function*

$$\alpha_\infty : \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] \mapsto \mathbb{R}^\nu\,, \quad \alpha_\infty(\omega, \mathtt{h}) = \omega + r_\varepsilon(\omega, \mathtt{h})\,, \quad \textit{with} \quad |r_\varepsilon|^{k_0,\gamma} \leq C\varepsilon\gamma^{-1}\,, \tag{5.19}$$

a family of embedded tori i_∞ defined for all $\omega \in \mathbb{R}^\nu$ and $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ satisfying the reversibility property (5.14) *and*

$$\| i_\infty(\varphi) - (\varphi, 0, 0) \|_{s_0}^{k_0,\gamma} \le C \varepsilon \gamma^{-1} , \tag{5.20}$$

a sequence of k_0 times differentiable functions $\mu_j^\infty : \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] \to \mathbb{R}$, $j \in \mathbb{N}^+ \setminus \mathbb{S}^+$, of the form

$$\mu_j^\infty(\omega, \mathtt{h}) = \mathtt{m}_{\frac{1}{2}}^\infty(\omega, \mathtt{h}) (j \tanh(\mathtt{h} j))^{\frac{1}{2}} + r_j^\infty(\omega, \mathtt{h}) \tag{5.21}$$

satisfying

$$|\mathtt{m}_{\frac{1}{2}}^\infty - 1|^{k_0,\gamma} \le C\varepsilon \, , \qquad \sup_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+} j^{\frac{1}{2}} |r_j^\infty|^{k_0,\gamma} \le C \varepsilon \gamma^{-\kappa_1} \tag{5.22}$$

such that for all $(\omega, \mathtt{h})$ in the Cantor like set

$$\begin{aligned} \mathcal{C}_\infty^\gamma := \Big\{ & (\omega, \mathtt{h}) \in \Omega \times [\mathtt{h}_1, \mathtt{h}_2] \, : \, |\omega \cdot \ell| \ge 8 \gamma \langle \ell \rangle^{-\tau}, \, \forall \ell \in \mathbb{Z}^\nu \setminus \{0\} , \\ & |\omega \cdot \ell + \mu_j^\infty(\omega, \mathtt{h})| \ge 4 \gamma j^{\frac{1}{2}} \langle \ell \rangle^{-\tau}, \, \forall \ell \in \mathbb{Z}^\nu, \, j \in \mathbb{N}^+ \setminus \mathbb{S}^+, \\ & |\omega \cdot \ell + \mu_j^\infty(\omega, \mathtt{h}) + \mu_{j'}^\infty(\omega, \mathtt{h})| \ge 4 \gamma (j^{\frac{1}{2}} + j'^{\frac{1}{2}}) \langle \ell \rangle^{-\tau}, \; \forall \ell \in \mathbb{Z}^\nu, \; j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ , \\ & |\omega \cdot \ell + \mu_j^\infty(\omega, \mathtt{h}) - \mu_{j'}^\infty(\omega, \mathtt{h})| \ge 4 \gamma j^{-\mathtt{d}} j'^{-\mathtt{d}} \langle \ell \rangle^{-\tau}, \; \forall \ell \in \mathbb{Z}^\nu, \; j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+, \, (\ell, j, j') \neq (0, j, j) \Big\} \end{aligned} \tag{5.23}$$

the function $i_\infty(\varphi) := i_\infty(\omega, \mathtt{h}, \varepsilon)(\varphi)$ is a solution of $\mathcal{F}(i_\infty, \alpha_\infty(\omega, \mathtt{h}), \omega, \mathtt{h}, \varepsilon) = 0$. As a consequence the embedded torus $\varphi \mapsto i_\infty(\varphi)$ is invariant for the Hamiltonian vector field $X_{H_{\alpha_\infty(\omega,\mathtt{h})}}$ and it is filled by quasi-periodic solutions with frequency ω.

Theorem 5.1 is proved in Section 16.1. The very weak second Melnikov non-resonance conditions in (5.23) can be verified for most parameters if $\mathtt{d}$ is large enough, i.e. $\mathtt{d} > \frac{3}{4} k_0^*$, see Theorem 5.2 below. The loss of derivatives produced by such small divisors is compensated in the reducibility scheme of Section 15 by the fact that in Sections 7-14 we will reduce the linearized operator to constant coefficients up to very regularizing terms $O(|D_x|^{-M})$ for some $M := M(\mathtt{d}, \tau)$, fixed in(15.16), large enough with respect to $\mathtt{d}$ and τ by (15.10).

5.2 Measure estimates

The aim is now to deduce Theorem 1.1 from Theorem 5.1.

By (5.19) the function $\alpha_\infty(\cdot, \mathtt{h})$ from Ω into the image $\alpha_\infty(\Omega, \mathtt{h})$ is invertible:

$$\beta = \alpha_\infty(\omega, \mathtt{h}) = \omega + r_\varepsilon(\omega, \mathtt{h}) \quad \Longleftrightarrow \quad \omega = \alpha_\infty^{-1}(\beta, \mathtt{h}) = \beta + \breve{r}_\varepsilon(\beta, \mathtt{h}) \quad \text{with} \quad |\breve{r}_\varepsilon|^{k_0,\gamma} \le C \varepsilon \gamma^{-1} . \tag{5.24}$$

We underline that the function $\alpha_\infty^{-1}(\cdot, \mathtt{h})$ is the inverse of $\alpha_\infty(\cdot, \mathtt{h})$, at any fixed value of $\mathtt{h}$ in $[\mathtt{h}_1, \mathtt{h}_2]$. Then, for any $\beta \in \alpha_\infty(\mathcal{C}_\infty^\gamma)$, Theorem 5.1 proves the existence of an embedded invariant torus filled by quasi-periodic solutions with Diophantine frequency $\omega = \alpha_\infty^{-1}(\beta, \mathtt{h})$ for the Hamiltonian

$$H_\beta = \beta \cdot I + \frac{1}{2} (z, \Omega z)_{L^2} + \varepsilon P \, .$$

Consider the curve of the unperturbed linear frequencies

$$[\mathtt{h}_1, \mathtt{h}_2] \ni \mathtt{h} \mapsto \vec{\omega}(\mathtt{h}) := (\sqrt{j \tanh(\mathtt{h} j)})_{j \in \mathbb{S}^+} \in \mathbb{R}^\nu \, .$$

In Theorem 5.2 below we prove that for "most" values of $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ the vector $(\alpha_\infty^{-1}(\vec{\omega}(\mathtt{h}), \mathtt{h}), \mathtt{h})$ is in $\mathcal{C}_\infty^\gamma$. Hence, for such values of $\mathtt{h}$ we have found an embedded invariant torus for the Hamiltonian H_ε in (5.11), filled by quasi-periodic solutions with Diophantine frequency $\omega = \alpha_\infty^{-1}(\vec{\omega}(\mathtt{h}), \mathtt{h})$.

This implies Theorem 1.1 together with the following measure estimate.

Theorem 5.2. (Measure estimates) *Let*

$$\gamma = \varepsilon^a \,, \quad 0 < a < \min\{a_0, 1/(k_0+\kappa_1)\} \,, \quad \tau > k_0^*(\nu+4) \,, \quad \mathtt{d} > \frac{3k_0^*}{4} \,, \tag{5.25}$$

where k_0^* is the index of non-degeneracy given by Proposition 4.4 and $k_0 = k_0^*+2$. Then the measure of the set

$$\mathcal{G}_\varepsilon = \big\{ \mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2] : \big(\alpha_\infty^{-1}(\vec{\omega}(\mathtt{h}), \mathtt{h}), \mathtt{h}\big) \in \mathcal{C}_\infty^\gamma \big\} \tag{5.26}$$

satisfies $|\mathcal{G}_\varepsilon| \to \mathtt{h}_2 - \mathtt{h}_1$ as $\varepsilon \to 0$.

The rest of this section is devoted to the proof of Theorem 5.2. By (5.24) the vector

$$\omega_\varepsilon(\mathtt{h}) := \alpha_\infty^{-1}(\vec{\omega}(\mathtt{h}), \mathtt{h}) = \vec{\omega}(\mathtt{h}) + \mathtt{r}_\varepsilon(\mathtt{h}) \,, \quad \mathtt{r}_\varepsilon(\mathtt{h}) := \breve{r}_\varepsilon(\vec{\omega}(\mathtt{h}), \mathtt{h}) \,, \tag{5.27}$$

satisfies

$$|\partial_{\mathtt{h}}^k \mathtt{r}_\varepsilon(\mathtt{h})| \le C\varepsilon\gamma^{-k-1} \quad \forall 0 \le k \le k_0 \,. \tag{5.28}$$

We also denote, with a small abuse of notation, for all $j \in \mathbb{N}^+ \setminus \mathbb{S}^+$,

$$\mu_j^\infty(\mathtt{h}) := \mu_j^\infty(\omega_\varepsilon(\mathtt{h}), \mathtt{h}) := \mathtt{m}_{\frac12}^\infty(\mathtt{h})(j \tanh(\mathtt{h}j))^{\frac12} + r_j^\infty(\mathtt{h}), \tag{5.29}$$

where

$$\mathtt{m}_{\frac12}^\infty(\mathtt{h}) := \mathtt{m}_{\frac12}^\infty(\omega_\varepsilon(\mathtt{h}), \mathtt{h}) \,, \qquad r_j^\infty(\mathtt{h}) := r_j^\infty(\omega_\varepsilon(\mathtt{h}), \mathtt{h}). \tag{5.30}$$

By (5.22), (5.30) and (5.27)-(5.28), using that $\varepsilon\gamma^{-k_0-1} \le 1$ (which by (5.25) is satisfied for ε small), we get

$$|\partial_{\mathtt{h}}^k(\mathtt{m}_{\frac12}^\infty(\mathtt{h}) - 1)| \le C\varepsilon\gamma^{-1-k} \,, \qquad \sup_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+} j^{\frac12} |\partial_{\mathtt{h}}^k r_j^\infty(\mathtt{h})| \le C\varepsilon\gamma^{-\kappa_1-k} \qquad \forall 0 \le k \le k_0 \,. \tag{5.31}$$

By (5.23), (5.27), (5.29), the Cantor set $\mathcal{G}_\varepsilon$ in (5.26) becomes

$$\begin{aligned} \mathcal{G}_\varepsilon := \Big\{ & \mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2] \colon |\omega_\varepsilon(\mathtt{h}) \cdot \ell| \ge 8\gamma\langle \ell \rangle^{-\tau}, \, \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \\ & |\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h})| \ge 4\gamma j^{\frac12} \langle \ell \rangle^{-\tau}, \, \forall \ell \in \mathbb{Z}^\nu, \, j \in \mathbb{N}^+ \setminus \mathbb{S}^+, \\ & |\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h}) + \mu_{j'}^\infty(\mathtt{h})| \ge 4\gamma(j^{\frac12} + j'^{\frac12})\langle \ell \rangle^{-\tau}, \; \forall \ell \in \mathbb{Z}^\nu, \; j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \,, \\ & |\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| \ge \frac{4\gamma\langle \ell \rangle^{-\tau}}{j^{\mathtt{d}} j'^{\mathtt{d}}}, \forall \ell \in \mathbb{Z}^\nu, j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+, (\ell, j, j') \ne (0, j, j) \Big\} \,. \end{aligned} \tag{5.32}$$

We estimate the measure of the complementary set

$$\mathcal{G}_\varepsilon^c := [\mathtt{h}_1, \mathtt{h}_2] \setminus \mathcal{G}_\varepsilon := \Big(\bigcup_{\ell \ne 0} R_\ell^{(0)} \Big) \cup \Big(\bigcup_{\ell, j} R_{\ell,j}^{(I)} \Big) \cup \Big(\bigcup_{\ell, j, j'} Q_{\ell j j'}^{(II)} \Big) \cup \Big(\bigcup_{(\ell, j, j') \ne (0, j, j)} R_{\ell j j'}^{(II)} \Big) \tag{5.33}$$

where the "resonant sets" are

$$R_\ell^{(0)} := \big\{ \mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2] : |\omega_\varepsilon(\mathtt{h}) \cdot \ell| < 8\gamma\langle \ell \rangle^{-\tau} \big\} \tag{5.34}$$

$$R_{\ell j}^{(I)} := \big\{ \mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2] : |\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h})| < 4\gamma j^{\frac12} \langle \ell \rangle^{-\tau} \big\} \tag{5.35}$$

$$Q_{\ell j j'}^{(II)} := \big\{ \mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2] : |\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h}) + \mu_{j'}^\infty(\mathtt{h})| < 4\gamma(j^{\frac12} + j'^{\frac12})\langle \ell \rangle^{-\tau} \big\} \tag{5.36}$$

$$R_{\ell j j'}^{(II)} := \Big\{ \mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2] : |\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| < \frac{4\gamma\langle \ell \rangle^{-\tau}}{j^{\mathtt{d}} j'^{\mathtt{d}}} \Big\} \tag{5.37}$$

with $j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+$. We first note that some of these sets are empty.

Lemma 5.3. *For ε, $\gamma \in (0, \gamma_0)$ small, we have that*

1. *If $R_{\ell j}^{(I)} \ne \emptyset$ then $j^{\frac12} \le C\langle \ell \rangle$.*

2. *If* $R^{(II)}_{\ell j j'} \neq \emptyset$ *then* $|j^{\frac12} - j'^{\frac12}| \leq C\langle \ell \rangle$. *Moreover,* $R^{(II)}_{0jj'} = \emptyset$*, for all* $j \neq j'$.

3. *If* $Q^{(II)}_{\ell j j'} \neq \emptyset$ *then* $j^{\frac12} + j'^{\frac12} \leq C\langle \ell \rangle$.

Proof. Let us consider the case of $R^{(II)}_{\ell j j'}$. If $R^{(II)}_{\ell j j'} \neq \emptyset$ there is $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ such that

$$|\mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| < \frac{4\gamma \langle \ell \rangle^{-\tau}}{j^{\mathtt{d}} j'^{\mathtt{d}}} + |\omega_\varepsilon(\mathtt{h}) \cdot \ell| \leq C\langle \ell \rangle\,. \tag{5.38}$$

On the other hand, (5.29), (5.31), and (4.13) imply

$$|\mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| \geq \mathtt{m}^\infty_{\frac12} c |\sqrt{j} - \sqrt{j'}| - C\varepsilon\gamma^{-\kappa_1} \geq \frac{c}{2}|\sqrt{j} - \sqrt{j'}| - 1\,. \tag{5.39}$$

Combining (5.38) and (5.39) we deduce $|j^{\frac12} - j'^{\frac12}| \leq C\langle \ell \rangle$.

Next we prove that $R^{(II)}_{0jj'} = \emptyset$, $\forall j \neq j'$. Recalling (5.29), (5.31), and the definition $\Omega_j(\mathtt{h}) = \sqrt{j \tanh(\mathtt{h} j)}$, we have

$$\begin{aligned} |\mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| &\geq \mathtt{m}^\infty_{\frac12}(\mathtt{h}) |\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})| - \frac{C\varepsilon\gamma^{-\kappa_1}}{j^{\frac12}} - \frac{C\varepsilon\gamma^{-\kappa_1}}{(j')^{\frac12}} \\ &\overset{(4.13)}{\geq} \frac{c}{2}|\sqrt{j} - \sqrt{j'}| - \frac{C\varepsilon\gamma^{-\kappa_1}}{j^{\frac12}} - \frac{C\varepsilon\gamma^{-\kappa_1}}{(j')^{\frac12}}\,. \end{aligned} \tag{5.40}$$

Now we observe that, for any fixed $j \in \mathbb{N}^+$, the minimum of $|\sqrt{j} - \sqrt{j'}|$ over all $j' \in \mathbb{N}^+ \setminus \{j\}$ is attained at $j' = j+1$. By symmetry, this implies that $|\sqrt{j} - \sqrt{j'}|$ is greater or equal than both $(\sqrt{j+1} + \sqrt{j})^{-1}$ and $(\sqrt{j'+1} + \sqrt{j'})^{-1}$. Hence, with $c_0 := 1/(1+\sqrt{2})$, one has

$$|\sqrt{j} - \sqrt{j'}| \geq c_0 \max\left\{\frac{1}{\sqrt{j}}, \frac{1}{\sqrt{j'}}\right\} \geq \frac{c_0}{2}\left(\frac{1}{\sqrt{j}} + \frac{1}{\sqrt{j'}}\right) \geq \frac{c_0}{j^{\frac14}(j')^{\frac14}} \qquad \forall j, j' \in \mathbb{N}^+,\ j \neq j'. \tag{5.41}$$

As a consequence of (5.40) and of the three inequalities in (5.41), for $\varepsilon\gamma^{-\kappa_1}$ small enough, we get for all $j \neq j'$

$$|\mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| \geq \frac{c}{8}|\sqrt{j} - \sqrt{j'}| \geq \frac{4\gamma}{j^{\mathtt{d}} j'^{\mathtt{d}}}\,,$$

for γ small, since $\mathtt{d} \geq 1/4$. This proves that $R^{(II)}_{0jj'} = \emptyset$, for all $j \neq j'$.

The statement for $R^{(I)}_{\ell j}$ and $Q^{(II)}_{\ell j j'}$ is elementary. $\square$

By Lemma 5.3, the last union in (5.33) becomes

$$\bigcup_{(\ell,j,j') \neq (0,j,j)} R^{(II)}_{\ell j j'} = \bigcup_{\substack{\ell \neq 0 \\ |\sqrt{j} - \sqrt{j'}| \leq C\langle \ell \rangle}} R^{(II)}_{\ell j j'}\,. \tag{5.42}$$

In order to estimate the measure of the sets (5.34)-(5.37) that are nonempty, the key point is to prove that the perturbed frequencies satisfy estimates similar to (4.9)-(4.12) in Proposition 4.4.

Lemma 5.4. (Perturbed transversality) *For* ε *small enough, for all* $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$*,*

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\omega_\varepsilon(\mathtt{h}) \cdot \ell\}| \geq \frac{\rho_0}{2}\langle \ell \rangle \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \tag{5.43}$$

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h})\}| \geq \frac{\rho_0}{2}\langle \ell \rangle \quad \forall \ell \in \mathbb{Z}^\nu,\ j \in \mathbb{N}^+ \setminus \mathbb{S}^+ : j^{\frac12} \leq C\langle \ell \rangle, \tag{5.44}$$

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})\}| \geq \frac{\rho_0}{2}\langle \ell \rangle \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\},\ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ : |j^{\frac12} - j'^{\frac12}| \leq C\langle \ell \rangle, \tag{5.45}$$

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h}) + \mu_{j'}^\infty(\mathtt{h})\}| \geq \frac{\rho_0}{2}\langle \ell \rangle \quad \forall \ell \in \mathbb{Z}^\nu,\ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ : j^{\frac12} + j'^{\frac12} \leq C\langle \ell \rangle, \tag{5.46}$$

where k_0^* *is the index of non-degeneracy given by Proposition 4.4.*

Proof. The most delicate estimate is (5.45). We split

$$\mu_j^\infty(\mathtt{h}) = \Omega_j(\mathtt{h}) + (\mu_j^\infty - \Omega_j)(\mathtt{h})$$

where $\Omega_j(\mathtt{h}) := j^{\frac{1}{2}}(\tanh(j\mathtt{h}))^{\frac{1}{2}}$. A direct calculation using (1.24) and (5.41) shows that, for $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$,

$$|\partial_{\mathtt{h}}^k \{\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})\}| \leq C_k |j^{\frac{1}{2}} - j'^{\frac{1}{2}}| \quad \forall\, k \geq 0. \tag{5.47}$$

Then, using (5.31), one has, for all $0 \leq k \leq k_0$,

$$\begin{aligned} |\partial_{\mathtt{h}}^k \{(\mu_j^\infty - \mu_{j'}^\infty)(\mathtt{h}) - (\Omega_j - \Omega_{j'})(\mathtt{h})\}| &\leq |\partial_{\mathtt{h}}^k \{(\mathtt{m}_{\frac{1}{2}}^\infty(\mathtt{h}) - 1)(\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h}))\}| \\ &\quad + |\partial_{\mathtt{h}}^k r_j^\infty(\mathtt{h})| + |\partial_{\mathtt{h}}^k r_{j'}^\infty(\mathtt{h})| \\ &\overset{(5.47)}{\leq} C_{k_0}\{\varepsilon\gamma^{-1-k}|j^{\frac{1}{2}} - j'^{\frac{1}{2}}| + \varepsilon\gamma^{-\kappa_1-k}(j^{-\frac{1}{2}} + (j')^{-\frac{1}{2}})\} \\ &\overset{(5.41)}{\leq} C'_{k_0}\varepsilon\gamma^{-\kappa_1-k}|j^{\frac{1}{2}} - j'^{\frac{1}{2}}|\,. \end{aligned} \tag{5.48}$$

Recall that $k_0 = k_0^* + 2$ (see (5.17)). By (5.28) and (5.48), using $|j^{\frac{1}{2}} - j'^{\frac{1}{2}}| \leq C\langle \ell \rangle$, we get

$$\begin{aligned} \max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})\}| &\geq \max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})\}| - C\varepsilon\gamma^{-(1+k_0^*)}|\ell| \\ &\quad - C\varepsilon\gamma^{-(k_0^*+\kappa_1)}|j^{\frac{1}{2}} - j'^{\frac{1}{2}}| \\ &\geq \max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})\}| - C\varepsilon\gamma^{-(k_0^*+\kappa_1)}\langle \ell \rangle \\ &\overset{(4.11)}{\geq} \rho_0 \langle \ell \rangle - C\varepsilon\gamma^{-(k_0^*+\kappa_1)}\langle \ell \rangle \geq \rho_0\langle \ell \rangle / 2 \end{aligned}$$

provided $\varepsilon\gamma^{-(k_0^*+\kappa_1)} \leq \rho_0/(2C)$, which, by (5.25), is satisfied for ε small enough. □

As an application of Rüssmann Theorem 17.1 in [57] we deduce the following

Lemma 5.5. (Estimates of the resonant sets) *The measure of the sets in* (5.34)-(5.37) *satisfies*

$$\begin{aligned} &|R_\ell^{(0)}| \lesssim \big(\gamma\langle \ell \rangle^{-(\tau+1)}\big)^{\frac{1}{k_0^*}} \quad \forall \ell \neq 0\,, \qquad |R_{\ell j}^{(I)}| \lesssim \big(\gamma j^{\frac{1}{2}}\langle \ell \rangle^{-(\tau+1)}\big)^{\frac{1}{k_0^*}}\,, \\ &|R_{\ell j j'}^{(II)}| \lesssim \Big(\gamma \frac{\langle \ell \rangle^{-(\tau+1)}}{j^{\mathtt{d}} j'^{\mathtt{d}}}\Big)^{\frac{1}{k_0^*}} \quad \forall \ell \neq 0, \qquad |Q_{\ell j j'}^{(II)}| \lesssim \big(\gamma(j^{\frac{1}{2}} + j'^{\frac{1}{2}})\langle \ell \rangle^{-(\tau+1)}\big)^{\frac{1}{k_0^*}}\,. \end{aligned}$$

Proof. We prove the estimate of $R_{\ell j j'}^{(II)}$ in (5.37). The other cases are simpler. We write

$$R_{\ell j j'}^{(II)} = \Big\{\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2] : |f_{\ell j j'}(\mathtt{h})| < \frac{4\gamma}{\langle \ell \rangle^{\tau+1} j^{\mathtt{d}} j'^{\mathtt{d}}}\Big\}$$

where $f_{\ell j j'}(\mathtt{h}) := (\omega_\varepsilon(\mathtt{h}) \cdot \ell + \mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h}))\langle \ell \rangle^{-1}$. By (5.42), we restrict to the case $|j^{\frac{1}{2}} - j'^{\frac{1}{2}}| \leq C\langle \ell \rangle$ and $\ell \neq 0$. By (5.45),

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k f_{\ell j j'}(\mathtt{h})| \geq \rho_0/2\,, \quad \forall \mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]\,.$$

In addition, (5.27)-(5.31) and Lemma 5.3 imply that $\max_{k \leq k_0} |\partial_{\mathtt{h}}^k f_{\ell j j'}(\mathtt{h})| \leq C$ for all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, provided $\varepsilon\gamma^{-(k_0+\kappa_1)}$ is small enough, namely, by (5.25), ε is small enough. In particular, $f_{\ell j j'}$ belongs to $\mathrm{Lip}(k_0)$, and therefore it is of class $\mathcal{C}^{k_0-1} = \mathcal{C}^{k_0^*+1}$. Thus Theorem 17.1 in [57] applies, whence the lemma follows. □

Proof of Theorem 5.2 completed. By Lemma 5.3 (in particular, recalling that $R^{(II)}_{\ell jj'}$ is empty for $\ell = 0$ and $j \neq j'$, see (5.42)) and Lemma 5.5, the measure of the set $\mathcal{G}^c_\varepsilon$ in (5.33) is estimated by

$$\begin{aligned} |\mathcal{G}^c_\varepsilon| &\le \sum_{\ell\neq 0} |R^{(0)}_\ell| + \sum_{\ell,j} |R^{(I)}_{\ell j}| + \sum_{(\ell,j,j')\neq(0,j,j)} |R^{(II)}_{\ell jj'}| + \sum_{\ell,j,j'} |Q^{(II)}_{\ell jj'}| \\ &\le \sum_{\ell\neq 0} |R^{(0)}_\ell| + \sum_{j\le C\langle\ell\rangle^2} |R^{(I)}_{\ell j}| + \sum_{\substack{\ell\neq 0\\ |\sqrt{j}-\sqrt{j'}|\le C\langle\ell\rangle}} |R^{(II)}_{\ell jj'}| + \sum_{j,j'\le C\langle\ell\rangle^2} |Q^{(II)}_{\ell jj'}| \\ &\lesssim \sum_\ell \Big(\frac{\gamma}{\langle\ell\rangle^{\tau+1}}\Big)^{\frac{1}{k_0^*}} + \sum_{j\le C\langle\ell\rangle^2} \Big(\frac{\gamma j^{\frac12}}{\langle\ell\rangle^{\tau+1}}\Big)^{\frac{1}{k_0^*}} + \sum_{|\sqrt{j}-\sqrt{j'}|\le C\langle\ell\rangle} \Big(\frac{\gamma}{\langle\ell\rangle^{\tau+1} j^{\mathtt{d}} j'^{\mathtt{d}}}\Big)^{\frac{1}{k_0^*}} + \sum_{j,j'\le C\langle\ell\rangle^2} \Big(\frac{\gamma(j^{\frac12}+j'^{\frac12})}{\langle\ell\rangle^{\tau+1}}\Big)^{\frac{1}{k_0^*}} \\ &\le C\gamma^{\frac{1}{k_0^*}} \Big\{ \sum_{\ell\in\mathbb{Z}^\nu} \frac{1}{\langle\ell\rangle^{\frac{\tau}{k_0^*}-4}} + \sum_{|\sqrt{j}-\sqrt{j'}|\le C\langle\ell\rangle} \frac{1}{\langle\ell\rangle^{\frac{\tau+1}{k_0^*}} j^{\frac{\mathtt{d}}{k_0^*}} j'^{\frac{\mathtt{d}}{k_0^*}}} \Big\}. \end{aligned} \tag{5.49}$$

The first series in (5.49) converges because $\frac{\tau}{k_0^*} - 4 > \nu$ by (5.25). For the second series in (5.49), we observe that the sum is symmetric in (j,j') and, for $j \le j'$, the bound $|\sqrt{j} - \sqrt{j'}| \le C\langle\ell\rangle$ implies that $j \le j' \le j + C^2\langle\ell\rangle^2 + 2C\sqrt{j}\langle\ell\rangle$. Since

$$\forall \ell, j, \qquad \sum_{j'=j}^{j+p} \frac{1}{j'^{\frac{\mathtt{d}}{k_0^*}}} \le \sum_{j'=j}^{j+p} \frac{1}{j^{\frac{\mathtt{d}}{k_0^*}}} = \frac{p+1}{j^{\frac{\mathtt{d}}{k_0^*}}}, \quad p := C^2\langle\ell\rangle^2 + 2C\sqrt{j}\langle\ell\rangle,$$

the second series in (5.49) converges because $\frac{\tau+1}{k_0^*} - 2 > \nu$ and $2\frac{\mathtt{d}}{k_0^*} - \frac12 > 1$ by (5.25). By (5.49) we get

$$|\mathcal{G}^c_\varepsilon| \le C\gamma^{\frac{1}{k_0^*}}.$$

In conclusion, for $\gamma = \varepsilon^a$, we find $|\mathcal{G}_\varepsilon| \ge \mathtt{h}_2 - \mathtt{h}_1 - C\varepsilon^{a/k_0^*}$ and the proof of Theorem 5.2 is concluded.

6 Approximate inverse

6.1 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 (5.13).

We first estimate the composition operator induced by $v(\theta, y)$ defined in (5.9). Since the functions $I_j \mapsto \sqrt{\xi_j + I_j}$, $\theta \mapsto \cos(\theta)$, $\theta \mapsto \sin(\theta)$ are analytic for $|I| \le r$ small, the composition Lemma 2.6 implies that, for all $\Theta, y \in H^s(\mathbb{T}^\nu, \mathbb{R}^\nu)$, $\|\Theta\|_{s_0}, \|y\|_{s_0} \le r$, setting $\theta(\varphi) := \varphi + \Theta(\varphi)$,

$$\|\partial_\theta^\alpha \partial_I^\beta v(\theta(\cdot), I(\cdot))\|_s^{k_0,\gamma} \lesssim_s 1 + \|\mathfrak{I}\|_s^{k_0,\gamma}, \quad \forall \alpha, \beta \in \mathbb{N}^\nu, \quad |\alpha| + |\beta| \le 3. \tag{6.1}$$

Lemma 6.1. *Let $\mathfrak{I}(\varphi)$ in (5.15) satisfy $\|\mathfrak{I}\|_{3s_0+2k_0+5}^{k_0,\gamma} \le 1$. Then the following estimates hold:*

$$\|X_P(i)\|_s^{k_0,\gamma} \lesssim_s 1 + \|\mathfrak{I}\|_{s+2s_0+2k_0+3}^{k_0,\gamma}, \tag{6.2}$$

and for all $\widehat{\imath} := (\widehat{\theta}, \widehat{I}, \widehat{z})$

$$\|d_i X_P(i)[\widehat{\imath}]\|_s^{k_0,\gamma} \lesssim_s \|\widehat{\imath}\|_{s+1}^{k_0,\gamma} + \|\mathfrak{I}\|_{s+2s_0+2k_0+4}^{k_0,\gamma} \|\widehat{\imath}\|_{s_0+1}^{k_0,\gamma}, \tag{6.3}$$

$$\|d_i^2 X_P(i)[\widehat{\imath}, \widehat{\imath}]\|_s^{k_0,\gamma} \lesssim_s \|\widehat{\imath}\|_{s+1}^{k_0,\gamma} \|\widehat{\imath}\|_{s_0+1}^{k_0,\gamma} + \|\mathfrak{I}\|_{s+2s_0+2k_0+5}^{k_0,\gamma} (\|\widehat{\imath}\|_{s_0+1}^{k_0,\gamma})^2. \tag{6.4}$$

Proof. By definition (5.11), $P = P_\varepsilon \circ A$, where A is defined in (5.9) and P_ε is defined in (5.3). Hence

$$X_P = \Big([\partial_I v(\theta, I)]^T \nabla P_\varepsilon(A(\theta, I, z)), -[\partial_\theta v(\theta, I)]^T \nabla P_\varepsilon(A(\theta, I, z)), \Pi^\perp_{\mathbb{S}^+} J \nabla P_\varepsilon(A(\theta, I, z)) \Big) \tag{6.5}$$

where $\Pi_{\mathbb{S}+}^{\perp}$ is the L^2-projector on the space $H_{\mathbb{S}+}^{\perp}$ defined in (5.5). Now $\nabla P_\varepsilon = -JX_{P_\varepsilon}$ (see (5.1)), where X_{P_ε} is the explicit Hamiltonian vector field in (5.2). The smallness condition of Proposition 3.1 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 (3.5) for the Dirichlet-Neumann operator (applied for $m, \alpha = 0$), the interpolation inequality (2.10), and (6.1), we get

$$\|\nabla P_\varepsilon(A(\theta(\cdot), I(\cdot), z(\cdot,\cdot)))\|_s^{k_0,\gamma} \lesssim_s \|A(\theta(\cdot), I(\cdot), z(\cdot,\cdot))\|_{s+2s_0+2k_0+3}^{k_0,\gamma} \lesssim_s 1 + \|\mathfrak{I}\|_{s+2s_0+2k_0+3}^{k_0,\gamma} .$$

Hence (6.2) follows by (6.5), interpolation and (6.1).

Estimates (6.3), (6.4) for $d_i X_P$ and $d_i^2 X_P$ follow by differentiating the expression of X_P in (6.5), applying the estimates of Proposition 3.1 on the Dirichlet-Neumann operator and estimate (6.1) on $v(\theta, y)$ and using the interpolation inequality (2.10). ☐

6.2 Almost-approximate inverse

In order to implement a convergent Nash-Moser scheme that leads to a solution of $\mathcal{F}(i,\alpha) = 0$ we construct an *almost-approximate right inverse* of the linearized operator

$$d_{i,\alpha}\mathcal{F}(i_0, \alpha_0)[\widehat{\imath}, \widehat{\alpha}] = \omega \cdot \partial_\varphi \widehat{\imath} - d_i X_{H_\alpha}(i_0(\varphi))[\widehat{\imath}] - (\widehat{\alpha}, 0, 0) .$$

Note that $d_{i,\alpha}\mathcal{F}(i_0, \alpha_0) = d_{i,\alpha}\mathcal{F}(i_0)$ is independent of α_0, see (5.13) and recall that the perturbation P does not depend on α.

We implement the general strategy in [16], [8], and we shall closely follow [21]. An invariant torus i_0 with Diophantine flow is isotropic (see e.g. [16]), namely the pull-back 1-form $i_0^*\Lambda$ is closed, where Λ is the 1-form in (5.7). This is tantamount to say that the 2-form $i_0^*\mathcal{W} = i_0^* d\Lambda = d i_0^*\Lambda = 0$. For an "approximately invariant" torus i_0 the 1-form $i_0^*\Lambda$ is only "approximately closed". In order to make this statement quantitative we consider

$$i_0^*\Lambda = \sum\nolimits_{k=1}^{\nu} a_k(\varphi) d\varphi_k , \quad a_k(\varphi) := -\big([\partial_\varphi \theta_0(\varphi)]^T I_0(\varphi)\big)_k - \frac{1}{2}(\partial_{\varphi_k} z_0(\varphi), J z_0(\varphi))_{L^2(\mathbb{T}_x)} \tag{6.6}$$

and we quantify how small is

$$i_0^*\mathcal{W} = d\, i_0^*\Lambda = \sum\nolimits_{1\leq k<j\leq \nu} A_{kj}(\varphi) d\varphi_k \wedge d\varphi_j , \quad A_{kj}(\varphi) := \partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi) \tag{6.7}$$

in terms of the "error function"

$$Z(\varphi) := (Z_1, Z_2, Z_3)(\varphi) := \mathcal{F}(i_0, \alpha_0)(\varphi) = \omega \cdot \partial_\varphi i_0(\varphi) - X_{H_\alpha}(i_0(\varphi), \alpha_0) . \tag{6.8}$$

Along this section we will always assume the following hypothesis, which will be verified at each step of the Nash-Moser iteration.

- ANSATZ. The map $(\omega, \mathtt{h}) \mapsto \mathfrak{I}_0(\omega, \mathtt{h}) := i_0(\varphi; \omega, \mathtt{h}) - (\varphi, 0, 0)$ is k_0 times differentiable with respect to the parameters $(\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, and for some $\mu := \mu(\tau, \nu) > 0$, $\gamma \in (0, 1)$,

$$\|\mathfrak{I}_0\|_{s_0+\mu}^{k_0,\gamma} + |\alpha_0 - \omega|^{k_0,\gamma} \leq C\varepsilon\gamma^{-1} , \tag{6.9}$$

  For some $\kappa := \kappa(\tau, \nu) > 0$, we shall always assume the smallness condition $\varepsilon\gamma^{-\kappa} \ll 1$.

We suppose that the torus $i_0(\omega, \mathtt{h})$ is defined for all the values of $(\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ because, in the Nash-Moser iteration we construct a k_0 times differentiable extension of each approximate solution on the whole $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$.

Lemma 6.2. $\|Z\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1} + \|\mathfrak{I}_0\|_{s+2}^{k_0,\gamma}$.

Proof. By (5.13), (6.2), (6.9). ☐

Lemma 6.3. *Assume that ω belongs to $\mathtt{DC}(\gamma,\tau)$ defined in* (2.13)*. Then the coefficients A_{kj} in* (6.7) *satisfy*

$$\|A_{kj}\|_s^{k_0,\gamma} \lesssim_s \gamma^{-1}\big(\|Z\|^{k_0,\gamma}_{s+\tau(k_0+1)+k_0+1} + \|Z\|^{k_0,\gamma}_{s_0+1}\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\tau(k_0+1)+k_0+1}\big)\,. \tag{6.10}$$

Proof. The coefficients A_{kj} satisfy the identity (see [16], 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 $\underline{e}_k$ denotes the k-th versor of $\mathbb{R}^\nu$. Then by (6.9) we get

$$\|\omega\cdot\partial_\varphi A_{kj}\|_s^{k_0,\gamma} \lesssim_s \|Z\|^{k_0,\gamma}_{s+1} + \|Z\|^{k_0,\gamma}_{s_0+1}\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+1}\,. \tag{6.11}$$

Then (6.10) follows applying $(\omega\cdot\partial_\varphi)^{-1}$, since, by Lemma 2.5, $\|(\omega\cdot\partial_\varphi)^{-1}g\|_s^{k_0,\gamma} \lesssim_s \gamma^{-1}\|g\|^{k_0,\gamma}_{s+\tau(k_0+1)+k_0}$. ☐

As in [16], [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 6.4. (Isotropic torus) *The torus $i_\delta(\varphi) := (\theta_0(\varphi), I_\delta(\varphi), z_0(\varphi))$ defined by*

$$I_\delta := I_0 + [\partial_\varphi\theta_0(\varphi)]^{-T}\rho(\varphi)\,, \qquad \rho_j(\varphi) := \Delta_\varphi^{-1}\sum_{k=1}^{\nu}\partial_{\varphi_j}A_{kj}(\varphi) \tag{6.12}$$

is isotropic. There is $\sigma := \sigma(\nu,\tau,k_0)$ such that

$$\|I_\delta - I_0\|_s^{k_0,\gamma} \leq \|I_0\|^{k_0,\gamma}_{s+1} \tag{6.13}$$
$$\|I_\delta - I_0\|_s^{k_0,\gamma} \lesssim_s \gamma^{-1}\big(\|Z\|^{k_0,\gamma}_{s+\sigma} + \|Z\|^{k_0,\gamma}_{s_0+\sigma}\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma}\big)\,, \tag{6.14}$$
$$\|\mathcal{F}(i_\delta,\alpha_0)\|_s^{k_0,\gamma} \lesssim_s \|Z\|^{k_0,\gamma}_{s+\sigma} + \|Z\|^{k_0,\gamma}_{s_0+\sigma}\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma} \tag{6.15}$$
$$\|\partial_i[i_\delta][\widehat{\imath}]\|_s^{k_0,\gamma} \lesssim_s \|\widehat{\imath}\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma}\|\widehat{\imath}\|^{k_0,\gamma}_{s_0}\,. \tag{6.16}$$

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. Estimates (6.13), (6.14) follow as in [8] by (6.12), (6.6), (6.7), (6.10), (6.9). The difference

$$\mathcal{F}(i_\delta,\alpha_0) - \mathcal{F}(i_0,\alpha_0) = \begin{pmatrix} 0 \\ \omega\cdot\partial_\varphi(I_\delta - I_0) \\ 0 \end{pmatrix} + \varepsilon\big(X_P(i_\delta) - X_P(i_0)\big)$$

where, as proved in [16], [8],

$$\omega\cdot\partial_\varphi(I_\delta - I_0) = [\partial_\varphi\theta_0(\varphi)]^{-T}\omega\cdot\partial_\varphi\rho(\varphi) - \big([\partial_\varphi\theta_0(\varphi)]^{-T}\big(\omega\cdot\partial_\varphi[\partial_\varphi\theta_0(\varphi)]^T\big)[\partial_\varphi\theta_0(\varphi)]^{-T}\big)\rho(\varphi)\,,$$
$$\omega\cdot\partial_\varphi[\partial_\varphi\theta_0(\varphi)] = \varepsilon\partial_\varphi(\partial_I P)(i_0(\varphi)) + \partial_\varphi Z_1(\varphi)\,.$$

Then (6.15) follows by (6.3), (6.14), (6.9), Lemma 6.2, (6.11), (6.10). The bound (6.16) follows by (6.12), (6.7), (6.6), (6.9). ☐

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 $\mathbb{T}^\nu\times\mathbb{R}^\nu\times H^\perp_{\mathbb{S}^+}$ defined by

$$\begin{pmatrix}\theta\\ I\\ z\end{pmatrix} := G_\delta\begin{pmatrix}\phi\\ y\\ w\end{pmatrix} := \begin{pmatrix}\theta_0(\phi)\\ I_\delta(\phi) + [\partial_\phi\theta_0(\phi)]^{-T}y - \big[(\partial_\theta\tilde{z}_0)(\theta_0(\phi))\big]^T Jw\\ z_0(\phi) + w\end{pmatrix} \tag{6.17}$$

where $\tilde{z}_0(\theta) := z_0(\theta_0^{-1}(\theta))$. It is proved in [16] that G_δ is symplectic, because the torus i_δ is isotropic (Lemma 6.4). 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 (5.12)) changes into

$$X_{K_\alpha} = (DG_\delta)^{-1}X_{H_\alpha}\circ G_\delta \qquad \text{where} \qquad K_\alpha := H_\alpha\circ G_\delta\,. \tag{6.18}$$

By (5.14) 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

$$K_\alpha(\phi, y, w) = K_{00}(\phi, \alpha) + K_{10}(\phi, \alpha) \cdot y + (K_{01}(\phi, \alpha), w)_{L^2(\mathbb{T}_x)} + \frac{1}{2} K_{20}(\phi) y \cdot y \\ + \big(K_{11}(\phi) y, w\big)_{L^2(\mathbb{T}_x)} + \frac{1}{2}\big(K_{02}(\phi) w, w\big)_{L^2(\mathbb{T}_x)} + K_{\geq 3}(\phi, y, w) \tag{6.19}$$

where $K_{\geq 3}$ collects the terms at least cubic in the variables (y, w). The Taylor coefficient $K_{00}(\phi, \alpha) \in \mathbb{R}$, $K_{10}(\phi, \alpha) \in \mathbb{R}^\nu$, $K_{01}(\phi, \alpha) \in H^\perp_{\mathbb{S}^+}$, $K_{20}(\phi)$ is a $\nu \times \nu$ real matrix, $K_{02}(\phi)$ is a linear self-adjoint operator of $H^\perp_{\mathbb{S}^+}$ and $K_{11}(\phi) \in \mathcal{L}(\mathbb{R}^\nu, H^\perp_{\mathbb{S}^+})$.

Note that, by (5.12) and (6.17), the only Taylor coefficients that depend on α are K_{00}, K_{10}, K_{01}.

The Hamilton equations associated to (6.19) are

$$\begin{cases} \dot{\phi} = K_{10}(\phi, \alpha) + K_{20}(\phi) y + K_{11}^T(\phi) w + \partial_y K_{\geq 3}(\phi, y, w) \\ \dot{y} = \partial_\phi K_{00}(\phi, \alpha) - [\partial_\phi K_{10}(\phi, \alpha)]^T y - [\partial_\phi K_{01}(\phi, \alpha)]^T w \\ \quad -\partial_\phi\big(\frac{1}{2} K_{20}(\phi) y \cdot y + (K_{11}(\phi) y, w)_{L^2(\mathbb{T}_x)} + \frac{1}{2}(K_{02}(\phi) w, w)_{L^2(\mathbb{T}_x)} + K_{\geq 3}(\phi, y, w)\big) \\ \dot{w} = J\big(K_{01}(\phi, \alpha) + K_{11}(\phi) y + K_{02}(\phi) w + \nabla_w K_{\geq 3}(\phi, y, w)\big) \end{cases} \tag{6.20}$$

where $\partial_\phi K_{10}^T$ is the $\nu \times \nu$ transposed matrix and $\partial_\phi K_{01}^T$, $K_{11}^T : H^\perp_{\mathbb{S}^+} \to \mathbb{R}^\nu$ are defined by the duality relation $(\partial_\phi K_{01}[\hat{\phi}], w)_{L^2_x} = \hat{\phi} \cdot [\partial_\phi K_{01}]^T w$, $\forall \hat{\phi} \in \mathbb{R}^\nu, w \in H^\perp_{\mathbb{S}^+}$, and similarly for K_{11}. Explicitly, for all $w \in H^\perp_{\mathbb{S}^+}$, and denoting by $\underline{e}_k$ the k-th versor of $\mathbb{R}^\nu$,

$$K_{11}^T(\phi) w = \sum\nolimits_{k=1}^{\nu} \big(K_{11}^T(\phi) w \cdot \underline{e}_k\big) \underline{e}_k = \sum\nolimits_{k=1}^{\nu} \big(w, K_{11}(\phi) \underline{e}_k\big)_{L^2(\mathbb{T}_x)} \underline{e}_k \in \mathbb{R}^\nu . \tag{6.21}$$

The coefficients K_{00}, K_{10}, K_{01} in the Taylor expansion (6.19) vanish on an exact solution (i.e. $Z = 0$).

Lemma 6.5. *We have*

$$\|\partial_\phi K_{00}(\cdot, \alpha_0)\|_s^{k_0,\gamma} + \|K_{10}(\cdot, \alpha_0) - \omega\|_s^{k_0,\gamma} + \|K_{01}(\cdot, \alpha_0)\|_s^{k_0,\gamma} \lesssim_s \|Z\|_{s+\sigma}^{k_0,\gamma} + \|Z\|_{s_0+\sigma}^{k_0,\gamma} \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} . \tag{6.22}$$

Proof. In Lemma 8 of [16] or Lemma 6.4 of [8] the following identities are proved

$$\begin{aligned} \partial_\phi K_{00}(\phi, \alpha_0) &= -[\partial_\phi \theta_0(\phi)]^T \big(- Z_{2,\delta} - [\partial_\phi I_\delta][\partial_\phi \theta_0]^{-1} Z_{1,\delta} - [(\partial_\theta \tilde{z}_0)(\theta_0(\phi))]^T J Z_{3,\delta} \\ &\quad - [(\partial_\theta \tilde{z}_0)(\theta_0(\phi))]^T J \partial_\phi z_0(\phi) [\partial_\phi \theta_0(\phi)]^{-1} Z_{1,\delta}\big) , \\ K_{10}(\phi, \alpha_0) &= \omega - [\partial_\phi \theta_0(\phi)]^{-1} Z_{1,\delta}(\phi) , \\ K_{01}(\phi, \alpha_0) &= J Z_{3,\delta} - J \partial_\phi z_0(\phi) [\partial_\phi \theta_0(\phi)]^{-1} Z_{1,\delta}(\phi) , \end{aligned}$$

where $Z_\delta = (Z_{1,\delta}, Z_{2,\delta}, Z_{3,\delta}) := \mathcal{F}(i_\delta, \alpha_0)$. Then (6.9), (6.14), (6.15) imply (6.22). ☐

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 6.6. *We have*

$$\|\partial_\alpha K_{00}\|_s^{k_0,\gamma} + \|\partial_\alpha K_{10} - \mathrm{Id}\|_s^{k_0,\gamma} + \|\partial_\alpha K_{01}\|_s^{k_0,\gamma} \lesssim_s \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} , \quad \|K_{20}\|_s^{k_0,\gamma} \lesssim_s \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) ,$$
$$\|K_{11} y\|_s^{k_0,\gamma} \lesssim_s \varepsilon\big(\|y\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \|y\|_{s_0}^{k_0,\gamma}\big) , \quad \|K_{11}^T w\|_s^{k_0,\gamma} \lesssim_s \varepsilon\big(\|w\|_{s+2}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \|w\|_{s_0+2}^{k_0,\gamma}\big) .$$

Proof. By [16], [8] we have

$$\begin{aligned} &\partial_\alpha K_{00}(\phi) = I_\delta(\phi) , \quad \partial_\alpha K_{10}(\phi) = [\partial_\phi \theta_0(\phi)]^{-1} , \quad \partial_\alpha K_{01}(\phi) = J \partial_\theta \tilde{z}_0(\theta_0(\phi)) , \\ &K_{20}(\varphi) = \varepsilon [\partial_\varphi \theta_0(\varphi)]^{-1} \partial_{II} P(i_\delta(\varphi)) [\partial_\varphi \theta_0(\varphi)]^{-T} , \\ &K_{11}(\varphi) = \varepsilon\big(\partial_I \nabla_z P(i_\delta(\varphi)) [\partial_\varphi \theta_0(\varphi)]^{-T} + J(\partial_\theta \tilde{z}_0)(\theta_0(\varphi))(\partial_{II} P)(i_\delta(\varphi)) [\partial_\varphi \theta_0(\varphi)]^{-T}\big) . \end{aligned}$$

Then (6.2), (6.9), (6.13) imply the lemma (the bound for K_{11}^T follows by (6.21)). ☐

Under the linear change of variables

$$DG_\delta(\varphi,0,0)\begin{pmatrix}\widehat{\phi}\\ \widehat{y}\\ \widehat{w}\end{pmatrix} := \begin{pmatrix}\partial_\phi\theta_0(\varphi) & 0 & 0\\ \partial_\phi I_\delta(\varphi) & [\partial_\phi\theta_0(\varphi)]^{-T} & -[(\partial_\theta\tilde{z}_0)(\theta_0(\varphi))]^T J\\ \partial_\phi z_0(\varphi) & 0 & I\end{pmatrix}\begin{pmatrix}\widehat{\phi}\\ \widehat{y}\\ \widehat{w}\end{pmatrix} \tag{6.23}$$

the linearized operator $d_{i,\alpha}\mathcal{F}(i_\delta)$ is transformed (approximately) into the one obtained when one linearizes the Hamiltonian system (6.20) at $(\phi,y,w)=(\varphi,0,0)$, differentiating also in α at α_0, and changing $\partial_t \rightsquigarrow \omega\cdot\partial_\varphi$, namely

$$\begin{pmatrix}\widehat{\phi}\\ \widehat{y}\\ \widehat{w}\\ \widehat{\alpha}\end{pmatrix} \mapsto \begin{pmatrix}\omega\cdot\partial_\varphi\widehat{\phi} - \partial_\phi K_{10}(\varphi)[\widehat{\phi}\,] - \partial_\alpha K_{10}(\varphi)[\widehat{\alpha}] - K_{20}(\varphi)\widehat{y} - K_{11}^T(\varphi)\widehat{w}\\ \omega\cdot\partial_\varphi\widehat{y} + \partial_{\phi\phi}K_{00}(\varphi)[\widehat{\phi}] + \partial_\phi\partial_\alpha K_{00}(\varphi)[\widehat{\alpha}] + [\partial_\phi K_{10}(\varphi)]^T\widehat{y} + [\partial_\phi K_{01}(\varphi)]^T\widehat{w}\\ \omega\cdot\partial_\varphi\widehat{w} - J\{\partial_\phi K_{01}(\varphi)[\widehat{\phi}] + \partial_\alpha K_{01}(\varphi)[\widehat{\alpha}] + K_{11}(\varphi)\widehat{y} + K_{02}(\varphi)\widehat{w}\}\end{pmatrix}. \tag{6.24}$$

As in [8], by (6.23), (6.9), (6.13), the induced composition operator satisfies: for all $\widehat{\imath} := (\widehat{\phi},\widehat{y},\widehat{w})$

$$\|DG_\delta(\varphi,0,0)[\widehat{\imath}]\|_s^{k_0,\gamma} + \|DG_\delta(\varphi,0,0)^{-1}[\widehat{\imath}]\|_s^{k_0,\gamma} \lesssim_s \|\widehat{\imath}\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|\widehat{\imath}\|_{s_0}^{k_0,\gamma}, \tag{6.25}$$

$$\|D^2G_\delta(\varphi,0,0)[\widehat{\imath}_1,\widehat{\imath}_2]\|_s^{k_0,\gamma} \lesssim_s \|\widehat{\imath}_1\|_s^{k_0,\gamma}\|\widehat{\imath}_2\|_{s_0}^{k_0,\gamma} + \|\widehat{\imath}_1\|_{s_0}^{k_0,\gamma}\|\widehat{\imath}_2\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|\widehat{\imath}_1\|_{s_0}^{k_0,\gamma}\|\widehat{\imath}_2\|_{s_0}^{k_0,\gamma}. \tag{6.26}$$

In order to construct an "almost-approximate" inverse of (6.24) we need that

$$\mathcal{L}_\omega := \Pi_{\mathbb{S}^+}^\perp\big(\omega\cdot\partial_\varphi - JK_{02}(\varphi)\big)_{|H_{\mathbb{S}^+}^\perp} \tag{6.27}$$

is "almost-invertible" up to remainders of size $O(N_{n-1}^{-\mathtt{a}})$ (see precisely (6.31)) where

$$N_n := K_n^p, \quad \forall n\geq 0, \tag{6.28}$$

and

$$K_n := K_0^{\chi^n}, \quad \chi := 3/2 \tag{6.29}$$

are the scales used in the nonlinear Nash-Moser iteration. Let $H_\perp^s(\mathbb{T}^{\nu+1}) := H^s(\mathbb{T}^{\nu+1})\cap H_{\mathbb{S}^+}^\perp$ (we recall that the phase space contains only functions even in x, see (5.4)).

- Almost-invertibility assumption. There exists a subset $\Lambda_o\subset \mathtt{DC}(\gamma,\tau)\times[\mathtt{h}_1,\mathtt{h}_2]$ such that, for all $(\omega,\mathtt{h})\in\Lambda_o$ the operator $\mathcal{L}_\omega$ in (6.27) may be decomposed as

$$\mathcal{L}_\omega = \mathcal{L}_\omega^< + \mathcal{R}_\omega + \mathcal{R}_\omega^\perp \tag{6.30}$$

where $\mathcal{L}_\omega^<$ is invertible. More precisely, there exist constants $K_0, M, \sigma, \mu(\mathtt{b}), \mathtt{a}, p>0$ such that for any $s_0\leq s\leq S$, the operators $\mathcal{R}_\omega$, $\mathcal{R}_\omega^\perp$ satisfy the estimates

$$\|\mathcal{R}_\omega h\|_s^{k_0,\gamma} \lesssim_S \varepsilon\gamma^{-2(M+1)}N_{n-1}^{-\mathtt{a}}\big(\|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big), \tag{6.31}$$

$$\|\mathcal{R}_\omega^\perp h\|_{s_0}^{k_0,\gamma} \lesssim_S K_n^{-b}\big(\|h\|_{s_0+b+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s_0+\mu(\mathtt{b})+\sigma+b}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big), \qquad \forall b>0, \tag{6.32}$$

$$\|\mathcal{R}_\omega^\perp h\|_s^{k_0,\gamma} \lesssim_S \|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}. \tag{6.33}$$

Moreover, for every function $g\in H_\perp^{s+\sigma}(\mathbb{T}^{\nu+1},\mathbb{R}^2)$ and such that $g(-\varphi) = -\rho g(\varphi)$, for every $(\omega,\mathtt{h})\in\Lambda_o$, there is a solution $h := (\mathcal{L}_\omega^<)^{-1}g\in H_\perp^s(\mathbb{T}^{\nu+1},\mathbb{R}^2)$ such that $h(-\varphi)=\rho h(\varphi)$, of the linear equation $\mathcal{L}_\omega^< h = g$. The operator $(\mathcal{L}_\omega^<)^{-1}$ satisfies for all $s_0\leq s\leq S$ the tame estimate

$$\|(\mathcal{L}_\omega^<)^{-1}g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\big(\|g\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|g\|_{s_0+\sigma}^{k_0,\gamma}\big). \tag{6.34}$$

This assumption shall be verified at the n-th step of the Nash-Moser nonlinear iteration in Section 16 by applying Theorem 15.12. It is obtained by the process of almost-diagonalization of $\mathcal{L}_\omega$ up to a remainder $\mathcal{R}_\omega$ of size $O(N_{n-1}^{-\mathtt{a}})$ and an operator $\mathcal{R}_\omega^\perp$ which acts on high frequencies (it contains the projector $\Pi_{K_n}^\perp$).

In order to find an almost-approximate inverse of the linear operator in (6.24) (and so of $d_{i,\alpha}\mathcal{F}(i_\delta)$), it is sufficient to almost-invert the operator

$$\mathbb{D}[\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha}] := \begin{pmatrix} \omega\cdot\partial_\varphi\widehat{\phi} - \partial_\alpha K_{10}(\varphi)[\widehat{\alpha}] - K_{20}(\varphi)\widehat{y} - K_{11}^T(\varphi)\widehat{w} \\ \omega\cdot\partial_\varphi\widehat{y} + \partial_\phi\partial_\alpha K_{00}(\varphi)[\widehat{\alpha}] \\ (\mathcal{L}_\omega^<)\widehat{w} - J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}] - JK_{11}(\varphi)\widehat{y} \end{pmatrix}. \tag{6.35}$$

The operator $\mathbb{D}$ in (6.35) is obtained by neglecting in (6.24) 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 6.5), and the small remainders $\mathcal{R}_\omega$, $\mathcal{R}_\omega^\perp$ appearing in (6.30). We look for an exact inverse of $\mathbb{D}$ by solving the system

$$\mathbb{D}[\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha}] = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \tag{6.36}$$

where (g_1, g_2, g_3) satisfy the reversibility property

$$g_1(\varphi) = g_1(-\varphi)\,, \quad g_2(\varphi) = -g_2(-\varphi)\,, \quad g_3(\varphi) = -(\rho g_3)(-\varphi)\,. \tag{6.37}$$

We first consider the second equation in (6.36), namely $\omega\cdot\partial_\varphi\widehat{y} = g_2 - \partial_\alpha\partial_\phi K_{00}(\varphi)[\widehat{\alpha}]$. By reversibility, the φ-average of the right hand side of this equation is zero, and so its solution is

$$\widehat{y} := (\omega\cdot\partial_\varphi)^{-1}\big(g_2 - \partial_\alpha\partial_\phi K_{00}(\varphi)[\widehat{\alpha}]\big)\,. \tag{6.38}$$

Then we consider the third equation $(\mathcal{L}_\omega^<)\widehat{w} = g_3 + JK_{11}(\varphi)\widehat{y} + J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}]$, which, by the inversion assumption (6.34), has a solution

$$\widehat{w} := (\mathcal{L}_\omega^<)^{-1}\big(g_3 + JK_{11}(\varphi)\widehat{y} + J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}]\big)\,. \tag{6.39}$$

Finally, we solve the first equation in (6.36), which, substituting (6.38), (6.39), becomes

$$\omega\cdot\partial_\varphi\widehat{\phi} = g_1 + M_1(\varphi)[\widehat{\alpha}] + M_2(\varphi)g_2 + M_3(\varphi)g_3\,, \tag{6.40}$$

where

$$M_1(\varphi) := \partial_\alpha K_{10}(\varphi) - M_2(\varphi)\partial_\alpha\partial_\phi K_{00}(\varphi) + M_3(\varphi)J\partial_\alpha K_{01}(\varphi)\,, \tag{6.41}$$

$$M_2(\varphi) := K_{20}(\varphi)[\omega\cdot\partial_\varphi]^{-1} + K_{11}^T(\varphi)(\mathcal{L}_\omega^<)^{-1}JK_{11}(\varphi)[\omega\cdot\partial_\varphi]^{-1}\,, \quad M_3(\varphi) := K_{11}^T(\varphi)(\mathcal{L}_\omega^<)^{-1}\,. \tag{6.42}$$

In order to solve equation (6.40) we have to choose $\widehat{\alpha}$ such that the right hand side has zero average. By Lemma 6.6, (6.9), the φ-averaged matrix is $\langle M_1\rangle = \mathrm{Id} + O(\varepsilon\gamma^{-1})$. Therefore, for $\varepsilon\gamma^{-1}$ small enough, $\langle M_1\rangle$ is invertible and $\langle M_1\rangle^{-1} = \mathrm{Id} + O(\varepsilon\gamma^{-1})$. Thus we define

$$\widehat{\alpha} := -\langle M_1\rangle^{-1}(\langle g_1\rangle + \langle M_2 g_2\rangle + \langle M_3 g_3\rangle)\,. \tag{6.43}$$

With this choice of $\widehat{\alpha}$, equation (6.40) has the solution

$$\widehat{\phi} := (\omega\cdot\partial_\varphi)^{-1}\big(g_1 + M_1(\varphi)[\widehat{\alpha}] + M_2(\varphi)g_2 + M_3(\varphi)g_3\big)\,. \tag{6.44}$$

In conclusion, we have obtained a solution $(\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha})$ of the linear system (6.36).

Proposition 6.7. *Assume* (6.9) *(with $\mu = \mu(\mathtt{b}) + \sigma$) and* (6.34). *Then, for all $(\omega, \mathtt{h}) \in \Lambda_o$, for all $g := (g_1, g_2, g_3)$ even in x and satisfying* (6.37), *system* (6.36) *has a solution $\mathbb{D}^{-1}g := (\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha})$, where $(\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha})$ are defined in* (6.44), (6.38), (6.39), (6.43), *which satisfies* (5.14) *and for any $s_0 \le s \le S$*

$$\|\mathbb{D}^{-1}g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\big(\|g\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|g\|_{s_0+\sigma}^{k_0,\gamma}\big). \tag{6.45}$$

Proof. Recalling (6.42), by Lemma 6.6, (6.34), (6.9), we get $\|M_2 g\|_{s_0}^{k_0,\gamma} + \|M_3 g\|_{s_0}^{k_0,\gamma} \leq C\|g\|_{s_0+\sigma}^{k_0,\gamma}$. Then, by (6.43) and $\langle M_1\rangle^{-1} = \mathrm{Id} + O(\varepsilon\gamma^{-1}) = O(1)$, we deduce $|\widehat{\alpha}|^{k_0,\gamma} \leq C\|g\|_{s_0+\sigma}^{k_0,\gamma}$ and (6.38) implies

$$\|\widehat{y}\|_s^{k_0,\gamma} \lesssim_s \gamma^{-1}\big(\|g\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|g\|_{s_0}^{k_0,\gamma}\big)\,.$$

Bound (6.45) is sharp for $\widehat{w}$ because $(\mathcal{L}_\omega^{<})^{-1} g_3$ in (6.39) is estimated using (6.34). Finally also $\widehat{\phi}$ satisfies (6.45) using (6.44), (6.42), (6.34) and Lemma 6.6. ☐

Finally we prove that the operator

$$\mathbf{T}_0 := \mathbf{T}_0(i_0) := (D\widetilde{G}_\delta)(\varphi,0,0)\circ\mathbb{D}^{-1}\circ(DG_\delta)(\varphi,0,0)^{-1} \tag{6.46}$$

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 6.8. (Almost-approximate inverse) *Assume the inversion assumption* (6.30)-(6.34). *Then, there exists $\bar{\sigma} := \bar{\sigma}(\tau,\nu,k_0) > 0$ such that, if* (6.9) *holds with $\mu = \mu(\mathtt{b}) + \bar{\sigma}$, then for all $(\omega,\mathtt{h}) \in \Lambda_o$, for all $g := (g_1,g_2,g_3)$ even in x and satisfying* (6.37)*, the operator $\mathbf{T}_0$ defined in* (6.46) *satisfies, for all $s_0 \leq s \leq S$,*

$$\|\mathbf{T}_0 g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\big(\|g\|_{s+\bar{\sigma}}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar{\sigma}}^{k_0,\gamma}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\big)\,. \tag{6.47}$$

Moreover $\mathbf{T}_0$ is an almost-approximate inverse of $d_{i,\alpha}\mathcal{F}(i_0)$, namely

$$d_{i,\alpha}\mathcal{F}(i_0)\circ\mathbf{T}_0 - \mathrm{Id} = \mathcal{P}(i_0) + \mathcal{P}_\omega(i_0) + \mathcal{P}_\omega^\perp(i_0) \tag{6.48}$$

where, for all $s_0 \leq s \leq S$,

$$\begin{aligned}
\|\mathcal{P}g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\Big(&\|\mathcal{F}(i_0,\alpha_0)\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\|g\|_{s+\bar{\sigma}}^{k_0,\gamma}\\
&+\big\{\|\mathcal{F}(i_0,\alpha_0)\|_{s+\bar{\sigma}}^{k_0,\gamma} + \|\mathcal{F}(i_0,\alpha_0)\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar{\sigma}}^{k_0,\gamma}\big\}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\Big),
\end{aligned}\tag{6.49}$$

$$\|\mathcal{P}_\omega g\|_s^{k_0,\gamma} \lesssim_S \varepsilon\gamma^{-2M-3}N_{n-1}^{-\mathtt{a}}\big(\|g\|_{s+\bar{\sigma}}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar{\sigma}}^{k_0,\gamma}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\big)\,, \tag{6.50}$$

$$\|\mathcal{P}_\omega^\perp g\|_{s_0}^{k_0,\gamma} \lesssim_{S,b} \gamma^{-1}K_n^{-b}\big(\|g\|_{s_0+\bar{\sigma}+b}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s_0+\mu(\mathtt{b})+\bar{\sigma}+b}^{k_0,\gamma}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\big)\,, \quad \forall b > 0\,, \tag{6.51}$$

$$\|\mathcal{P}_\omega^\perp g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\big(\|g\|_{s+\bar{\sigma}}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar{\sigma}}^{k_0,\gamma}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\big)\,. \tag{6.52}$$

Proof. Bound (6.47) follows from (6.46), (6.45), (6.25). By (5.13), since $X_\mathcal{N}$ does not depend on I, and i_δ differs by i_0 only in the I component (see (6.12)), we have

$$\mathcal{E}_0 := d_{i,\alpha}\mathcal{F}(i_0) - d_{i,\alpha}\mathcal{F}(i_\delta) = \varepsilon\int_0^1 \partial_I d_i X_P(\theta_0, I_\delta + s(I_0 - I_\delta), z_0)[I_0 - I_\delta, \Pi[\,\cdot\,]]ds \tag{6.53}$$

where Π is the projection $(\widehat{\imath},\widehat{\alpha}) \mapsto \widehat{\imath}$. Denote by $\mathtt{u} := (\phi,y,w)$ the symplectic coordinates induced by G_δ in (6.17). Under the symplectic map G_δ, the nonlinear operator $\mathcal{F}$ in (5.13) is transformed into

$$\mathcal{F}(G_\delta(\mathtt{u}(\varphi)),\alpha) = DG_\delta(\mathtt{u}(\varphi))\big(\mathcal{D}_\omega\mathtt{u}(\varphi) - X_{K_\alpha}(\mathtt{u}(\varphi),\alpha)\big) \tag{6.54}$$

where $K_\alpha = H_\alpha\circ G_\delta$, see (6.18) and (6.20). Differentiating (6.54) at the trivial torus $\mathtt{u}_\delta(\varphi) = G_\delta^{-1}(i_\delta)(\varphi) = (\varphi,0,0)$, at $\alpha = \alpha_0$, we get

$$d_{i,\alpha}\mathcal{F}(i_\delta) = DG_\delta(\mathtt{u}_\delta)\big(\omega\cdot\partial_\varphi - d_{\mathtt{u},\alpha}X_{K_\alpha}(\mathtt{u}_\delta,\alpha_0)\big)D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1} + \mathcal{E}_1\,, \tag{6.55}$$

$$\mathcal{E}_1 := D^2G_\delta(\mathtt{u}_\delta)\big[DG_\delta(\mathtt{u}_\delta)^{-1}\mathcal{F}(i_\delta,\alpha_0),\ DG_\delta(\mathtt{u}_\delta)^{-1}\Pi[\,\cdot\,]\big] \tag{6.56}$$

In expanded form $\omega\cdot\partial_\varphi - d_{\mathtt{u},\alpha}X_{K_\alpha}(\mathtt{u}_\delta,\alpha_0)$ is provided by (6.24). By (6.35), (6.27), (6.30) and Lemma 6.5 we split

$$\omega\cdot\partial_\varphi - d_{\mathtt{u},\alpha}X_K(\mathtt{u}_\delta,\alpha_0) = \mathbb{D} + R_Z + \mathbb{R}_\omega + \mathbb{R}_\omega^\perp \tag{6.57}$$

where

$$R_Z[\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha}] := \begin{pmatrix} -\partial_\phi K_{10}(\varphi,\alpha_0)[\widehat{\phi}] \\ \partial_{\phi\phi}K_{00}(\varphi,\alpha_0)[\widehat{\phi}] + [\partial_\phi K_{10}(\varphi,\alpha_0)]^T\widehat{y} + [\partial_\phi K_{01}(\varphi,\alpha_0)]^T\widehat{w} \\ -J\{\partial_\phi K_{01}(\varphi,\alpha_0)[\widehat{\phi}]\} \end{pmatrix},$$

and

$$\mathbb{R}_\omega[\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha}] := \begin{pmatrix} 0 \\ 0 \\ \mathcal{R}_\omega[\widehat{w}] \end{pmatrix}, \qquad \mathbb{R}_\omega^\perp[\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha}] := \begin{pmatrix} 0 \\ 0 \\ \mathcal{R}_\omega^\perp[\widehat{w}] \end{pmatrix}.$$

By (6.53), (6.55), (6.56), (6.57) we get the decomposition

$$d_{i,\alpha}\mathcal{F}(i_0) = DG_\delta(\mathtt{u}_\delta) \circ \mathbb{D} \circ D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1} + \mathcal{E} + \mathcal{E}_\omega + \mathcal{E}_\omega^\perp \tag{6.58}$$

where

$$\mathcal{E} := \mathcal{E}_0 + \mathcal{E}_1 + DG_\delta(\mathtt{u}_\delta)R_Z D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}, \quad \mathcal{E}_\omega := DG_\delta(\mathtt{u}_\delta)\mathbb{R}_\omega D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}, \tag{6.59}$$

$$\mathcal{E}_\omega^\perp := DG_\delta(\mathtt{u}_\delta)\mathbb{R}_\omega^\perp D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}. \tag{6.60}$$

Applying $\mathbf{T}_0$ defined in (6.46) to the right hand side in (6.58) (recall that $\mathtt{u}_\delta(\varphi) := (\varphi, 0, 0)$), since $\mathbb{D}\circ\mathbb{D}^{-1} = \mathrm{Id}$ (Proposition 6.7), we get

$$d_{i,\alpha}\mathcal{F}(i_0) \circ \mathbf{T}_0 - \mathrm{Id} = \mathcal{P} + \mathcal{P}_\omega + \mathcal{P}_\omega^\perp,$$
$$\mathcal{P} := \mathcal{E} \circ \mathbf{T}_0, \quad \mathcal{P}_\omega := \mathcal{E}_\omega \circ \mathbf{T}_0, \quad \mathcal{P}_\omega^\perp := \mathcal{E}_\omega^\perp \circ \mathbf{T}_0.$$

By (6.9), (6.22), (6.13), (6.14), (6.15), (6.25)-(6.26) we get the estimate

$$\|\mathcal{E}[\widehat{\imath},\widehat{\alpha}]\|_s^{k_0,\gamma} \lesssim_s \|Z\|_{s_0+\sigma}^{k_0,\gamma}\|\widehat{\imath}\|_{s+\sigma}^{k_0,\gamma} + \|Z\|_{s+\sigma}^{k_0,\gamma}\|\widehat{\imath}\|_{s_0+\sigma}^{k_0,\gamma} + \|Z\|_{s_0+\sigma}^{k_0,\gamma}\|\widehat{\imath}\|_{s_0+\sigma}^{k_0,\gamma}\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}, \tag{6.61}$$

where $Z := \mathcal{F}(i_0,\alpha_0)$, recall (6.8). Then (6.49) follows from (6.47), (6.61), (6.9). Estimates (6.50), (6.51), (6.52) follow by (6.31)-(6.33), (6.47), (6.25), (6.13), (6.9). □

7 The linearized operator in the normal directions

In order to write an explicit expression of the linear operator $\mathcal{L}_\omega$ defined in (6.27) we have to express the operator $K_{02}(\phi)$ in terms of the original water waves Hamiltonian vector field.

Lemma 7.1. *The operator $K_{02}(\phi)$ is*

$$K_{02}(\phi) = \Pi_{\mathbb{S}^+}^\perp \partial_u \nabla_u H(T_\delta(\phi)) + \varepsilon R(\phi) \tag{7.1}$$

where H is the water waves Hamiltonian defined in (1.7) *(with gravity constant $g = 1$ and depth h replaced by $\mathtt{h}$), evaluated at the torus*

$$T_\delta(\phi) := \varepsilon A(i_\delta(\phi)) = \varepsilon A(\theta_0(\phi), I_\delta(\phi), z_0(\phi)) = \varepsilon v(\theta_0(\phi), I_\delta(\phi)) + \varepsilon z_0(\phi) \tag{7.2}$$

with $A(\theta, I, z)$, $v(\theta, I)$ defined in (5.9)*. The operator $K_{02}(\phi)$ is even and reversible. The remainder $R(\phi)$ has the "finite dimensional" form*

$$R(\phi)[h] = \sum\nolimits_{j\in\mathbb{S}^+} \big(h\,,\,g_j\big)_{L^2_x}\chi_j, \quad \forall h \in H_{\mathbb{S}^+}^\perp, \tag{7.3}$$

for functions $g_j, \chi_j \in H_{\mathbb{S}^+}^\perp$ which satisfy the tame estimates: for some $\sigma := \sigma(\tau,\nu) > 0$, $\forall s \geq s_0$,

$$\|g_j\|_s^{k_0,\gamma} + \|\chi_j\|_s^{k_0,\gamma} \lesssim_s 1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{k_0,\gamma}, \quad \|\partial_i g_j[\widehat{\imath}]\|_s + \|\partial_i\chi_j[\widehat{\imath}]\|_s \lesssim_s \|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma}. \tag{7.4}$$

Proof. The lemma follows as in Lemma 6.1 in [21]. □

By Lemma 7.1 the linear operator $\mathcal{L}_\omega$ defined in (6.27) has the form

$$\mathcal{L}_\omega = \Pi_{\mathbb{S}^+}^\perp(\mathcal{L} + \varepsilon R)_{|H_{\mathbb{S}^+}^\perp} \qquad \text{where} \qquad \mathcal{L} := \omega\cdot\partial_\varphi - J\partial_u \nabla_u H(T_\delta(\varphi)) \tag{7.5}$$

is obtained linearizing the original water waves system (1.14), (1.6) at the torus $u = (\eta, \psi) = T_\delta(\varphi)$ defined in (7.2), changing $\partial_t \rightsquigarrow \omega\cdot\partial_\varphi$. The function $\eta(\varphi, x)$ is even(φ)even(x) and $\psi(\varphi, x)$ is odd(φ)even(x).

Using formula (3.1), the linearized operator of (1.14) is represented by the 2×2 operator matrix

$$\mathcal{L} := \omega\cdot\partial_\varphi + \begin{pmatrix} \partial_x V + G(\eta)B & -G(\eta) \\ (1 + BV_x) + BG(\eta)B & V\partial_x - BG(\eta) \end{pmatrix} \tag{7.6}$$

where B, V are defined in (3.2). The function B is odd(φ)even(x) and V is odd(φ)odd(x). The operator $\mathcal{L}$ acts on $H^1(\mathbb{T})\times H^1(\mathbb{T})$.

The operators $\mathcal{L}_\omega$ and $\mathcal{L}$ are real, even and reversible.

We are going to make several transformations, whose aim is to conjugate the linearized operator to a constant coefficients operator, up to a remainder that is small in size and regularizing at a conveniently high order. It is convenient to ignore all projections at first, and consider the linearized operator as an operator on the whole of $H^1(\mathbb{T})\times H^1(\mathbb{T})$. At the end of the conjugation procedure, we shall restrict ourselves to the phase space $H_0^1(\mathbb{T})\times \dot H^1(\mathbb{T})$ and perform the projection on the normal directions $H_{\mathbb{S}^+}^\perp$. The finite dimensional remainder εR transforms under conjugation into an operator of the same form and therefore it will be dealt with only once at the end of Section 14.

For the sequel we will always assume the following ansatz (that will be satisfied by the approximate solutions obtained along the nonlinear Nash-Moser iteration of Section 16): for some constant $\mu_0 := \mu_0(\tau, \nu) > 0$, $\gamma \in (0, 1)$,

$$\|\mathfrak{I}_0\|_{s_0+\mu_0}^{k_0,\gamma} \leq 1 \ , \qquad \text{and so, by (6.13),} \ \|\mathfrak{I}_\delta\|_{s_0+\mu_0}^{k_0,\gamma} \leq 2\,. \tag{7.7}$$

In order to estimate the variation of the eigenvalues with respect to the approximate invariant torus, we need also to estimate the derivatives (or the variation) with respect to the torus $i(\varphi)$ in another low norm $\|\ \|_{s_1}$, for all the Sobolev indices s_1 such that

$$s_1 + \sigma_0 \leq s_0 + \mu_0\,, \quad \text{for some} \quad \sigma_0 := \sigma_0(\tau, \nu) > 0\,. \tag{7.8}$$

Thus by (7.7) we have

$$\|\mathfrak{I}_0\|_{s_1+\sigma_0}^{k_0,\gamma} \leq 1 \qquad \text{and so, by (6.13),} \ \|\mathfrak{I}_\delta\|_{s_1+\sigma_0}^{k_0,\gamma} \leq 2\,. \tag{7.9}$$

The constants μ_0 and σ_0 represent the *loss of derivatives* accumulated along the reduction procedure of Sections 8-13. What is important is that they are independent of the Sobolev index s. Along Sections 7-13, we shall denote by $\sigma := \sigma(k_0, \tau, \nu) > 0$ a constant (which possibly increases from lemma to lemma) representing the loss of derivatives along the finitely many steps of the reduction procedure.

As a consequence of Moser composition Lemma 2.6, the Sobolev norm of the function $u = T_\delta$ defined in (7.2) satisfies, $\forall s \geq s_0$,

$$\|u\|_s^{k_0,\gamma} = \|\eta\|_s^{k_0,\gamma} + \|\psi\|_s^{k_0,\gamma} \leq \varepsilon C(s)\big(1 + \|\mathfrak{I}_0\|_s^{k_0,\gamma}\big) \tag{7.10}$$

(the function A defined in (5.9) is smooth). Similarly

$$\|\partial_i u[\hat\imath]\|_{s_1} \lesssim_{s_1} \varepsilon\|\hat\imath\|_{s_1}, \qquad \|\Delta_{12} u\|_{s_1} \lesssim_{s_1} \varepsilon\|i_2 - i_1\|_{s_1} \tag{7.11}$$

where we denote $\Delta_{12} u := u(i_2) - u(i_1)$; we will systematically use this notation.

In the next sections we shall also assume that, for some $\kappa := \kappa(\tau, \nu) > 0$, we have

$$\varepsilon\gamma^{-\kappa} \leq \delta(S)\,,$$

where $\delta(S) > 0$ is a constant small enough and S will be fixed in (16.12). We recall that $\mathfrak{I}_0 := \mathfrak{I}_0(\omega, \mathtt{h})$ is defined for all $(\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ by the extension procedure that we perform along the Nash-Moser nonlinear iteration. Moreover all the functions appearing in $\mathcal{L}$ in (7.6) 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 2.3.

7.1 Linearized good unknown of Alinhac

Following [1], [21] we conjugate the linearized operator $\mathcal{L}$ in (7.6) by the multiplication operator

$$\mathcal{Z} := \begin{pmatrix} 1 & 0 \\ B & 1 \end{pmatrix}, \quad \mathcal{Z}^{-1} = \begin{pmatrix} 1 & 0 \\ -B & 1 \end{pmatrix}, \tag{7.12}$$

where $B = B(\varphi, x)$ is the function defined in (3.2), obtaining

$$\mathcal{L}_0 := \mathcal{Z}^{-1} \mathcal{L} \mathcal{Z} = \omega \cdot \partial_\varphi + \begin{pmatrix} \partial_x V & -G(\eta) \\ a & V \partial_x \end{pmatrix} \tag{7.13}$$

where a is the function

$$a := a(\varphi, x) := 1 + (\omega \cdot \partial_\varphi B) + V B_x \,. \tag{7.14}$$

All a, B, V are real valued periodic functions of (φ, x) — variable coefficients — and satisfy

$$B = \mathrm{odd}(\varphi)\mathrm{even}(x), \quad V = \mathrm{odd}(\varphi)\mathrm{odd}(x), \quad a = \mathrm{even}(\varphi)\mathrm{even}(x) \,.$$

The matrix $\mathcal{Z}$ in (7.12) amounts to introduce, as in Lannes [45]-[46], a linearized version of the *good unknown of Alinhac*, working with the variables (η, ς) with $\varsigma := \psi - B\eta$, instead of (η, ψ).

Lemma 7.2. *The maps $\mathcal{Z}^{\pm 1} - \mathrm{Id}$ are even, reversibility preserving and $\mathcal{D}^{k_0}$-tame with tame constant satisfying, for all $s \geq s_0$,*

$$\mathfrak{M}_{\mathcal{Z}^{\pm 1} - \mathrm{Id}}(s) \,, \mathfrak{M}_{(\mathcal{Z}^{\pm 1} - \mathrm{Id})^*}(s) \lesssim_s \varepsilon \big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) \,. \tag{7.15}$$

The operator $\mathcal{L}_0$ is even and reversible. There is $\sigma := \sigma(\tau, \nu) > 0$ such that the functions

$$\|a - 1\|_s^{k_0,\gamma} + \|V\|_s^{k_0,\gamma} + \|B\|_s^{k_0,\gamma} \lesssim_s \varepsilon \big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) \,. \tag{7.16}$$

Moreover

$$\|\partial_i a[\hat\imath]\|_{s_1} + \|\partial_i V[\hat\imath]\|_{s_1} + \|\partial_i B[\hat\imath]\|_{s_1} \lesssim_{s_1} \varepsilon \|\hat\imath\|_{s_1+\sigma} \tag{7.17}$$

$$\|\partial_i(\mathcal{Z}^{\pm 1}[\hat\imath])h\|_{s_1} \,, \|\partial_i((\mathcal{Z}^{\pm 1})^*[\hat\imath])h\|_{s_1} \lesssim_{s_1} \varepsilon \|\hat\imath\|_{s_1+\sigma} \|h\|_{s_1} \,. \tag{7.18}$$

Proof. The proof is the same as the one of Lemma 6.3 in [21]. □

We expand $\mathcal{L}_0$ in (7.13) as

$$\mathcal{L}_0 = \omega \cdot \partial_\varphi + \begin{pmatrix} V\partial_x & 0 \\ 0 & V\partial_x \end{pmatrix} + \begin{pmatrix} V_x & -G(\eta) \\ a & 0 \end{pmatrix} . \tag{7.19}$$

In the next section we deal with the first order operator $\omega \cdot \partial_\varphi + V\partial_x$.

8 Straightening the first order vector field

The aim of this section is to conjugate the variable coefficients operator $\omega \cdot \partial_\varphi + V(\varphi, x)\partial_x$ to the constant coefficients vector field $\omega \cdot \partial_\varphi$, namely to find a change of variable $\mathcal{B}$ such that

$$\mathcal{B}^{-1}\big(\omega \cdot \partial_\varphi + V(\varphi, x)\partial_x\big)\mathcal{B} = \omega \cdot \partial_\varphi \,. \tag{8.1}$$

Quasi-periodic transport equation. We consider a φ-dependent family of diffeomorphisms of $\mathbb{T}_x$ of the space variable

$$y = x + \beta(\varphi, x)$$

where the function $\beta : \mathbb{T}_\varphi^\nu \times \mathbb{T}_x \to \mathbb{R}$ is odd in x, even in φ, and $|\beta_x(\varphi, x)| < 1/2$ for all $(\varphi, x) \in \mathbb{T}^{\nu+1}$. We denote by $\mathcal{B}$ the corresponding composition operator, namely

$$\mathcal{B} : h \mapsto \mathcal{B}h, \quad (\mathcal{B}h)(\varphi, x) := h(\varphi, x + \beta(\varphi, x)) \,. \tag{8.2}$$

Let us compute the conjugated operator in the left hand side in (8.1). The conjugate $\mathcal{B}^{-1} f \mathcal{B}$ of a multiplication operator $f : u \mapsto f(\varphi, x)u$ is the multiplication operator $(\mathcal{B}^{-1} f) : u \mapsto (\mathcal{B}^{-1} f)(\varphi, y)u$. The conjugate of the differential operators ∂_x and $\omega \cdot \partial_\varphi$ by the change of variable $\mathcal{B}$ are

$$\mathcal{B}^{-1} \partial_x \mathcal{B} = \big(1 + \mathcal{B}^{-1} \beta_x\big)\, \partial_y, \qquad \mathcal{B}^{-1}\, \omega \cdot \partial_\varphi \, \mathcal{B} = \omega \cdot \partial_\varphi + (\mathcal{B}^{-1} \omega \cdot \partial_\varphi \beta)\, \partial_y .$$

Therefore $\omega \cdot \partial_\varphi + V(\varphi, x)\partial_x$ is transformed into

$$\mathcal{B}^{-1}\big(\omega \cdot \partial_\varphi + V(\varphi, x)\partial_x\big)\mathcal{B} = \omega \cdot \partial_\varphi + c(\varphi, y)\, \partial_y \tag{8.3}$$

where $c(\varphi, y)$ is the periodic function

$$c(\varphi, y) = \mathcal{B}^{-1}\big(\omega \cdot \partial_\varphi \beta + V(1 + \beta_x)\big)(\varphi, y)\,. \tag{8.4}$$

In view of (8.3)-(8.4) we obtain (8.1) if $\beta(\varphi, x)$ solves the equation

$$\omega \cdot \partial_\varphi \beta(\varphi, x) + V(\varphi, x)(1 + \beta_x(\varphi, x)) = 0\,, \tag{8.5}$$

which can be interpreted as a *quasi-periodic transport* equation.

Quasi-periodic characteristic equation. Instead of solving directly (8.5) we solve the equation satisfied by the inverse diffeomorphism

$$x + \beta(\varphi, x) = y \qquad \Longleftrightarrow \qquad x = y + \breve\beta(\varphi, y)\,, \qquad \forall x, y \in \mathbb{R},\ \varphi \in \mathbb{T}^\nu\,. \tag{8.6}$$

It turns out that equation (8.5) for $\beta(\varphi, x)$ is equivalent to the following equation for $\breve\beta(\varphi, y)$:

$$\omega \cdot \partial_\varphi \breve\beta(\varphi, y) = V(\varphi, y + \breve\beta(\varphi, y)) \tag{8.7}$$

which is a quasi-periodic version of the *characteristic* equation $\dot x = V(\omega t, x)$.

Remark 8.1. *We can give a geometric interpretation of equation* (8.7) *in terms of conjugation of vector fields on the torus* $\mathbb{T}^\nu \times \mathbb{T}$. *Under the diffeomorphism of* $\mathbb{T}^\nu \times \mathbb{T}$ *defined by*

$$\begin{pmatrix} \varphi \\ x \end{pmatrix} = \begin{pmatrix} \psi \\ y + \breve\beta(\psi, y) \end{pmatrix}, \quad \textit{the system} \quad \frac{d}{dt}\begin{pmatrix} \varphi \\ x \end{pmatrix} = \begin{pmatrix} \omega \\ V(\varphi, x) \end{pmatrix}$$

transforms into

$$\frac{d}{dt}\begin{pmatrix} \psi \\ y \end{pmatrix} = \begin{pmatrix} \omega \\ \big\{ - \omega \cdot \partial_\varphi \breve\beta(\psi, y) + V(\varphi, y + \breve\beta(\psi, y))\big\}\big(1 + \breve\beta_y(\psi, y)\big)^{-1} \end{pmatrix}.$$

The vector field in the new coordinates reduces to $(\omega, 0)$ *if and only if* (8.7) *holds. In the new variables the solutions are simply given by* $y(t) = c$, $c \in \mathbb{R}$, *and all the solutions of the scalar quasi-periodically forced differential equation* $\dot x = V(\omega t, x)$ *are time quasi-periodic of the form* $x(t) = c + \breve\beta(\omega t, c)$.

In the rest of the section we solve equation (8.7), for $V(\varphi, x)$ small, and for ω in the set of Diophantine vectors $\mathtt{DC}(\gamma, \tau)$ defined in (2.13), by applying the Nash-Moser-Hörmander implicit function theorem in Appendix B.

We rename $\breve\beta \to u$, $y \to x$, and write equation (8.7) as

$$F(u)(\varphi, x) := \omega \cdot \partial_\varphi u(\varphi, x) - V(\varphi, x + u(\varphi, x)) = 0\,. \tag{8.8}$$

The linearized operator at a given function $u(\varphi, x)$ is

$$F'(u)h := \omega \cdot \partial_\varphi h - q(\varphi, x)h, \quad q(\varphi, x) := V_x(\varphi, x + u(\varphi, x))\,. \tag{8.9}$$

In the next lemma we solve the linear problem $F'(u)h = f$.

Lemma 8.2. (Linearized quasi-periodic characteristic equation) *Let $\varsigma := 3k_0 + 2\tau(k_0+1) + 2 = 2\mu + k_0 + 2$, where μ is the loss in (2.18) (with $k+1 = k_0$), and let $\omega \in \mathtt{DC}(2\gamma,\tau)$. Assume that the periodic function u is even(φ)odd(x), that V is odd(φ)odd(x), and*

$$\|u\|^{k_0,\gamma}_{s_0+\varsigma} + \gamma^{-1}\|V\|^{k_0,\gamma}_{s_0+\varsigma} \le \delta_0 \tag{8.10}$$

with δ_0 small enough. Then, given a periodic function f which is odd(φ)odd(x), the linearized equation

$$F'(u)h = f \tag{8.11}$$

has a unique periodic solution $h(\varphi,x)$ which is even(φ)odd(x) having zero average in φ, i.e.

$$\langle h\rangle_\varphi(x) := \frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} h(\varphi,x)\,d\varphi = 0 \quad \forall x\in\mathbb{T}. \tag{8.12}$$

This defines a right inverse of the linearized operator $F'(u)$, which we denote by $h = F'(u)^{-1}f$. The right inverse $F'(u)^{-1}$ satisfies

$$\|F'(u)^{-1}f\|^{k_0,\gamma}_s \lesssim_s \gamma^{-1}\big(\|f\|^{k_0,\gamma}_{s+\varsigma} + \gamma^{-1}(\|V\|^{k_0,\gamma}_{s+\varsigma} + \|u\|^{k_0,\gamma}_{s+\varsigma}\|V\|^{k_0,\gamma}_{s_0+\varsigma})\|f\|^{k_0,\gamma}_{s_0}\big) \tag{8.13}$$

for all $s \ge s_0$, where $\|\cdot\|^{k_0,\gamma}_s$ denotes the norm of $\mathrm{Lip}(k_0,\mathtt{DC}(2\gamma,\tau),s,\gamma)$.

Proof. Given f, we have to solve the linear equation $\omega\cdot\partial_\varphi h - qh = f$, where q is the function defined in (8.9). From the parity of u, V it follows that q is odd(φ)even(x). By variation of constants, we look for solutions of the form $h = we^v$, and we find (recalling (2.14))

$$v := (\omega\cdot\partial_\varphi)^{-1}q, \quad w := w_0 + g, \quad w_0 := (\omega\cdot\partial_\varphi)^{-1}(e^{-v}f), \quad g = g(x) := -\frac{\langle w_0 e^v\rangle_\varphi}{\langle e^v\rangle_\varphi}.$$

This choice of g, and hence of w, is the only one matching the zero average requirement (8.12); this gives uniqueness of the solution. Moreover

$$v = \mathrm{even}(\varphi)\mathrm{even}(x), \quad w_0 = \mathrm{even}(\varphi)\mathrm{odd}(x), \quad g(x) = \mathrm{odd}(x),$$

whence h is even(φ)odd(x). Using (2.10), (2.11), (2.18), (2.19), (8.10), and (2.9) one has

$$\begin{aligned}
\|v\|^{k_0,\gamma}_s &\lesssim_s \gamma^{-1}\|q\|^{k_0,\gamma}_{s+\mu} \lesssim_s \gamma^{-1}\big(\|V\|^{k_0,\gamma}_{s+\mu+k_0+1} + \|u\|^{k_0,\gamma}_{s+\mu}\|V\|^{k_0,\gamma}_{s_0+k_0+2}\big),\\
\|w\|^{k_0,\gamma}_s &\lesssim_s \gamma^{-1}\big(\|f\|^{k_0,\gamma}_{s+\mu} + \|v\|^{k_0,\gamma}_{s+\mu}\|f\|^{k_0,\gamma}_{s_0}\big).
\end{aligned}$$

Using again (2.10), (2.19), (8.10), and (2.9), the proof of (8.13) is complete. ☐

We now prove the existence of a solution of equation (8.8) by means of the Nash-Moser-Hörmander theorem proved in [10], whose statement is given in Appendix B. The main advantage of using such a result consists in providing estimate (8.16) of the high norm of the solution u in terms of the high norm of V with a fixed loss of regularity p.

Theorem 8.3. (Solution of the quasi-periodic characteristic equation (8.8)**)** *Let ς be the constant defined in Lemma 8.2, and let $s_2 := 2s_0 + 3\varsigma + 1$, $p := 3\varsigma + 2$. Assume that V is odd(φ)odd(x). There exist $\delta\in(0,1), C>0$ depending on ς, s_0 such that, for all $\omega\in\mathtt{DC}(2\gamma,\tau)$, if $V \in \mathrm{Lip}(k_0,\mathtt{DC}(2\gamma,\tau),s_2+p,\gamma)$ satisfies*

$$\gamma^{-1}\|V\|^{k_0,\gamma}_{s_2+p} \le \delta, \tag{8.14}$$

then there exists a solution $u \in \mathrm{Lip}(k_0,\mathtt{DC}(2\gamma,\tau),s_2,\gamma)$ of $F(u) = 0$. The solution u is even(φ)odd(x), it has zero average in φ, and satisfies

$$\|u\|^{k_0,\gamma}_{s_2} \le C\gamma^{-1}\|V\|^{k_0,\gamma}_{s_2+p}. \tag{8.15}$$

If, in addition, $V \in \mathrm{Lip}(k_0,\mathtt{DC}(2\gamma,\tau),s+p,\gamma)$ for $s > s_2$, then $u\in\mathrm{Lip}(k_0,\mathtt{DC}(2\gamma,\tau),s,\gamma)$, with

$$\|u\|^{k_0,\gamma}_s \le C_s\gamma^{-1}\|V\|^{k_0,\gamma}_{s+p} \tag{8.16}$$

for some constant C_s depending on s,ς,s_0, independent of V,γ.

Proof. We apply Theorem B.1 of Appendix B. For $a, b \geq 0$, we define

$$E_a := \big\{u \in \mathrm{Lip}(k_0, \mathtt{DC}(2\gamma,\tau), 2s_0 + a, \gamma) : u = \mathrm{even}(\varphi)\mathrm{odd}(x),\ \langle u\rangle_\varphi(x) = 0\big\}, \quad \|u\|_{E_a} := \|u\|^{k_0,\gamma}_{2s_0+a}, \quad (8.17)$$

$$F_b := \big\{g \in \mathrm{Lip}(k_0, \mathtt{DC}(2\gamma,\tau), 2s_0 + b, \gamma) : g = \mathrm{odd}(\varphi)\mathrm{odd}(x)\big\}, \qquad \|g\|_{F_b} := \|g\|^{k_0,\gamma}_{2s_0+b} \quad (8.18)$$

(s_0 is in the last term of (8.13), while $2s_0$ appears in the composition estimate (2.11)). We consider Fourier truncations at powers of 2 as smoothing operators, namely

$$S_n \ :\ u(\varphi,x) = \sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} u_{\ell j} e^{\mathrm{i}(\ell\cdot\varphi + jx)} \quad \mapsto \quad (S_n u)(\varphi,x) := \sum_{\langle \ell,j\rangle \leq 2^n} u_{\ell j} e^{\mathrm{i}(\ell\cdot\varphi + jx)} \tag{8.19}$$

on both spaces E_a and F_b. Hence both E_a and F_b satisfy (B.1)-(B.5), and the operators R_n defined in (B.6) give the dyadic decomposition $2^n < \langle \ell, j\rangle \leq 2^{n+1}$. Since S_n in (8.19) are "crude" Fourier truncations, (B.8) holds with "=" instead of "≤" and $C = 1$. As a consequence, every $g \in F_\beta$ satisfies the first inequality in (B.12) with $A = 1$ (it becomes, in fact, an equality), and, similarly, if $g \in F_{\beta+c}$ then (B.15) holds with $A_c = 1$ (and "=").

We denote by $\mathcal{V}$ the composition operator $\mathcal{V}(u)(\varphi,x) := V(\varphi, x + u(\varphi,x))$, and define $\Phi(u) := \omega\cdot\partial_\varphi u - \mathcal{V}(u)$, namely we take the nonlinear operator F in (8.8) as the operator Φ of Theorem B.1. By Lemma 2.4, if $\|u\|^{k_0,\gamma}_{2s_0+1} \leq \delta_{2.4}$ (where we denote by $\delta_{2.4}$ the constant δ of Lemma 2.4), then $\mathcal{V}(u)$ satisfies (2.11), namely for all $s \geq s_0$

$$\|\mathcal{V}(u)\|^{k_0,\gamma}_s \lesssim_s \|V\|^{k_0,\gamma}_{s+k_0} + \|u\|^{k_0,\gamma}_s \|V\|^{k_0,\gamma}_{s_0+k_0+1}, \tag{8.20}$$

and its second derivative $\mathcal{V}''(u)[v,w] = V_{xx}(\varphi, x + u(\varphi,x))vw$ satisfies

$$\begin{aligned}\|\mathcal{V}''(u)[v,w]\|^{k_0,\gamma}_s \lesssim_s \|V\|^{k_0,\gamma}_{s_0+k_0+3}\Big(\|v\|^{k_0,\gamma}_s\|w\|^{k_0,\gamma}_{s_0} + \|v\|^{k_0,\gamma}_{s_0}\|w\|^{k_0,\gamma}_s\Big)&\\ + \big\{\|V\|^{k_0,\gamma}_{s_0+k_0+3}\|u\|^{k_0,\gamma}_s + \|V\|^{k_0,\gamma}_{s+k_0+2}\big\}\|v\|^{k_0,\gamma}_{s_0}\|w\|^{k_0,\gamma}_{s_0}.&\end{aligned} \tag{8.21}$$

We fix μ, U of Theorem B.1 as $\mu := 1$, $U := \{u \in E_1 : \|u\|_{E_1} \leq \delta_{2.4}\}$. Thus Φ maps $U \to F_0$ and $U \cap E_{a+\mu} \to F_a$ for all $a \in [0, a_2 - 1]$, provided that $\|V\|^{k_0,\gamma}_{2s_0+a_2-1+k_0} < \infty$ (a_2 will be fixed below in (8.25)). Moreover, for all $a \in [0, a_2-1]$, Φ is of class $C^2(U \cap E_{a+\mu}, F_a)$ and it satisfies (B.10) with $a_0 := 0$,

$$M_1(a) := C(a)\|V\|^{k_0,\gamma}_{s_0+k_0+3}, \quad M_2(a) := M_1(a), \quad M_3(a) := C(a)\|V\|^{k_0,\gamma}_{2s_0+k_0+2+a}. \tag{8.22}$$

We fix a_1, δ_1 of Theorem B.1 as $a_1 := \varsigma$, where $\varsigma = 3k_0 + 2\tau(k_0+1) + 2$ is the constant appearing in Lemma 8.2, and $\delta_1 := \frac{1}{2}\delta_{8.2}$, where $\delta_{8.2}$ is the constant δ_0 of Lemma 8.2. If $\gamma^{-1}\|V\|^{k_0,\gamma}_{s_0+\varsigma} \leq \delta_1$ and $\|v\|_{E_{a_1}} \leq \delta_1$, then, by Lemma 8.2, the right inverse $\Psi(v) := F'(v)^{-1}$ is well defined, and it satisfies

$$\|\Psi(v)g\|_{E_a} \leq L_1(a)\|g\|_{F_{a+\varsigma}} + (L_2(a)\|v\|_{E_{a+\varsigma}} + L_3(a))\|g\|_{F_0} \tag{8.23}$$

where

$$L_1(a) := C(a)\gamma^{-1}, \quad L_2(a) := C(a)\gamma^{-2}\|V\|^{k_0,\gamma}_{s_0+\varsigma}, \quad L_3(a) := C(a)\gamma^{-2}\|V\|^{k_0,\gamma}_{2s_0+a+\varsigma}. \tag{8.24}$$

We fix α, β, a_2 of Theorem B.1 as

$$\beta := 4\varsigma + 1, \quad \alpha := 3\varsigma + 1, \quad a_2 := 5\varsigma + 3, \tag{8.25}$$

so that (B.9) is satisfied. Bound (8.23) implies (B.11) for all $a \in [a_1, a_2]$ provided that $\|V\|^{k_0,\gamma}_{2s_0+a_2+\varsigma} < \infty$.

All the hypotheses of the first part of Theorem B.1 are satisfied. As a consequence, there exists a constant $\delta_{B.14}$ (given by (B.14) with $A = 1$) such that, if $\|g\|_{F_\beta} \leq \delta_{B.14}$, then the equation $\Phi(u) = \Phi(0) + g$ has a solution $u \in E_\alpha$, with bound (B.13). In particular, the result applies to $g = V$, in which case the equation $\Phi(u) = \Phi(0) + g$ becomes $\Phi(u) = 0$. We have to verify the smallness condition $\|g\|_{F_\beta} \leq \delta_{B.14}$. Using (8.22), (8.24), (8.14), we verify that $\delta_{B.14} \geq C\gamma$. Thus, the smallness condition $\|g\|_{F_\beta} \leq \delta_{B.14}$ is satisfied if $\|V\|^{k_0,\gamma}_{2s_0+a_2+\varsigma}\gamma^{-1}$ is smaller than some δ depending on ς, s_0. This is assumption (8.14), since $2s_0 + a_2 + \varsigma = s_2 + p$. Then (B.13), recalling (8.25), gives $\|u\|^{k_0,\gamma}_{s_2} \leq C\gamma^{-1}\|V\|^{k_0,\gamma}_{s_2+\varsigma}$, which implies (8.15) since $p \geq \varsigma$.

We finally prove estimate (8.16). Let $c > 0$. If, in addition, $\|V\|^{k_0,\gamma}_{2s_0+a_2+c+\varsigma} < \infty$, then all the assumptions of the second part of Theorem B.1 are satisfied. By (8.22), (8.24) and (8.14), we estimate the constants defined in (B.17)-(B.18) as

$$\mathcal{G}_1 \le C_c \gamma^{-2} \|V\|^{k_0,\gamma}_{2s_0+a_2+c+\varsigma}, \quad \mathcal{G}_2 \le C_c \gamma^{-1}, \quad z \le C_c$$

for some constant C_c depending on c. Bound (B.16) implies (8.16) with $s = s_2 + c$ (the highest norm of V in (8.16) does not come from the term $\|V\|_{F_{\beta+c}}$ of (B.16), but from the factor $\mathcal{G}_1$). The proof is complete. $\square$

The next lemma deals with the dependence of the solution u of (8.8) on V (actually it would be enough to estimate this Lipschitz dependence only in the "low" norm s_1 introduced in (7.8)).

Lemma 8.4 (Lipschitz dependence of u on V)**.** *Let ς, s_2, p be as defined in Theorem 8.3. Let V_1, V_2 satisfy* (8.14)*, and let u_1, u_2 be the solutions of*

$$\omega \cdot \partial_\varphi u_i - V_i(\varphi, x + u_i(\varphi, x)) = 0, \quad i = 1, 2,$$

given by Theorem 8.3. Then for all $s \ge s_2 - \mu$ (where μ is the constant defined in (2.18)*)*

$$\|u_1 - u_2\|^{k_0,\gamma}_s \lesssim_s \gamma^{-1} \|V_1 - V_2\|^{k_0,\gamma}_{s+\mu+k_0} + \gamma^{-2} \max_{i=1,2} \|V_i\|^{k_0,\gamma}_{s+2\mu+p} \|V_1 - V_2\|^{k_0,\gamma}_{s_2+k_0}. \tag{8.26}$$

Proof. The difference $h := u_1 - u_2$ is even(φ)odd(x), it has zero average in φ and it solves $\omega \cdot \partial_\varphi h - ah = b$, where

$$a(\varphi, x) := \int_0^1 (\partial_x V_1)(\varphi, x + tu_1 + (1-t)u_2)\, dt, \quad b(\varphi, x) := (V_1 - V_2)(\varphi, x + u_2).$$

The function a is odd(φ)even(x) and b is odd(φ)odd(x). Then, by variation of constants and uniqueness, $h = we^v$, where (as in Lemma 8.2)

$$v := (\omega \cdot \partial_\varphi)^{-1} a, \quad w := w_0 + g, \quad w_0 := (\omega \cdot \partial_\varphi)^{-1}(e^{-v} b), \quad g = g(x) := -\frac{\langle w_0 e^v \rangle_\varphi}{\langle e^v \rangle_\varphi}.$$

By (2.11), (8.14), (8.15), (8.16), one has

$$\|a\|^{k_0,\gamma}_s \lesssim_s \|V_{1,2}\|^{k_0,\gamma}_{s+p}, \quad \|b\|^{k_0,\gamma}_s \lesssim_s \|V_1 - V_2\|^{k_0,\gamma}_{s+k_0} + \gamma^{-1} \|V_2\|^{k_0,\gamma}_{s+p} \|V_1 - V_2\|^{k_0,\gamma}_{s_0+k_0+1} \quad \forall s \ge s_2,$$

where $\|V_{1,2}\|^{k_0,\gamma}_s := \max_{i=1,2} \|V_i\|^{k_0,\gamma}_s$, and, like in Theorem 8.3, $s_2 := 2s_0 + 3\varsigma + 1$, $p := 3\varsigma + 2$. By (2.18) and (2.19),

$$\|v\|^{k_0,\gamma}_s \lesssim_s \gamma^{-1} \|V_{1,2}\|^{k_0,\gamma}_{s+\mu+p}, \quad \|e^v\|^{k_0,\gamma}_s \lesssim_s 1 + \gamma^{-1} \|V_{1,2}\|^{k_0,\gamma}_{s+\mu+p} \quad \forall s \ge s_2 - \mu,$$

where μ is defined in (2.18). Then

$$\|w_0\|^{k_0,\gamma}_s \lesssim_s \gamma^{-1} \|V_1 - V_2\|^{k_0,\gamma}_{s+\mu+k_0} + \gamma^{-2} \|V_{1,2}\|^{k_0,\gamma}_{s+2\mu+p} \|V_1 - V_2\|^{k_0,\gamma}_{s_2+k_0}$$

for all $s \ge s_2 - \mu$, and $w_0 e^v, g, h$ satisfy the same bound. $\square$

In Theorem 8.3, for any $\lambda = (\omega, \mathtt{h}) \in \mathtt{DC}(2\gamma, \tau) \times [\mathtt{h}_1, \mathtt{h}_2]$ we have constructed a periodic function $u = \breve{\beta}$ that solves (8.8), namely the quasi-periodic characteristic equation (8.7), so that the periodic function β, defined by the inverse diffeomorphism in (8.6), solves the quasi-periodic transport equation (8.5).

By Theorem A.2 we define an extension $\mathcal{E}_k(u) = \mathcal{E}_k(\breve{\beta}) =: \breve{\beta}_{ext}$ (with $k + 1 = k_0$) to the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$. By the linearity of the extension operator $\mathcal{E}_k$ and by the norm equivalence (A.6), the difference of the extended functions $\mathcal{E}_k(u_1) - \mathcal{E}_k(u_2)$ also satisfies the same estimate (8.26) as $u_1 - u_2$.

We define an extension β_{ext} of β to the whole space $\lambda \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ by

$$y = x + \beta_{ext}(\varphi, x) \quad \Leftrightarrow \quad x = y + \breve{\beta}_{ext}(\varphi, y) \quad \forall x, y \in \mathbb{T}, \ \varphi \in \mathbb{T}^\nu$$

(note that, in general, β_{ext} and $\mathcal{E}_k(\beta)$ are two different extensions of β outside $\mathtt{DC}(\gamma,\tau)\times[\mathtt{h}_1,\mathtt{h}_2]$). The extended functions $\beta_{ext},\breve\beta_{ext}$ induce the operators $\mathcal{B}_{ext},\mathcal{B}_{ext}^{-1}$ by

$$(\mathcal{B}_{ext}h)(\varphi,x):=h(\varphi,x+\beta_{ext}(\varphi,x)),\qquad (\mathcal{B}_{ext}^{-1}h)(\varphi,y):=h(\varphi,y+\breve\beta_{ext}(\varphi,y)),\qquad \mathcal{B}_{ext}\circ\mathcal{B}_{ext}^{-1}=\mathrm{Id},$$

and they are defined for $\lambda\in\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$.

Notation: for simplicity, in the sequel we will drop the subscript "*ext*" and we rename

$$\beta_{ext}:=\beta,\quad \breve\beta_{ext}:=\breve\beta,\quad \mathcal{B}_{ext}:=\mathcal{B},\quad \mathcal{B}_{ext}^{-1}:=\mathcal{B}^{-1}. \tag{8.27}$$

We have the following estimates on the transformations $\mathcal{B}$ and $\mathcal{B}^{-1}$.

Lemma 8.5. *Let $\beta,\breve\beta$ be defined in* (8.27)*. There exists $\sigma:=\sigma(\tau,\nu,k_0)$ such that, if* (7.7) *holds with $\mu_0\geq\sigma$, then for any $s\geq s_2$,*

$$\|\beta\|_s^{k_0,\gamma},\|\breve\beta\|_s^{k_0,\gamma}\lesssim_s\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{8.28}$$

The operators $A=\mathcal{B}^{\pm1}-\mathrm{Id},(\mathcal{B}^{\pm1}-\mathrm{Id})^*$ satisfy the estimates

$$\|Ah\|_s^{k_0,\gamma}\lesssim_s\varepsilon\gamma^{-1}\Big(\|h\|_{s+k_0+1}^{k_0,\gamma}+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|h\|_{s_0+k_0+2}^{k_0,\gamma}\Big)\quad\forall s\geq s_2. \tag{8.29}$$

Let i_1,i_2 be two given embedded tori. Then, denoting $\Delta_{12}\beta=\beta(i_2)-\beta(i_1)$ and similarly for the other quantities, we have

$$\|\Delta_{12}\beta\|_{s_1},\|\Delta_{12}\breve\beta\|_{s_1}\lesssim_{s_1}\varepsilon\gamma^{-1}\|i_1-i_2\|_{s_1+\sigma}\,, \tag{8.30}$$

$$\|(\Delta_{12}A)[h]\|_{s_1}\lesssim_{s_1}\varepsilon\gamma^{-1}\|i_1-i_2\|_{s_1+\sigma}\|h\|_{s_1+1}\,,\quad A\in\{\mathcal{B}^{\pm1},(\mathcal{B}^{\pm1})^*\}\,, \tag{8.31}$$

where s_1 is introduced in (7.8)*.*

Proof. The bound (8.28) for $\breve\beta$ follows, recalling that $\breve\beta=u$, by (8.16) and (7.16). Estimate (8.28) for β follows by that for $\breve\beta$, applying inequality (2.12). We now prove estimate (8.29) for $\mathcal{B}-\mathrm{Id}$. We have

$$(\mathcal{B}-\mathrm{Id})h=\beta\int_0^1\mathcal{B}_\tau[h_x]\,d\tau\,,\qquad \mathcal{B}_\tau[f](\varphi,x):=f(\varphi,x+\tau\beta(\varphi,x))\,.$$

Then (8.29) follows by applying (2.11) to the operator $\mathcal{B}_\tau$, using the estimates on β, ansatz (7.7) and the interpolation estimate (2.10). The estimate for $\mathcal{B}^{-1}-\mathrm{Id}$ is obtained similarly. The estimate on the adjoint operators follows because

$$\mathcal{B}^*h(\varphi,y)=(1+\breve\beta(\varphi,y))h(\varphi,y+\breve\beta(\varphi,y)),\quad (\mathcal{B}^{-1})^*h(\varphi,x)=(1+\beta(\varphi,x))h(\varphi,x+\beta(\varphi,x))\,.$$

Estimate (8.30) for $\Delta_{12}\breve\beta$ follows by Lemma 8.4, and by (7.16)-(7.17). The estimate for $\Delta_{12}\beta$ is proved using the identities

$$\beta_k(\varphi,x)+\breve\beta_k(\varphi,x+\beta_k(\varphi,x))=0,\quad \beta_k:=\beta(i_k)\,,\quad k=1,2\,,$$

whence

$$\Delta_{12}\beta(\varphi,x)=-\Delta_{12}\breve\beta(\varphi,x+\beta_1(\varphi,x))+\Big(\breve\beta_2(\varphi,x+\beta_1(\varphi,x))-\breve\beta_2(\varphi,x+\beta_2(\varphi,x))\Big)\,,$$

so that, by (7.9), (8.29), (8.28) and by the estimates on composition of functions of Lemma 2.4, one gets

$$\|\Delta_{12}\beta\|_{s_1}\lesssim_{s_1}\|\Delta_{12}\breve\beta\|_{s_1}+\varepsilon\gamma^{-1}\|\Delta_{12}\beta\|_{s_1}\,.$$

Estimate (8.30) for $\Delta_{12}\beta$ follows by taking $\varepsilon\gamma^{-1}$ small enough with respect to some constant $C(s_1)>0$. □

We now conjugate the whole operator $\mathcal{L}_0$ in (7.13) by the diffeomorphism $\mathcal{B}$.

Lemma 8.6. *Let $\beta, \breve{\beta}, \mathcal{B}, \mathcal{B}^{-1}$ be defined in (8.27). For all $\lambda \in \mathtt{DC}(\gamma,\tau) \times [\mathtt{h}_1, \mathtt{h}_2]$, the transformation $\mathcal{B}$ conjugates the operator $\mathcal{L}_0$ defined in (7.13) to*

$$\mathcal{L}_1 := \mathcal{B}^{-1}\mathcal{L}_0\mathcal{B} = \omega\cdot\partial_\varphi + \begin{pmatrix} a_1 & -a_2\partial_y\mathcal{H}T_{\mathtt{h}} + \mathcal{R}_1 \\ a_3 & 0 \end{pmatrix}, \tag{8.32}$$

$$T_{\mathtt{h}} := \tanh(\mathtt{h}|D_y|) := \mathrm{Op}\big(\tanh(\mathtt{h}\chi(\xi)|\xi|)\big), \tag{8.33}$$

where a_1, a_2, a_3 are the functions

$$a_1(\varphi,y) := (\mathcal{B}^{-1}V_x)(\varphi,y), \quad a_2(\varphi,y) := 1 + (\mathcal{B}^{-1}\beta_x)(\varphi,y), \quad a_3(\varphi,y) := (\mathcal{B}^{-1}a)(\varphi,y), \tag{8.34}$$

and $\mathcal{R}_1$ is a pseudo-differential operator of order $OPS^{-\infty}$. Formula (8.34) defines the functions a_1, a_2, a_3 on the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$. The operator $\mathcal{R}_1$ admits an extension to $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ as well, which we also denote by $\mathcal{R}_1$. The real valued functions β, a_1, a_2, a_3 have parity

$$\beta = even(\varphi)odd(x); \qquad a_1 = odd(\varphi)even(y); \qquad a_2, a_3 = even(\varphi)even(y). \tag{8.35}$$

There exists $\sigma = \sigma(\tau,\nu,k_0) > 0$ such that for any $m, \alpha \geq 0$, assuming (7.7) with $\mu_0 \geq \sigma + m + \alpha$, for any $s \geq s_0$, on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ the following estimates hold:

$$\|a_1\|_s^{k_0,\gamma} + \|a_2 - 1\|_s^{k_0,\gamma} + \|a_3 - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big), \tag{8.36}$$

$$|\!|\!| \mathcal{R}_1 |\!|\!|_{-m,s,\alpha}^{k_0,\gamma} \lesssim_{m,s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+m+\alpha}^{k_0,\gamma}\big). \tag{8.37}$$

Finally, given two tori i_1, i_2, we have

$$\|\Delta_{12}a_1\|_{s_1} + \|\Delta_{12}a_2\|_{s_1} + \|\Delta_{12}a_3\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma}, \tag{8.38}$$

$$|\!|\!| \Delta_{12}\mathcal{R}_1 |\!|\!|_{-m,s_1,\alpha} \lesssim_{m,s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma+m+\alpha}. \tag{8.39}$$

Proof. By (7.19) and (8.3)-(8.5) we have that

$$\mathcal{L}_1 := \mathcal{B}^{-1}\mathcal{L}_0\mathcal{B} = \omega\cdot\partial_\varphi + \begin{pmatrix} a_1 & -\mathcal{B}^{-1}G(\eta)\mathcal{B} \\ a_3 & 0 \end{pmatrix} \tag{8.40}$$

where the functions a_1 and a_3 are defined in (8.34). We now conjugate the Dirichlet-Neumann operator $G(\eta)$ under the diffeomorphism $\mathcal{B}$. Following Proposition 3.1, we write

$$G(\eta) = |D_x|\tanh(\mathtt{h}|D_x|) + \mathcal{R}_G = \partial_x\mathcal{H}T_{\mathtt{h}} + \mathcal{R}_G, \qquad T_{\mathtt{h}} := \tanh(\mathtt{h}|D_x|), \tag{8.41}$$

where $\mathcal{R}_G$ is an integral operator in $OPS^{-\infty}$. Recall the decomposition (3.51), which is

$$T_{\mathtt{h}} = \mathrm{Id} + \mathrm{Op}(r_{\mathtt{h}}), \qquad r_{\mathtt{h}}(\xi) := -\frac{2}{1 + e^{2\mathtt{h}|\xi|\chi(\xi)}} \in S^{-\infty}. \tag{8.42}$$

Since $\mathcal{B}^{-1}\,\partial_x\,\mathcal{B} = a_2\partial_y$ where the function a_2 is defined in (8.34), we have

$$\begin{aligned}
\mathcal{B}^{-1}\partial_x\mathcal{H}T_{\mathtt{h}}\mathcal{B} &= (\mathcal{B}^{-1}\partial_x\mathcal{B})(\mathcal{B}^{-1}\mathcal{H}\mathcal{B})(\mathcal{B}^{-1}T_{\mathtt{h}}\mathcal{B}) \\
&= a_2\partial_y\{\mathcal{H} + (\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H})\}(\mathcal{B}^{-1}T_{\mathtt{h}}\mathcal{B}) \\
&= a_2\partial_y\mathcal{H}\{T_{\mathtt{h}} + [\mathcal{B}^{-1}\mathrm{Op}(r_{\mathtt{h}})\mathcal{B} - (T_{\mathtt{h}} - \mathrm{Id})]\} + a_2\partial_y(\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H})(\mathcal{B}^{-1}T_{\mathtt{h}}\mathcal{B}) \\
&= a_2\partial_y\mathcal{H}T_{\mathtt{h}} + a_2\partial_y\mathcal{H}[\mathcal{B}^{-1}\mathrm{Op}(r_{\mathtt{h}})\mathcal{B} - \mathrm{Op}(r_{\mathtt{h}})] + a_2\partial_y(\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H})(\mathcal{B}^{-1}T_{\mathtt{h}}\mathcal{B}).
\end{aligned} \tag{8.43}$$

Therefore by (8.41)-(8.43) we get

$$-\mathcal{B}^{-1}G(\eta)\mathcal{B} = -a_2\partial_y\mathcal{H}T_{\mathtt{h}} + \mathcal{R}_1, \tag{8.44}$$

where $\mathcal{R}_1$ is the operator in $OPS^{-\infty}$ defined by

$$\begin{aligned}
&\mathcal{R}_1 := \mathcal{R}_1^{(1)} + \mathcal{R}_1^{(2)} + \mathcal{R}_1^{(3)} \\
&\mathcal{R}_1^{(1)} := -\mathcal{B}^{-1}\mathcal{R}_G\mathcal{B}\,, \\
&\mathcal{R}_1^{(2)} := -a_2\partial_y\mathcal{H}[\mathcal{B}^{-1}\mathrm{Op}(r_{\mathtt{h}})\mathcal{B} - \mathrm{Op}(r_{\mathtt{h}})]\,, \\
&\mathcal{R}_1^{(3)} := -a_2\partial_y(\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H})\mathcal{B}^{-1}T_{\mathtt{h}}\mathcal{B}\,.
\end{aligned} \tag{8.45}$$

Notice that $\mathcal{B}^{-1}\mathcal{R}_G\mathcal{B}$ and $\mathcal{B}^{-1}\mathrm{Op}(r_{\mathtt{h}})\mathcal{B}$ are in $OPS^{-\infty}$ since $\mathcal{R}_G$ and $\mathrm{Op}(r_{\mathtt{h}})$, defined in (8.41)-(8.42), are in $OPS^{-\infty}$. The operator $\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H}$ is in $OPS^{-\infty}$ by Lemma 2.18.

In conclusion, (8.40) and (8.44) imply (8.32)-(8.34), for all λ in the Cantor set $\mathtt{DC}(\gamma,\tau) \times [\mathtt{h}_1,\mathtt{h}_2]$. By formulas (8.45), $\mathcal{R}_1$ is defined on the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1,\mathtt{h}_2]$.

Estimates (8.36), (8.38) for a_1, a_2, a_3 on $\mathbb{R}^\nu \times [\mathtt{h}_1,\mathtt{h}_2]$ follow by (7.16), (7.17) and Lemma 8.5. We now prove the bounds (8.37), (8.39). We estimate separately the three terms in (8.45).

Estimate of $\mathcal{R}_1^{(1)}$. By Proposition 3.1 and Lemma 2.16, $\mathcal{B}^{-1}\mathcal{R}_G\mathcal{B}$ is an integral operator with $\mathcal{C}^\infty$ kernel

$$\breve{K}_G(\varphi,x,z) := \big(1 + \partial_z\breve{\beta}(\varphi,z)\big)K_G(\varphi, x + \breve{\beta}(\varphi,x), z + \breve{\beta}(\varphi,z))\,,$$

where K_G is the $\mathcal{C}^\infty$ kernel of $\mathcal{R}_G$. Applying (2.57), (2.60), Proposition 3.1, and using (7.10), (7.11) and the estimates of Lemma 8.5, we get (8.37) and (8.39) for $\mathcal{R}_1^{(1)}$.

Estimate of $\mathcal{R}_1^{(2)}$. Since the symbol $r_{\mathtt{h}} \in S^{-\infty}$ (see (8.42)), by Lemma 2.16 the operator $\mathcal{B}^{-1}\mathrm{Op}(r_{\mathtt{h}})\mathcal{B} - \mathrm{Op}(r_{\mathtt{h}})$ is an integral operator with $\mathcal{C}^\infty$ kernel

$$\big(1 + \partial_z\breve{\beta}(\varphi,z)\big)K_{r_{\mathtt{h}}}(y + \breve{\beta}(\varphi,y), z + \breve{\beta}(\varphi,z)) - K_{r_{\mathtt{h}}}(y,z)\,,$$

where $K_{r_{\mathtt{h}}}$ is the $\mathcal{C}^\infty$ kernel associated to $r_{\mathtt{h}}$ (see (2.63)). Hence the kernel associated to $\mathcal{R}_1^{(2)}$ is given by

$$K_1^{(2)}(\varphi,y,z) := a_2(\varphi,y)\mathcal{H}_y\partial_y\Big(\big(1 + \partial_z\breve{\beta}(\varphi,z)\big)K_{r_{\mathtt{h}}}(y + \breve{\beta}(\varphi,y), z + \breve{\beta}(\varphi,z)) - K_{r_{\mathtt{h}}}(y,z)\Big)$$

(note that $\mathcal{H}_y$ is the Hilbert transform with respect to the variable y). By Lemmata 2.16, 2.19, by the estimates of Lemma 8.5 and using also (7.10), (7.11), (8.36), (8.38), one gets

$$\|K_1^{(2)}\|^{k_0,\gamma}_{\mathcal{C}^{s+m+\alpha}} \lesssim_{s,m,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+m+\alpha+\sigma}\big)\,, \quad \|\Delta_{12}K_1^{(2)}\|_{\mathcal{C}^{s_1+m+\alpha}} \lesssim_{s_1,m,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+m+\alpha+\sigma}$$

for $\alpha, m \geq 0$, for some $\sigma = \sigma(\tau,\nu,k_0) > 0$. Estimates (8.37), (8.39) for $\mathcal{R}_1^{(2)}$ follow by Lemma 2.15.

Estimate of $\mathcal{R}_1^{(3)}$. Let $K_{\mathcal{B}}$ be the $\mathcal{C}^\infty$ kernel of the operator $\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H}$ given in (2.61), (2.62) with β instead of p and $\breve{\beta}$ instead of q. One has

$$\begin{aligned}
\mathcal{R}_1^{(3)}u(\varphi,y) &= -a_2(\varphi,y)\partial_y\int_{\mathbb{T}} K_{\mathcal{B}}(\varphi,y,z)(\mathcal{B}^{-1}T_{\mathtt{h}}\mathcal{B}u)(\varphi,z)\,dz \\
&= -a_2(\varphi,y)\partial_y\int_{\mathbb{T}} \big(\mathcal{B}^*T_{\mathtt{h}}(\mathcal{B}^{-1})^*K_{\mathcal{B}}(\varphi,y,z)\big)u(\varphi,z)\,dz
\end{aligned} \tag{8.46}$$

using that $T_{\mathtt{h}}^* = T_{\mathtt{h}}$. Hence $\mathcal{R}_1^{(3)}$ is an integral operator with kernel $K_1^{(3)}$ given by

$$K_1^{(3)}(\varphi,x,z) := -a_2(\varphi,y)\partial_y\Big(\mathcal{B}^*T_{\mathtt{h}}(\mathcal{B}^{-1})^*K_{\mathcal{B}}(\varphi,y,z)\Big)\,.$$

Then by Lemmata 2.18, 8.5 and by (7.10), (7.11), (8.36), (8.38), we get

$$\|K_1^{(3)}\|^{k_0,\gamma}_{\mathcal{C}^{s+m+\alpha}} \lesssim_{s,m,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+m+\alpha+\sigma}\big)\,, \quad \|\Delta_{12}K_1^{(3)}\|_{\mathcal{C}^{s_1+m+\alpha}} \lesssim_{s_1,m,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+m+\alpha+\sigma}$$

for $m, \alpha \geq 0$, for some $\sigma = \sigma(\tau,\nu,k_0) > 0$. Thus estimates (8.37), (8.39) for $\mathcal{R}_1^{(3)}$ follow by Lemma 2.15. □

Remark 8.7. *We stress that the conjugation identity* (8.32) *holds only on the Cantor set* $\mathtt{DC}(\gamma,\tau)\times[\mathtt{h}_1,\mathtt{h}_2]$. *It is technically convenient to consider the extension of* $a_1, a_2, a_3, \mathcal{R}_1$ *to the whole parameter space* $\mathbb{R}^\nu \times [\mathtt{h}_1,\mathtt{h}_2]$, *in order to directly use the results of Section 2.3 expressed by means of classical derivatives with respect to the parameter* λ. *Formulas* (8.34) *and* (8.45) *define* $a_1, a_2, a_3, \mathcal{R}_1$ *on the whole parameter space* $\mathbb{R}^\nu \times [\mathtt{h}_1,\mathtt{h}_2]$. *Note that the resulting extended operator* $\mathcal{L}_1$ *in the right hand side of* (8.32) *is defined on* $\mathbb{R}^\nu \times [\mathtt{h}_1,\mathtt{h}_2]$, *and in general it is different from* $\mathcal{B}^{-1}\mathcal{L}_0\mathcal{B}$ *outside* $\mathtt{DC}(\gamma,\tau)\times[\mathtt{h}_1,\mathtt{h}_2]$.

In the sequel we rename in (8.32)-(8.35) the space variable y by x.

9 Change of the space variable

We consider a φ-independent diffeomorphism of the torus $\mathbb{T}$ of the form

$$y = x + \alpha(x) \qquad \text{with inverse} \qquad x = y + \breve{\alpha}(y) \tag{9.1}$$

where α is a $\mathcal{C}^\infty(\mathbb{T}_x)$ real valued function, independent of φ, satisfying $\|\alpha_x\|_{L^\infty} \le 1/2$. We also make the following ansatz on α that will be verified when we choose it in Section 12, see formula (12.25): the function α is odd(x) and $\alpha = \alpha(\lambda) = \alpha(\lambda, i_0(\lambda))$, $\lambda \in \mathbb{R}^{\nu+1}$ is k_0 times differentiable with respect to the parameter $\lambda \in \mathbb{R}^{\nu+1}$ with $\partial_\lambda^k \alpha \in \mathcal{C}^\infty(\mathbb{T})$ for any $k \in \mathbb{N}^{\nu+1}$, $|k| \le k_0$, and it satisfies the estimate

$$\begin{aligned} &\|\alpha\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \forall s \ge s_0\,, \\ &\|\Delta_{12}\alpha\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,, \end{aligned} \tag{9.2}$$

for some $\sigma = \sigma(k_0,\tau,\nu) > 0$. By (9.2) and Lemma 2.4, arguing as in the proof of Lemma 8.5, one gets

$$\begin{aligned} &\|\breve{\alpha}\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \forall s \ge s_0\,, \\ &\|\Delta_{12}\breve{\alpha}\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,, \end{aligned} \tag{9.3}$$

for some $\sigma = \sigma(k_0,\tau,\nu) > 0$. Furthermore, the function $\breve{\alpha}(y)$ is odd(y).

We conjugate the operator $\mathcal{L}_1$ in (8.32) by the composition operator

$$(\mathcal{A}u)(\varphi, x) := u(\varphi, x + \alpha(x)), \quad (\mathcal{A}^{-1}u)(\varphi, y) := u(\varphi, y + \breve{\alpha}(y))\,. \tag{9.4}$$

By (8.32), using that the operator $\mathcal{A}$ is φ-independent, recalling expansion (3.51), namely

$$T_{\mathtt{h}} = \mathrm{Id} + \mathrm{Op}(r_{\mathtt{h}})\,, \quad r_{\mathtt{h}}(\xi) = -\frac{2}{1 + e^{2\mathtt{h}|\xi|\chi(\xi)}} \in S^{-\infty}\,,$$

and arguing as in (8.43) to compute the conjugation $\mathcal{A}^{-1}\big(- a_2\partial_x \mathcal{H} T_{\mathtt{h}}\big)\mathcal{A}$, one has

$$\mathcal{L}_2 := \mathcal{A}^{-1}\mathcal{L}_1\mathcal{A} = \omega\cdot\partial_\varphi + \begin{pmatrix} a_4 & -a_5\partial_y \mathcal{H} T_{\mathtt{h}} + \mathcal{R}_2 \\ a_6 & 0 \end{pmatrix}, \tag{9.5}$$

where a_4, a_5, a_6 are the functions

$$a_4(\varphi, y) := (\mathcal{A}^{-1}a_1)(\varphi, y) = a_1(\varphi, y + \breve{\alpha}(y))\,, \tag{9.6}$$

$$a_5(\varphi, y) := \big(\mathcal{A}^{-1}(a_2(1+\alpha_x))\big)(\varphi, y) = \{a_2(\varphi, x)(1 + \alpha_x(x))\}|_{x = y + \breve{\alpha}(y)} \tag{9.7}$$

$$a_6(\varphi, y) := (\mathcal{A}^{-1}a_3)(\varphi, y) = a_3(\varphi, y + \breve{\alpha}(y)) \tag{9.8}$$

and $\mathcal{R}_2$ is the operator in $OPS^{-\infty}$ given by

$$\mathcal{R}_2 := -a_5\partial_y\mathcal{H}\big[\mathcal{A}^{-1}\mathrm{Op}(r_{\mathtt{h}})\mathcal{A} - \mathrm{Op}(r_{\mathtt{h}})\big] - a_5\partial_y(\mathcal{A}^{-1}\mathcal{H}\mathcal{A} - \mathcal{H})(\mathcal{A}^{-1}T_{\mathtt{h}}\mathcal{A}) + \mathcal{A}^{-1}\mathcal{R}_1\mathcal{A}\,. \tag{9.9}$$

Lemma 9.1. *There exists a constant $\sigma = \sigma(k_0, \tau, \nu) > 0$ such that, if (7.7) holds with $\mu_0 \geq \sigma$, then the following holds: the operators $A \in \{\mathcal{A}^{\pm 1} - \mathrm{Id}, (\mathcal{A}^{\pm 1} - \mathrm{Id})^*\}$ are even and reversibility preserving and satisfy*

$$\begin{aligned} &\|Ah\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(\|h\|_{s+k_0+1}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|h\|_{s_0+k_0+2}^{k_0,\gamma}\big)\,, \quad \forall s \geq s_0\,, \\ &\|(\Delta_{12}A)h\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma}\|h\|_{s_1+1}\,. \end{aligned} \tag{9.10}$$

The real valued functions a_4, a_5, a_6 in (9.6)-(9.8) *satisfy*

$$a_4 = odd(\varphi)even(y), \qquad a_5, a_6 = even(\varphi)even(y)\,, \tag{9.11}$$

and

$$\begin{aligned} &\|a_4\|_s^{k_0,\gamma}\,, \|a_5 - 1\|_s^{k_0,\gamma}\,, \|a_6 - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) \\ &\|\Delta_{12}a_4\|_{s_1}\,, \|\Delta_{12}a_5\|_{s_1}\,, \|\Delta_{12}a_6\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma}\,. \end{aligned} \tag{9.12}$$

The remainder $\mathcal{R}_2$ defined in (9.9) *is an even and reversible pseudo-differential operator in $OPS^{-\infty}$. Moreover, for any $m, \alpha \geq 0$, and assuming* (7.7) *with $\sigma + m + \alpha \leq \mu_0$, the following estimates hold:*

$$\begin{aligned} &|\mathcal{R}_2|_{-m,s,\alpha}^{k_0,\gamma} \lesssim_{m,s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+m+\alpha}^{k_0,\gamma}\big)\,, \quad \forall s \geq s_0 \\ &|\Delta_{12}\mathcal{R}_2|_{-m,s_1,\alpha} \lesssim_{m,s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma+m+\alpha}\,. \end{aligned} \tag{9.13}$$

Proof. The transformations $\mathcal{A}^{\pm 1} - \mathrm{Id}, (\mathcal{A}^{\pm 1} - \mathrm{Id})^*$ are even and reversibility preserving because α and $\breve{\alpha}$ are odd functions. Estimate (9.10) can be proved by using (9.2), (9.3), arguing as in the proof of Lemma 8.5.

Estimate (9.12) follows by definitions (9.6)-(9.8), by estimates (9.2), (9.3), (9.10), (8.36), (8.38), and by applying Lemma 2.4.

Estimates (9.13) of the remainder $\mathcal{R}_2$ follow by using the same arguments we used in Lemma 8.6 to get estimates (8.37), (8.39) for the remainder $\mathcal{R}_1$. □

In the sequel we rename in (9.5)-(9.9) the space variable y by x.

10 Symmetrization of the highest order

The aim of this section is to conjugate the operator $\mathcal{L}_2$ defined in (9.5) to a new operator $\mathcal{L}_4$ in which the highest order derivatives appear in the off-diagonal entries with the same order and opposite coefficients (see (10.10)-(10.14)). In the complex variables $(u, \bar{u})$ that we will introduce in Section 11, this amounts to the symmetrization of the linear operator at the highest order, see (11.1)-(11.3).

We first conjugate $\mathcal{L}_2$ by the real, even and reversibility preserving transformation

$$\mathcal{M}_2 := \begin{pmatrix} \Lambda_{\mathtt{h}} & 0 \\ 0 & \Lambda_{\mathtt{h}}^{-1} \end{pmatrix}, \tag{10.1}$$

where $\Lambda_{\mathtt{h}}$ is the Fourier multiplier, acting on the periodic functions,

$$\Lambda_{\mathtt{h}} := \pi_0 + |D|^{\frac{1}{4}} T_{\mathtt{h}}^{\frac{1}{4}}\,, \qquad \text{with inverse} \qquad \Lambda_{\mathtt{h}}^{-1} = \pi_0 + |D|^{-\frac{1}{4}} T_{\mathtt{h}}^{-\frac{1}{4}}\,, \tag{10.2}$$

with $T_{\mathtt{h}} = \tanh(\mathtt{h}|D|)$ and π_0 defined in (2.35). The conjugated operator is

$$\mathcal{L}_3 := \mathcal{M}_2^{-1}\mathcal{L}_2\mathcal{M}_2 = \omega\cdot\partial_\varphi + \begin{pmatrix} \Lambda_{\mathtt{h}}^{-1}a_4\Lambda_{\mathtt{h}} & \Lambda_{\mathtt{h}}^{-1}(-a_5\partial_x\mathcal{H}T_{\mathtt{h}} + \mathcal{R}_2)\Lambda_{\mathtt{h}}^{-1} \\ \Lambda_{\mathtt{h}}a_6\Lambda_{\mathtt{h}} & 0 \end{pmatrix} =: \omega\cdot\partial_\varphi + \begin{pmatrix} A_3 & B_3 \\ C_3 & 0 \end{pmatrix}. \tag{10.3}$$

We develop the operators in (10.3) up to order $-1/2$. First we write

$$A_3 = \Lambda_{\mathtt{h}}^{-1}a_4\Lambda_{\mathtt{h}} = a_4 + \mathcal{R}_{A_3} \qquad \text{where} \qquad \mathcal{R}_{A_3} := [\Lambda_{\mathtt{h}}^{-1}, a_4]\Lambda_{\mathtt{h}} \in OPS^{-1} \tag{10.4}$$

by Lemma 2.11. Using that $|D|^m \pi_0 = \pi_0 |D|^m = 0$ for any $m \in \mathbb{R}$ and that $\pi_0^2 = \pi_0$ on the periodic functions, one has

$$
\begin{aligned}
C_3 &= \Lambda_{\mathtt{h}} a_6 \Lambda_{\mathtt{h}} = a_6 \Lambda_{\mathtt{h}}^2 + [\Lambda_{\mathtt{h}}, a_6]\Lambda_{\mathtt{h}} = a_6(\pi_0 + |D|^{\frac14} T_{\mathtt{h}}^{\frac14})^2 + [\Lambda_{\mathtt{h}}, a_6]\Lambda_{\mathtt{h}} \\
&= a_6 |D|^{\frac12} T_{\mathtt{h}}^{\frac12} + \pi_0 + \mathcal{R}_{C_3} \qquad \text{where} \qquad \mathcal{R}_{C_3} := (a_6 - 1)\pi_0 + [\Lambda_{\mathtt{h}}, a_6]\Lambda_{\mathtt{h}} \,.
\end{aligned} \tag{10.5}
$$

Using that $|D| = \mathcal{H}\partial_x$, (10.2) and $|D|\pi_0 = 0$ on the periodic functions, we write B_3 in (10.3) as

$$
\begin{aligned}
B_3 &= \Lambda_{\mathtt{h}}^{-1}(-a_5 \partial_x \mathcal{H} T_{\mathtt{h}} + \mathcal{R}_2)\Lambda_{\mathtt{h}}^{-1} = -a_5 |D| T_{\mathtt{h}} \Lambda_{\mathtt{h}}^{-2} - [\Lambda_{\mathtt{h}}^{-1}, a_5]|D| T_{\mathtt{h}} \Lambda_{\mathtt{h}}^{-1} + \Lambda_{\mathtt{h}}^{-1} \mathcal{R}_2 \Lambda_{\mathtt{h}}^{-1} \\
&= -a_5 |D| T_{\mathtt{h}} \big(\pi_0 + |D|^{-\frac14} T_{\mathtt{h}}^{-\frac14}\big)^2 - [\Lambda_{\mathtt{h}}^{-1}, a_5]|D| T_{\mathtt{h}} \Lambda_{\mathtt{h}}^{-1} + \Lambda_{\mathtt{h}}^{-1} \mathcal{R}_2 \Lambda_{\mathtt{h}}^{-1} \\
&= -a_5 |D|^{\frac12} T_{\mathtt{h}}^{\frac12} + \mathcal{R}_{B_3} \qquad \text{where} \qquad \mathcal{R}_{B_3} := -[\Lambda_{\mathtt{h}}^{-1}, a_5]|D| T_{\mathtt{h}} \Lambda_{\mathtt{h}}^{-1} + \Lambda_{\mathtt{h}}^{-1} \mathcal{R}_2 \Lambda_{\mathtt{h}}^{-1} \,.
\end{aligned} \tag{10.6}
$$

In the next lemma we provide some estimates on $\Lambda_{\mathtt{h}}$ and the remainders $\mathcal{R}_{A_3}$, $\mathcal{R}_{B_3}$, $\mathcal{R}_{C_3}$.

Lemma 10.1. *The operators $\Lambda_{\mathtt{h}} \in OPS^{\frac14}$, $\Lambda_{\mathtt{h}}^{-1} \in OPS^{-\frac14}$ and $\mathcal{R}_{A_3}, \mathcal{R}_{B_3}, \mathcal{R}_{C_3} \in OPS^{-\frac12}$. Furthermore, there exists $\sigma(k_0, \tau, \nu) > 0$ such that for any $\alpha > 0$, assuming (7.7) with $\mu_0 \geq \sigma + \alpha$, then for all $s \geq s_0$,*

$$
|\!|\!| \Lambda_{\mathtt{h}} |\!|\!|^{k_0,\gamma}_{\frac14, s, \alpha} \,, |\!|\!| \Lambda_{\mathtt{h}}^{-1} |\!|\!|^{k_0,\gamma}_{-\frac14, s, \alpha} \lesssim_\alpha 1 \,, \tag{10.7}
$$

$$
|\!|\!| \mathcal{R} |\!|\!|^{k_0,\gamma}_{-\frac12, s, \alpha} \lesssim_{s,\alpha} \varepsilon \gamma^{-1} \big(1 + \| \mathfrak{I}_0 \|^{k_0,\gamma}_{s+\sigma+\alpha}\big) \,, \quad |\!|\!| \Delta_{12} \mathcal{R} |\!|\!|_{-\frac12, s_1, \alpha} \lesssim_{s_1,\alpha} \varepsilon \gamma^{-1} \| \Delta_{12} i \|_{s_1+\sigma+\alpha} \tag{10.8}
$$

for all $\mathcal{R} \in \{\mathcal{R}_{A_3}, \mathcal{R}_{B_3}, \mathcal{R}_{C_3}\}$. The operator $\mathcal{L}_3$ in (10.3) *is real, even and reversible.*

Proof. The lemma follows by the definitions of $\mathcal{R}_{A_3}$, $\mathcal{R}_{B_3}$, $\mathcal{R}_{C_3}$ in (10.4), (10.6), (10.5), by Lemmata 2.10 and 2.11, recalling (2.41) and using (9.12), (9.13). □

Consider now a transformation $\mathcal{M}_3$ of the form

$$
\mathcal{M}_3 := \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathcal{M}_3^{-1} = \begin{pmatrix} p^{-1} & 0 \\ 0 & 1 \end{pmatrix}, \tag{10.9}
$$

where $p(\varphi, x)$ is a real-valued periodic function, with $p-1$ small, that we shall fix in (10.14). The conjugated operator is

$$
\mathcal{L}_4 := \mathcal{M}_3^{-1} \mathcal{L}_3 \mathcal{M}_3 = \omega \cdot \partial_\varphi + \begin{pmatrix} p^{-1}(\omega \cdot \partial_\varphi p) + p^{-1} A_3 p & p^{-1} B_3 \\ C_3 p & 0 \end{pmatrix} = \omega \cdot \partial_\varphi + \begin{pmatrix} A_4 & B_4 \\ C_4 & 0 \end{pmatrix} \tag{10.10}
$$

where, recalling (10.4), (10.6), (10.5), one has

$$
A_4 = \breve{a}_4 + \mathcal{R}_{A_4} \,, \quad \breve{a}_4 := a_4 + p^{-1}(\omega \cdot \partial_\varphi p) \,, \quad \mathcal{R}_{A_4} := p^{-1} \mathcal{R}_{A_3} p \tag{10.11}
$$

$$
B_4 = -p^{-1} a_5 |D|^{\frac12} T_{\mathtt{h}}^{\frac12} + \mathcal{R}_{B_4} \,, \quad \mathcal{R}_{B_4} := p^{-1} \mathcal{R}_{B_3} \tag{10.12}
$$

$$
C_4 = a_6 p |D|^{\frac12} T_{\mathtt{h}}^{\frac12} + \pi_0 + \mathcal{R}_{C_4} \,, \quad \mathcal{R}_{C_4} := a_6 [|D|^{\frac12} T_{\mathtt{h}}^{\frac12}, p] + \pi_0(p-1) + \mathcal{R}_{C_3} p \tag{10.13}
$$

and therefore $\mathcal{R}_{A_4}, \mathcal{R}_{B_4}, \mathcal{R}_{C_4} \in OPS^{-\frac12}$. The coefficients of the highest order term in B_4 in (10.12) and C_4 in (10.13) are opposite if $a_6 p = p^{-1} a_5$. Therefore we fix the real valued function

$$
p := \sqrt{\frac{a_5}{a_6}} \,, \qquad a_6 p = p^{-1} a_5 = \sqrt{a_5 a_6} =: a_7 \,. \tag{10.14}
$$

Lemma 10.2. *There exists $\sigma := \sigma(\tau, \nu, k_0) > 0$ such that for any $\alpha > 0$, assuming (7.7) with $\mu_0 \geq \sigma + \alpha$, then for any $s \geq s_0$ the following holds. The transformation $\mathcal{M}_3$ defined in (10.9) is real, even and reversibility preserving and satisfies*

$$
|\!|\!| \mathcal{M}_3^{\pm 1} - \mathrm{Id} |\!|\!|^{k_0,\gamma}_{0,s,0} \lesssim_s \varepsilon \gamma^{-1} \big(1 + \| \mathfrak{I}_0 \|^{k_0,\gamma}_{s+\sigma}\big) \,. \tag{10.15}
$$

The real valued functions $\breve{a}_4, a_7$ *defined in* (10.11), (10.14) *satisfy*

$$\breve{a}_4 = \text{odd}(\varphi)\text{even}(x)\,, \qquad a_7 = \text{even}(\varphi)\text{even}(x)\,, \tag{10.16}$$

and, for any $s \geq s_0$,

$$\|\breve{a}_4\|_s^{k_0,\gamma}, \|a_7 - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{10.17}$$

The remainders $\mathcal{R}_{A_4}, \mathcal{R}_{B_4}, \mathcal{R}_{C_4} \in OPS^{-\frac{1}{2}}$ *defined in* (10.11)-(10.13) *satisfy*

$$|\mathcal{R}|_{-\frac{1}{2},s,\alpha}^{k_0,\gamma} \lesssim_{s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\alpha}^{k_0,\gamma}\big)\,, \qquad \mathcal{R} \in \{\mathcal{R}_{A_4}, \mathcal{R}_{B_4}, \mathcal{R}_{C_4}\}\,. \tag{10.18}$$

Let i_1, i_2 *be given embedded tori. Then*

$$|\Delta_{12}\mathcal{M}_3^{\pm 1}|_{0,s_1,0} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,, \tag{10.19}$$

$$\|\Delta_{12}\breve{a}_4\|_{s_1}, \|\Delta_{12} a_7\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,, \tag{10.20}$$

$$|\Delta_{12}\mathcal{R}|_{-\frac{1}{2},s_1,\alpha} \lesssim_{s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma+\alpha}\,, \qquad \mathcal{R} \in \{\mathcal{R}_{A_4}, \mathcal{R}_{B_4}, \mathcal{R}_{C_4}\}\,. \tag{10.21}$$

The operator $\mathcal{L}_4$ *in* (10.10) *is real, even and reversible.*

Proof. By (9.11), the functions a_5, a_6 are even(φ)even(x), and therefore p is even(φ)even(x). Moreover, since a_4 is odd(φ)even(x), we deduce (10.16). Since p is even(φ)even(x), the transformation $\mathcal{M}_3$ is real, even and reversibility preserving.

By definition (10.14), Lemma 2.6, the interpolation estimate (2.10) and applying estimates (9.12) on a_5 and a_6, one gets that p satisfies the estimates

$$\|p^{\pm 1} - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \|\Delta_{12} p^{\pm 1}\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma} \tag{10.22}$$

for some $\sigma = \sigma(\tau, \nu, k_0) > 0$. Hence estimates (10.15), (10.19) for $\mathcal{M}_3^{\pm 1}$ follow by definition (10.9), using estimates (2.41), (10.22).

Estimates (10.17), (10.20) for $\breve{a}_4, a_7$ follow by definitions (10.11), (10.14) and applying estimates (9.12) on a_4, a_5 and a_6, estimates (10.22) on p, Lemma 2.6 and the interpolation estimate (2.10).

Estimates (10.18), (10.21) follow by definitions (10.11)-(10.13), estimate (2.41), Lemmata 2.10 and 2.11, bounds (9.12) on a_4, a_5, a_6, (10.22) on p, and Lemma 10.1. □

11 Symmetrization of the lower orders

To symmetrize the linear operator $\mathcal{L}_4$ in (10.10), with p fixed in (10.14), at lower orders, it is convenient to introduce the complex coordinates $(u, \bar{u}) := \mathcal{C}^{-1}(\eta, \psi)$, with $\mathcal{C}$ defined in (2.67), namely $u = \eta + \mathrm{i}\psi$, $\bar{u} = \eta - \mathrm{i}\psi$. In these complex coordinates the linear operator $\mathcal{L}_4$ becomes, using (2.68) and (10.14),

$$\mathcal{L}_5 := \mathcal{C}^{-1}\mathcal{L}_4\mathcal{C} = \omega\cdot\partial_\varphi + \mathrm{i} a_7 |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}}\Sigma + a_8\mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_5 + \mathcal{Q}_5\,, \qquad a_8 := \frac{\breve{a}_4}{2}\,, \tag{11.1}$$

where the real valued functions $a_7, \breve{a}_4$ are defined in (10.14), (10.11) and satisfy (10.16),

$$\Sigma := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \Pi_0 := \frac{1}{2}\begin{pmatrix} \pi_0 & \pi_0 \\ -\pi_0 & -\pi_0 \end{pmatrix}, \qquad \mathbb{I}_2 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{11.2}$$

π_0 is defined in (2.35), and

$$\begin{gathered} \mathcal{P}_5 := \begin{pmatrix} P_5 & 0 \\ 0 & \overline{P_5} \end{pmatrix}, \quad \mathcal{Q}_5 := \begin{pmatrix} 0 & Q_5 \\ \overline{Q_5} & 0 \end{pmatrix}, \\ P_5 := \frac{1}{2}\big\{\mathcal{R}_{A_4} + \mathrm{i}(\mathcal{R}_{C_4} - \mathcal{R}_{B_4})\big\}\,, \qquad Q_5 := a_8 + \frac{1}{2}\big\{\mathcal{R}_{A_4} + \mathrm{i}(\mathcal{R}_{C_4} + \mathcal{R}_{B_4})\big\}\,. \end{gathered} \tag{11.3}$$

By the estimates of Lemma 10.2 we have

$$\|a_7 - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \|\Delta_{12} a_7\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma} \tag{11.4}$$

$$\|a_8\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \|\Delta_{12} a_8\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,, \tag{11.5}$$

$$|\mathcal{P}_5|_{-\frac12,s,\alpha}^{k_0,\gamma}\,,\ |\mathcal{Q}_5|_{0,s,\alpha}^{k_0,\gamma} \lesssim_{s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\alpha}^{k_0,\gamma}\big) \tag{11.6}$$

$$|\Delta_{12}\mathcal{P}_5|_{-\frac12,s_1,\alpha}\,,\ |\Delta_{12}\mathcal{Q}_5|_{0,s_1,\alpha} \lesssim_{s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma+\alpha}\,. \tag{11.7}$$

Now we define inductively a finite number of transformations to remove all the terms of orders $\geq -M$ from the off-diagonal operator $\mathcal{Q}_5$. The constant M will be fixed in (15.16).

Let $\mathcal{L}_5^{(0)} := \mathcal{L}_5$, $P_5^{(0)} := P_5$ and $Q_5^{(0)} := Q_5$. In the rest of the section we prove the following inductive claim:

- Symmetrization of $\mathcal{L}_5^{(0)}$ in decreasing orders. For $m \geq 0$, there is a real, even and reversible operator of the form

$$\mathcal{L}_5^{(m)} := \omega\cdot\partial_\varphi + \mathrm{i} a_7 |D|^{\frac12} T_{\mathtt h}^{\frac12}\Sigma + a_8 \mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_5^{(m)} + \mathcal{Q}_5^{(m)}\,, \tag{11.8}$$

where

$$\mathcal{P}_5^{(m)} = \begin{pmatrix} P_5^{(m)} & 0 \\ 0 & \overline{P}_5^{(m)} \end{pmatrix}\,, \quad \mathcal{Q}_5^{(m)} = \begin{pmatrix} 0 & Q_5^{(m)} \\ \overline{Q}_5^{(m)} & 0 \end{pmatrix}\,,$$
$$P_5^{(m)} = \mathrm{Op}(p_m) \in OPS^{-\frac12}\,, \quad Q_5^{(m)} = \mathrm{Op}(q_m) \in OPS^{-\frac{m}{2}}\,. \tag{11.9}$$

For any $\alpha \in \mathbb{N}$, assuming (7.7) with $\mu_0 \geq \aleph_4(m,\alpha) + \sigma$, where the increasing constants $\aleph_4(m,\alpha)$ are defined inductively by

$$\aleph_4(0,\alpha) := \alpha\,, \qquad \aleph_4(m+1,\alpha) := \aleph_4(m,\alpha+1) + \frac{m}{2} + 2\alpha + 4\,, \tag{11.10}$$

we have

$$|\mathcal{P}_5^{(m)}|_{-\frac12,s,\alpha}^{k_0,\gamma}\,,\ |\mathcal{Q}_5^{(m)}|_{-\frac{m}{2},s,\alpha} \lesssim_{m,s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_4(m,\alpha)+\sigma}^{k_0,\gamma}\big)\,, \tag{11.11}$$

$$|\Delta_{12}\mathcal{P}_5^{(m)}|_{-\frac12,s_1,\alpha}\,,\ |\Delta_{12}\mathcal{Q}_5^{(m)}|_{-\frac{m}{2},s_1,\alpha} \lesssim_{m,s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\aleph_4(m,\alpha)+\sigma}\,. \tag{11.12}$$

For $m \geq 1$, there exist real, even, reversibility preserving, invertible maps Φ_{m-1} of the form

$$\Phi_{m-1} := \mathbb{I}_2 + \Psi_{m-1}\,, \qquad \Psi_{m-1} := \begin{pmatrix} 0 & \psi_{m-1}(\varphi,x,D) \\ \overline{\psi_{m-1}(\varphi,x,D)} & 0 \end{pmatrix}\,, \tag{11.13}$$

with $\psi_{m-1}(\varphi,x,D)$ in $OPS^{-\frac{m-1}{2}-\frac12}$, such that

$$\mathcal{L}_5^{(m)} = \Phi_{m-1}^{-1}\mathcal{L}_5^{(m-1)}\Phi_{m-1}\,. \tag{11.14}$$

Initialization. The real, even and reversible operator $\mathcal{L}_5^{(0)} = \mathcal{L}_5$ in (11.1) satisfies the assumptions (11.8)-(11.12) for $m = 0$ by (11.6)-(11.7).

Inductive step. We conjugate $\mathcal{L}_5^{(m)}$ in (11.8) by a real operator of the form (see (11.13))

$$\Phi_m := \mathbb{I}_2 + \Psi_m\,, \qquad \Psi_m := \begin{pmatrix} 0 & \psi_m(\varphi,x,D) \\ \overline{\psi_m(\varphi,x,D)} & 0 \end{pmatrix}\,,$$
$$\psi_m(\varphi,x,D) := \mathrm{Op}(\psi_m) \in OPS^{-\frac{m}{2}-\frac12}\,. \tag{11.15}$$

We compute

$$\begin{aligned} \mathcal{L}_5^{(m)}\Phi_m &= \Phi_m\big(\omega\cdot\partial_\varphi + \mathrm{i} a_7|D|^{\frac12}T_{\mathtt h}^{\frac12}\Sigma + a_8\mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_5^{(m)}\big) \\ &\quad + \big[\mathrm{i} a_7|D|^{\frac12}T_{\mathtt h}^{\frac12}\Sigma + a_8\mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_5^{(m)}, \Psi_m\big] + (\omega\cdot\partial_\varphi\Psi_m) + \mathcal{Q}_5^{(m)} + \mathcal{Q}_5^{(m)}\Psi_m\,. \end{aligned} \tag{11.16}$$

In the next lemma we choose Ψ_m to decrease the order of the off-diagonal operator $\mathcal{Q}_5^{(m)}$.

Lemma 11.1. *Let*

$$\psi_m(\varphi,x,\xi) := \begin{cases} -\dfrac{\chi(\xi)q_m(\varphi,x,\xi)}{2\mathrm{i}a_7(\varphi,x)|\xi|^{\frac12}\tanh^{\frac12}(\mathtt{h}|\xi|)} & \text{if} \quad |\xi|>\frac13\,, \\ 0 & \text{if} \quad |\xi|\le\frac13\,, \end{cases} \qquad \psi_m\in S^{-\frac{m}{2}-\frac12}\,, \tag{11.17}$$

where the cut-off function χ *is defined in (2.16). Then the operator* Ψ_m *in (11.15) solves*

$$\mathrm{i}\big[a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma,\Psi_m\big]+\mathcal{Q}_5^{(m)}=\mathcal{Q}_{\psi_m} \tag{11.18}$$

where

$$\mathcal{Q}_{\psi_m} := \begin{pmatrix} 0 & q_{\psi_m}(\varphi,x,D) \\ \overline{q_{\psi_m}(\varphi,x,D)} & \end{pmatrix}, \quad q_{\psi_m}\in S^{-\frac{m}{2}-1}\,. \tag{11.19}$$

Moreover, there exists $\sigma(k_0,\tau,\nu)>0$ *such that, for any* $\alpha>0$, *if (7.7) holds with* $\mu_0\ge\aleph_4(m,\alpha+1)+\alpha+\frac{m}{2}+\sigma+4$, *then*

$$|q_{\psi_m}(\varphi,x,D)|^{k_0,\gamma}_{-\frac{m}{2}-1,s,\alpha}\lesssim_{s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\aleph_4(m,\alpha+1)+\frac{m}{2}+\alpha+\sigma+4}\big)\,. \tag{11.20}$$

The map Ψ_m *is real, even, reversibility preserving and*

$$|\psi_m(\varphi,x,D)|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha)}\big)\,, \tag{11.21}$$

$$|\Delta_{12}\psi_m(\varphi,x,D)|_{-\frac{m}{2}-\frac12,s_1,\alpha}\lesssim_{m,s_1,\alpha}\varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma+\aleph_4(m,\alpha)}\,, \tag{11.22}$$

$$|\Delta_{12}q_{\psi_m}(\varphi,x,D)|_{-\frac{m}{2}-1,s_1,\alpha}\lesssim_{m,s_1,\alpha}\varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\aleph_4(m,\alpha+1)+\frac{m}{2}+\alpha+\sigma+4}\,. \tag{11.23}$$

Proof. We first note that in (11.17) the denominator $a_7|\xi|^{\frac12}\tanh(\mathtt{h}|\xi|)^{\frac12}\ge c|\xi|^{\frac12}$ with $c>0$ for all $|\xi|\ge 1/3$, since $a_7-1=O(\varepsilon\gamma^{-1})$ by (10.17) and (7.7). Thus the symbol ψ_m is well defined and, by (11.17), (2.47) and (11.11), (10.17), Lemma 2.6, (7.7) we have, for all $s\ge s_0$,

$$\begin{aligned}|\psi_m(\varphi,x,D)|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}&\lesssim_{m,s,\alpha}|\mathcal{Q}_5^{(m)}|^{k_0,\gamma}_{-\frac{m}{2},s,\alpha}+\|a_7-1\|^{k_0,\gamma}_s|\mathcal{Q}_5^{(m)}|^{k_0,\gamma}_{-\frac{m}{2},s_0,\alpha}\\&\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha)}\big)\end{aligned}$$

proving (11.21).

Recalling the definition (11.2) of Σ, the vector valued commutator $\mathrm{i}[a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma,\Psi_m]$ is given by

$$\mathrm{i}\big[a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma,\Psi_m\big]=\begin{pmatrix}0 & A\\ \bar{A} & 0\end{pmatrix}, \qquad A:=\mathrm{i}\big(a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\mathrm{Op}(\psi_m)+\mathrm{Op}(\psi_m)a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\big)\,. \tag{11.24}$$

By (11.24), in order to solve (11.18) with a remainder $\mathcal{Q}_{\psi_m}\in OPS^{-\frac{m}{2}-1}$ as in (11.19), we have to solve the equation

$$\mathrm{i}a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\mathrm{Op}(\psi_m)+\mathrm{i}\mathrm{Op}(\psi_m)a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}+\mathrm{Op}(q_m)=\mathrm{Op}(q_{\psi_m})\in OPS^{-\frac{m}{2}-1}\,. \tag{11.25}$$

By (2.30), applied with $N=1$, $A=a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}$, $B=\psi_m(\varphi,x,D)$, and (2.33), we have the expansion

$$\begin{aligned}&a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\psi_m(\varphi,x,D)+\psi_m(\varphi,x,D)a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\\&=\mathrm{Op}\Big(2a_7|\xi|^{\frac12}\tanh^{\frac12}(\mathtt{h}|\xi|)\psi_m(\varphi,x,\xi)\Big)+\mathrm{Op}(\mathfrak{q}_{\psi_m}(\varphi,x,\xi))\end{aligned} \tag{11.26}$$

where, using that $a_7\chi(\xi)|\xi|^{\frac12}\tanh^{\frac12}(\mathtt{h}\chi(\xi)|\xi|)\in S^{\frac12}$ and $\psi_m(\varphi,x,\xi)\in S^{-\frac{m}{2}-\frac12}$, the symbol

$$\mathfrak{q}_{\psi_m}=r_{1,AB}+r_{1,BA}+2a_7|\xi|^{\frac12}\big(\tanh^{\frac12}(\mathtt{h}\chi(\xi)|\xi|)\chi(\xi)-\tanh^{\frac12}(\mathtt{h}|\xi|)\big)\psi_m(\varphi,x,\xi)\in S^{-\frac{m}{2}-1}\,, \tag{11.27}$$

recalling that $1-\chi(\xi)\in S^{-\infty}$ by (2.16). The symbol $\psi_m(\varphi,x,\xi)$ in (11.17) is the solution of

$$2\mathrm{i}a_7|\xi|^{\frac12}\tanh^{\frac12}(\mathtt{h}|\xi|)\psi_m(\varphi,x,\xi)+\chi(\xi)q_m(\varphi,x,\xi)=0\,, \tag{11.28}$$

and therefore, by (11.26)-(11.28), the remainder $q_{\psi_m}(\varphi,x,\xi)$ in (11.25) is

$$q_{\psi_m}(\varphi,x,\xi)=\mathrm{i}\mathfrak{q}_{\psi_m}(\varphi,x,\xi)+(1-\chi(\xi))q_m(\varphi,x,\xi)\in S^{-\frac{m}{2}-1}\,. \tag{11.29}$$

This proves (11.18)-(11.19).

We now prove (11.20). We first estimate (11.27). By (2.46) (applied with $N=1$, $A=a_7|D|^{\frac12}T_{\mathtt h}^{\frac12}$, $B=\psi_m(\varphi,x,D)$, $m=1/2$, $m'=-\frac{m}{2}-\frac12$ and also by inverting the role of A and B), and the estimates (11.21), (11.4), (7.7) we have

$$|\mathfrak{q}_{\psi_m}(\varphi,x,D)|^{k_0,\gamma}_{-\frac{m}{2}-1,s,\alpha}\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha+1)+\frac{m}{2}+\alpha+4}\big) \tag{11.30}$$

and the estimate (11.20) for $q_{\psi_m}(\varphi,x,D)$ follows by (11.29) using (11.11), recalling that $1-\chi(\xi)\in S^{-\infty}$ and by applying (2.47) with $g(D)=1-\chi(D)$ and $A=q_m(\varphi,x,D)$. Bounds (11.22)-(11.23) follow by similar arguments and by a repeated use of the triangular inequality.

Finally, the map Ψ_m defined by (11.15), (11.17) is real, even and reversibility preserving because $\mathcal{Q}_5^{(m)}$ is real, even, reversible and a_7 is even(φ)even(x) (see (10.16)). □

For $\varepsilon\gamma^{-1}$ small enough, by (11.21) and (7.7) the operator Φ_m is invertible, and, by Lemma 2.14,

$$|\Phi_m^{-1}-\mathbb{I}_2|^{k_0,\gamma}_{0,s,\alpha}\lesssim_{s,\alpha}|\Psi_m|^{k_0,\gamma}_{0,s,\alpha}\lesssim_{s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha)}\big)\,. \tag{11.31}$$

By (11.16) and (11.18), the conjugated operator is

$$\mathcal{L}_5^{(m+1)}:=\Phi_m^{-1}\mathcal{L}_5^{(m)}\Phi_m=\omega\cdot\partial_\varphi+\mathrm{i}a_7|D|^{\frac12}T_{\mathtt h}^{\frac12}\Sigma+a_8\mathbb{I}_2+\mathrm{i}\Pi_0+\mathcal{P}_5^{(m)}+\breve{\mathcal{P}}_{m+1} \tag{11.32}$$

where $\breve{\mathcal{P}}_{m+1}:=\Phi_m^{-1}\mathcal{P}^*_{m+1}$ and

$$\mathcal{P}^*_{m+1}:=\mathcal{Q}_{\psi_m}+\big[\mathrm{i}\Pi_0,\Psi_m\big]+\big[a_8\mathbb{I}_2+\mathcal{P}_5^{(m)},\Psi_m\big]+(\omega\cdot\partial_\varphi\Psi_m)+\mathcal{Q}_5^{(m)}\Psi_m\,. \tag{11.33}$$

Thus (11.14) at order $m+1$ is proved. Note that $\breve{\mathcal{P}}_{m+1}$ and Π_0 are the only operators in (11.32) containing off-diagonal terms.

Lemma 11.2. *The operator $\breve{\mathcal{P}}_{m+1}\in OPS^{-\frac{m}{2}-\frac12}$. Furthermore, for any $\alpha>0$, assuming (7.7) with $\mu_0\geq\sigma+\aleph_4(m+1,\alpha)$, the following estimates hold:*

$$|\breve{\mathcal{P}}_{m+1}|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m+1,\alpha)}\big)\,,\quad\forall s\geq s_0\,, \tag{11.34}$$

$$|\Delta_{12}\breve{\mathcal{P}}_{m+1}|_{-\frac{m}{2}-\frac12,s_1,\alpha}\lesssim_{m,s_1,\alpha}\varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma+\aleph_4(m+1,\alpha)} \tag{11.35}$$

where the constant $\aleph_4(m+1,\alpha)$ is defined in (11.10).

Proof. We prove the estimate (11.34). The operator $\mathcal{Q}_{\psi_m}$ defined in (11.19) is in $OPS^{-\frac{m}{2}-1}\subset OPS^{-\frac{m}{2}-\frac12}$ and satisfies (11.20). The operator $[\Pi_0,\Psi_m]\in OPS^{-\infty}$ satisfies, by (11.21),

$$|[\Pi_0,\Psi_m]|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha)}\big)\,.$$

Recalling (11.9), (11.15), we have

$$\big[a_8\mathbb{I}_2+\mathcal{P}_5^{(m)},\Psi_m\big]=\begin{pmatrix}0&A\\ \overline{A}&0\end{pmatrix}\,,\quad A:=\big(a_8+P_5^{(m)}\big)\mathrm{Op}(\psi_m)-\mathrm{Op}(\psi_m)\big(a_8+\overline{P_5^{(m)}}\big)\,,$$

and (2.45), (11.5), (11.11), (11.21) imply

$$|\big[a_8\mathbb{I}_2+\mathcal{P}_5^{(m)},\Psi_m\big]|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha)+\frac{m}{2}+\frac12+\alpha}\big)\,.$$

The operator $(\omega\cdot\partial_\varphi\Psi_m)\in OPS^{-\frac{m}{2}-\frac12}$ satisfies

$$|\omega\cdot\partial_\varphi\Psi_m|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}\lesssim|\Psi_m|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s+1,\alpha}\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha)+1}\big)$$

by (11.21).

Finally $\mathcal{Q}_5^{(m)}\Psi_m\in OPS^{-m-\frac12}\subset OPS^{-\frac{m}{2}-\frac12}$ and by (2.40) and (2.45) (applied with $A=\mathcal{Q}_5^{(m)}$, $B=\Psi_m$, $(-\frac{m}{2},-\frac{m}{2}-\frac12)$ instead of (m,m')), (11.11), (11.21) we get

$$|\mathcal{Q}_5^{(m)}\Psi_m|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}\le|\mathcal{Q}_5^{(m)}\Psi_m|^{k_0,\gamma}_{-m-\frac12,s,\alpha}\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha)+\alpha+\frac{m}{2}}\big)\,.$$

Collecting all the previous estimates we deduce that $\mathcal{P}^*_{m+1}$ defined in (11.33) is in $OPS^{-\frac{m}{2}-\frac12}$ (the highest order term is $\omega\cdot\partial_\varphi\Psi_m$) and

$$|\mathcal{P}^*_{m+1}|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha+1)+\frac{m}{2}+\alpha+4}\big)\,. \tag{11.36}$$

In conclusion (2.45) (applied with $m=0$, $m'=-\frac{m}{2}-\frac12$), (11.31), (11.36) imply

$$\begin{aligned}|\breve{\mathcal{P}}_{m+1}|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}&=|\Phi_m^{-1}\mathcal{P}^*_{m+1}|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha}\\&\lesssim_{m,s,\alpha}|\Phi_m^{-1}|^{k_0,\gamma}_{0,s,\alpha}|\mathcal{P}^*_{m+1}|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s_0+\alpha,\alpha}+|\Phi_m^{-1}|^{k_0,\gamma}_{0,s_0,\alpha}|\mathcal{P}^*_{m+1}|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s+\alpha,\alpha}\\&\lesssim_{m,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha+1)+\frac{m}{2}+2\alpha+4}\big)\end{aligned}$$

which is (11.34), recalling (11.10). Estimate (11.35) can be proved by similar arguments. □

The operator $\mathcal{L}_5^{(m+1)}$ in (11.32) has the same form (11.8) as $\mathcal{L}_5^{(m)}$ with diagonal operators $\mathcal{P}_5^{(m+1)}$ and off-diagonal operators $\mathcal{Q}_5^{(m+1)}$ like in (11.9), with

$$\mathcal{P}_5^{(m+1)}+\mathcal{Q}_5^{(m+1)}=\mathcal{P}_5^{(m)}+\breve{\mathcal{P}}_{m+1}\,,$$

satisfying (11.11)-(11.12) at the step $m+1$ thanks to (11.34)-(11.35) and (11.11)-(11.12) at the step m. This proves the inductive claim.

Applying it $2M$ times (the constant M will be fixed in (15.16)), we derive the following lemma.

Lemma 11.3. *For any $\alpha>0$, assuming (7.7) with $\mu_0\ge\aleph_5(M,\alpha)+\sigma$ where the constant $\aleph_5(M,\alpha):=\aleph_4(2M,\alpha)$ is defined recursively by (11.10), the following holds. The real, even, reversibility preserving, invertible map*

$$\mathbf{\Phi}_M:=\Phi_0\circ\ldots\circ\Phi_{2M-1} \tag{11.37}$$

where Φ_m, $m=0,\ldots,2M-1$, are defined in (11.15), satisfies

$$|\mathbf{\Phi}_M^{\pm1}-\mathbb{I}_2|^{k_0,\gamma}_{0,s,0}\,,\ |(\mathbf{\Phi}_M^{\pm1}-\mathbb{I}_2)^*|^{k_0,\gamma}_{0,s,0}\lesssim_{s,M}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_5(M,0)}\big)\,,\quad\forall s\ge s_0\,, \tag{11.38}$$

$$|\Delta_{12}\mathbf{\Phi}_M^{\pm1}|_{0,s_1,0}\,,\ |\Delta_{12}(\mathbf{\Phi}_M^{\pm1})^*|_{0,s_1,0}\lesssim_{M,s_1}\varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma+\aleph_5(M,0)}\,. \tag{11.39}$$

The map $\mathbf{\Phi}_M$ conjugates $\mathcal{L}_5$ to the real, even and reversible operator

$$\mathcal{L}_6:=\mathbf{\Phi}_M^{-1}\mathcal{L}_5\mathbf{\Phi}_M=\omega\cdot\partial_\varphi+\mathrm{i}a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma+a_8\mathbb{I}_2+\mathrm{i}\Pi_0+\mathcal{P}_6+\mathcal{Q}_6 \tag{11.40}$$

where the functions a_7,a_8 are defined in (10.14), (11.1), and

$$\mathcal{P}_6:=\begin{pmatrix}P_6&0\\0&\overline{P_6}\end{pmatrix}\in OPS^{-\frac12}\,,\quad\mathcal{Q}_6:=\begin{pmatrix}0&Q_6\\\overline{Q_6}&0\end{pmatrix}\in OPS^{-M} \tag{11.41}$$

given by $\mathcal{P}_6:=\mathcal{P}_5^{(2M)}$, $\mathcal{Q}_6:=\mathcal{Q}_5^{(2M)}$ in (11.8)-(11.9) for $m=2M$, satisfy

$$|\mathcal{P}_6|^{k_0,\gamma}_{-\frac12,s,\alpha}+|\mathcal{Q}_6|^{k_0,\gamma}_{-M,s,\alpha}\lesssim_{M,s,\alpha}\varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_5(M,\alpha)}\big)\,,\quad\forall s\ge s_0\,, \tag{11.42}$$

$$|\Delta_{12}\mathcal{P}_6|_{-\frac12,s_1,\alpha}+|\Delta_{12}\mathcal{Q}_6|_{-M,s_1,\alpha}\lesssim_{M,s_1,\alpha}\varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma+\aleph_5(M,\alpha)}\,. \tag{11.43}$$

Proof. Estimates (11.42)-(11.43) are (11.11)-(11.12) at $m = 2M$. Let us prove (11.38). For all $0 \leq m \leq 2M - 1$, $s \geq s_0$, we have

$$|\!|\!| \Phi_m - \mathbb{I}_2 |\!|\!|_{0,s,0}^{k_0,\gamma} \overset{(11.15)}{=} |\!|\!| \Psi_m |\!|\!|_{0,s,0}^{k_0,\gamma} \overset{(11.21)}{\lesssim_s} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_4(m,0)}^{k_0,\gamma}\big) \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_4(2M,0)}^{k_0,\gamma}\big) ,$$

then the estimate (11.38) for $\mathbf{\Phi}_M - \mathbb{I}_2$ follows by applying iteratively the estimate (2.45) of Lemma 2.10. The estimate for $\mathbf{\Phi}_M^{-1} - \mathbb{I}_2$ follows since by (11.37), $\mathbf{\Phi}_M^{-1} = \Phi_{2M-1}^{-1} \circ \ldots \circ \Phi_0^{-1}$, using (11.31) and applying iteratively the estimate (2.45) of Lemma 2.10. The estimate on the adjoint operator $(\mathbf{\Phi}_M^{\pm 1} - \mathbb{I}_2)^*$ follows similarly since Lemma 2.13 implies $|\!|\!| (\Phi_m^{\pm 1} - \mathbb{I}_2)^* |\!|\!|_{0,s,0}^{k_0,\gamma} \lesssim_M |\!|\!| \Phi_m^{\pm 1} - \mathbb{I}_2 |\!|\!|_{0,s+s_0,0}^{k_0,\gamma}$. Also (11.39) is proved analogously. □

12 Reduction of the highest order

We have obtained the operator $\mathcal{L}_6$ in (11.40), where $\mathcal{P}_6$ is in $OPS^{-\frac{1}{2}}$ and the off-diagonal term $\mathcal{Q}_6$ is in OPS^{-M}. The goal of this section is to reduce to *constant coefficient* the leading term $\mathrm{i}a_7(\varphi,x)|D|^{\frac{1}{2}}T_{\mathtt{h}}^{\frac{1}{2}}\Sigma$. To this end, we study how the operator $\mathcal{L}_6$ transforms under the action of the flow of the φ-dependent family of pseudo-PDEs

$$\begin{cases} \partial_\tau u = \mathrm{i}\beta(\varphi,x)|D|^{\frac{1}{2}} u \\ u(0,\varphi,x) = u_0(\varphi,x) \end{cases} \tag{12.1}$$

where $\beta(\varphi,x)$ is a real valued smooth function, which will be defined in (12.21). We also write

$$A := A(\varphi) := \beta(\varphi,x)|D|^{\frac{1}{2}} . \tag{12.2}$$

Let $\Phi(\tau) := \Phi(\tau,\varphi)$ denote the time τ flow of equation (12.1), namely

$$\begin{cases} \partial_\tau \Phi(\tau) = \mathrm{i}\beta(\varphi,x)|D|^{\frac{1}{2}}\Phi(\tau) \\ \Phi(0) = \mathrm{Id} . \end{cases} \tag{12.3}$$

Since the function $\beta(\varphi,x)$ is real valued, usual energy estimates imply that the flow $\Phi(\tau,\varphi)$ is a bounded operator mapping H_x^s to H_x^s. In the Appendix of [21] it is proved that the flow $\Phi(\tau,\varphi)$ also satisfies tame estimates in $H_{\varphi,x}^s$ as well as $\partial_\varphi^r \partial_\lambda^k \Phi(\tau,\varphi)$ with losses of $(|r|+|k|)/2$ derivatives, see Section 2.7.

Let $\Phi := \Phi(\varphi) := \Phi(1,\varphi)$ be the time one flow of (12.1). Note that $\Phi^{-1} = \overline{\Phi}$ (see Section 2.7) and

$$\Phi\pi_0 = \pi_0 = \Phi^{-1}\pi_0 \tag{12.4}$$

because, when the datum $u_0(\varphi,x) = u_0(\varphi)$ in (12.1) does not depend on x, the solution of (12.1) is $u(\tau,\varphi,x) = u_0(\varphi)$ for all τ, x.

We write the operator $\mathcal{L}_6$ in (11.40) as

$$\mathcal{L}_6 = \omega\cdot\partial_\varphi + \mathrm{i}\Pi_0 + \begin{pmatrix} P_6^{(0)} & Q_6 \\ \overline{Q}_6 & \overline{P}_6^{(0)} \end{pmatrix}$$

where Π_0 is defined in (11.2), Q_6 in (11.41), and

$$P_6^{(0)} := P_6^{(0)}(\varphi,x,D) := \mathrm{i}a_7|D|^{\frac{1}{2}}T_{\mathtt{h}}^{\frac{1}{2}} + a_8 + P_6 \tag{12.5}$$

with P_6 defined in (11.41). Conjugating $\mathcal{L}_6$ with the real operator

$$\mathbf{\Phi} := \begin{pmatrix} \Phi & 0 \\ 0 & \overline{\Phi} \end{pmatrix} \tag{12.6}$$

we get, since $\mathbf{\Phi}^{-1}\Pi_0\mathbf{\Phi} = \Pi_0\mathbf{\Phi}$ by (12.4),

$$\mathcal{L}_7 := \mathbf{\Phi}^{-1}\mathcal{L}_6\mathbf{\Phi} = \omega\cdot\partial_\varphi + \mathbf{\Phi}^{-1}\big(\omega\cdot\partial_\varphi\mathbf{\Phi}\big) + \mathrm{i}\Pi_0\mathbf{\Phi} + \begin{pmatrix} \Phi^{-1}P_6^{(0)}\Phi & \Phi^{-1}Q_6\overline{\Phi} \\ \overline{\Phi}^{-1}\overline{Q}_6\Phi & \overline{\Phi}^{-1}\overline{P}_6^{(0)}\overline{\Phi} \end{pmatrix} . \tag{12.7}$$

Let us study the operator

$$L_7 := \omega \cdot \partial_\varphi + \Phi^{-1}\big(\omega \cdot \partial_\varphi \Phi\big) + \Phi^{-1} P_6^{(0)} \Phi \,. \tag{12.8}$$

Analysis of the term $\Phi^{-1} P_6^{(0)} \Phi$. Recalling (12.3) and (12.2) the operator

$$P(\tau,\varphi) := \Phi(\tau,\varphi)^{-1} P_6^{(0)} \Phi(\tau,\varphi)$$

satisfies the equation

$$\partial_\tau P(\tau,\varphi) = -\mathrm{i}\Phi(\tau,\varphi)^{-1}\big[A(\varphi), P_6^{(0)}\big]\Phi(\tau,\varphi)\,.$$

Iterating this formula, and using the notation

$$\mathrm{Ad}_{A(\varphi)} P_6^{(0)} := \big[A(\varphi), P_6^{(0)}\big]\,,$$

we obtain the following Lie series expansion of the conjugated operator

$$\begin{aligned}
\Phi(1,\varphi)^{-1} P_6^{(0)} \Phi(1,\varphi) = P_6^{(0)} - \mathrm{i}[A, P_6^{(0)}] + \sum_{n=2}^{2M} \frac{(-\mathrm{i})^n}{n!} \mathrm{Ad}^n_{A(\varphi)} P_6^{(0)} \\
+ \frac{(-\mathrm{i})^{2M+1}}{(2M)!} \int_0^1 (1-\tau)^{2M} \Phi(\tau,\varphi)^{-1} \mathrm{Ad}^{2M+1}_{A(\varphi)} P_6^{(0)} \Phi(\tau,\varphi)\, d\tau \,.
\end{aligned} \tag{12.9}$$

The order M of the expansion will be fixed in (15.16). We remark that (12.9) is an expansion in operators with decreasing orders (and size) because each commutator with $A(\varphi) = \beta(\varphi,x)|D|^{\frac12}$ gains $\frac12$ order (and it has the size of β). By (12.2) and (12.5),

$$-\mathrm{i}\big[A, P_6^{(0)}\big] = \big[\beta|D|^{\frac12}, a_7|D|^{\frac12}\big] + \big[\beta|D|^{\frac12}, a_7|D|^{\frac12}(T_{\mathtt{h}}^{\frac12} - \mathrm{Id})\big] - \mathrm{i}\big[\beta|D|^{\frac12}, a_8 + P_6\big]\,. \tag{12.10}$$

Moreover, by (2.48), (2.49) one has

$$\begin{aligned}
[\beta|D|^{\frac12}, a_7|D|^{\frac12}] &= \mathrm{Op}\Big(- \mathrm{i}\{\beta\chi(\xi)|\xi|^{\frac12}, a_7\chi(\xi)|\xi|^{\frac12}\} + \mathtt{r}_2(\beta\chi(\xi)|\xi|^{\frac12}, a_7\chi(\xi)|\xi|^{\frac12})\Big) \\
&= \mathrm{i}\mathrm{Op}\Big(\big((\partial_x\beta)a_7 - \beta(\partial_x a_7)\big)\chi(\xi)|\xi|^{\frac12}\partial_\xi\big(\chi(\xi)|\xi|^{\frac12}\big)\Big) + \mathrm{Op}\big(\mathtt{r}_2(\beta\chi(\xi)|\xi|^{\frac12}, a_7\chi(\xi)|\xi|^{\frac12})\big) \\
&= \mathrm{i}\big((\partial_x\beta)a_7 - \beta(\partial_x a_7)\big)\mathrm{Op}\Big(\frac12\chi^2(\xi)\mathrm{sign}(\xi) + \chi(\xi)\partial_\xi\chi(\xi)|\xi|\Big) + \mathrm{Op}\big(\mathtt{r}_2(\beta\chi(\xi)|\xi|^{\frac12}, a_7\chi(\xi)|\xi|^{\frac12})\big)
\end{aligned} \tag{12.11}$$

where the symbol $\mathtt{r}_2(\beta\chi(\xi)|\xi|^{\frac12}, a_7\chi(\xi)|\xi|^{\frac12}) \in S^{-1}$ is defined according to (2.50). Therefore (12.10), (12.11) imply the expansion

$$-\mathrm{i}\big[A, P_6^{(0)}\big] = -\frac12\big((\partial_x\beta)a_7 - \beta(\partial_x a_7)\big)\mathcal{H} + R_{A,P_6^{(0)}} \tag{12.12}$$

where the remainder

$$\begin{aligned}
R_{A,P_6^{(0)}} &:= \mathrm{i}\big((\partial_x\beta)a_7 - \beta(\partial_x a_7)\big)\mathrm{Op}\,\Big(\chi(\xi)\partial_\xi\chi(\xi)|\xi| + \frac12(\chi^2(\xi) - \chi(\xi))\mathrm{sign}(\xi)\Big) \\
&\quad + \mathrm{Op}\big(\mathtt{r}_2(\beta\chi(\xi)|\xi|^{\frac12}, a_7\chi(\xi)|\xi|^{\frac12})\big) + \big[\beta|D|^{\frac12}, a_7|D|^{\frac12}(T_{\mathtt{h}}^{\frac12} - \mathrm{Id})\big] - \mathrm{i}\big[\beta|D|^{\frac12}, a_8 + P_6\big]
\end{aligned} \tag{12.13}$$

is an operator of order $-\frac12$ (because of the term $[\beta|D|^{\frac12}, a_8]$).

Analysis of the term $\omega \cdot \partial_\varphi + \Phi^{-1}\{\omega \cdot \partial_\varphi \Phi\} = \Phi^{-1} \circ \omega \cdot \partial_\varphi \circ \Phi$. We argue as above, differentiating

$$\begin{aligned}
\partial_\tau\big\{\Phi(\tau,\varphi)^{-1} \circ \omega \cdot \partial_\varphi \circ \Phi(\tau,\varphi)\big\} &= -\mathrm{i}\Phi(\tau,\varphi)^{-1}\big[A(\varphi), \omega\cdot\partial_\varphi\big]\Phi(\tau,\varphi) \\
&= -\mathrm{i}\Phi(\tau,\varphi)^{-1}\big(\mathrm{Ad}_{A(\varphi)}\omega\cdot\partial_\varphi\big)\Phi(\tau,\varphi)\,.
\end{aligned}$$

Therefore, by iteration, we get the Lie series expansion

$$\Phi(1,\varphi)^{-1}\circ\omega\cdot\partial_\varphi\circ\Phi(1,\varphi)=\omega\cdot\partial_\varphi-\mathrm{i}\mathrm{Ad}_{A(\varphi)}\omega\cdot\partial_\varphi+\frac{(-\mathrm{i})^2}{2}\mathrm{Ad}^2_{A(\varphi)}\omega\cdot\partial_\varphi+\sum_{n=3}^{2M+1}\frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}^n_{A(\varphi)}\omega\cdot\partial_\varphi$$
$$+\frac{(-\mathrm{i})^{2M+2}}{(2M+1)!}\int_0^1(1-\tau)^{2M+1}\Phi(\tau,\varphi)^{-1}\big(\mathrm{Ad}^{2M+2}_{A(\varphi)}\omega\cdot\partial_\varphi\big)\Phi(\tau,\varphi)\,d\tau. \tag{12.14}$$

We compute the commutator

$$\mathrm{Ad}_{A(\varphi)}\omega\cdot\partial_\varphi=\big[A(\varphi),\omega\cdot\partial_\varphi\big]=-(\omega\cdot\partial_\varphi A(\varphi))\overset{(12.2)}{=}-(\omega\cdot\partial_\varphi\beta(\varphi,x))|D|^{1/2} \tag{12.15}$$

and, using (2.48), (2.49),

$$\begin{aligned}\mathrm{Ad}^2_{A(\varphi)}\omega\cdot\partial_\varphi&=\big[(\omega\cdot\partial_\varphi A(\varphi)),A(\varphi)\big]=\big[(\omega\cdot\partial_\varphi\beta)|D|^{\frac12},\beta|D|^{\frac12}\big]\\&=\mathrm{Op}\Big(-\mathrm{i}\big\{(\omega\cdot\partial_\varphi\beta)\chi(\xi)|\xi|^{\frac12},\beta\chi(\xi)|\xi|^{\frac12}\big\}+\mathtt{r}_2((\omega\cdot\partial_\varphi\beta)\chi(\xi)|\xi|^{\frac12},\beta\chi(\xi)|\xi|^{\frac12})\Big).\end{aligned}$$

According to (2.49) the term with the Poisson bracket is

$$-\mathrm{i}\big\{(\omega\cdot\partial_\varphi\beta)\chi(\xi)|\xi|^{\frac12},\beta\chi(\xi)|\xi|^{\frac12}\big\}=\mathrm{i}\big(\beta\,\omega\cdot\partial_\varphi\beta_x-\beta_x\,\omega\cdot\partial_\varphi\beta\big)\Big(\frac12\chi(\xi)^2\mathrm{sign}(\xi)+\chi(\xi)\partial_\xi\chi(\xi)|\xi|\Big)$$

and therefore

$$\frac{(-\mathrm{i})^2}{2}\mathrm{Ad}^2_{A(\varphi)}\omega\cdot\partial_\varphi=\frac14\big(\beta\,\omega\cdot\partial_\varphi\beta_x-\beta_x\,\omega\cdot\partial_\varphi\beta\big)\mathcal{H}+R_{A,\omega\cdot\partial_\varphi} \tag{12.16}$$

where

$$\begin{aligned}R_{A,\omega\cdot\partial_\varphi}:=&-\frac{\mathrm{i}}{4}\big(\beta\,\omega\cdot\partial_\varphi\beta_x-\beta_x\,\omega\cdot\partial_\varphi\beta\big)\,\mathrm{Op}\Big((\chi(\xi)^2-\chi(\xi))\mathrm{sign}(\xi)+2\chi(\xi)\partial_\xi\chi(\xi)|\xi|\Big)\\&-\frac12\mathrm{Op}\Big(\mathtt{r}_2\big((\omega\cdot\partial_\varphi\beta)\chi(\xi)|\xi|^{\frac12},\beta\chi(\xi)|\xi|^{\frac12}\big)\Big).\end{aligned} \tag{12.17}$$

is an operator in OPS^{-1} (the first line of (12.17) reduces to the zero operator when acting on the periodic functions, because $\chi^2-\chi$ and $\partial_\xi\chi$ vanish on $\mathbb{Z}$).

Finally, by (12.14), (12.15) and (12.16), we get

$$\begin{aligned}\Phi(1,\varphi)^{-1}\circ\omega\cdot\partial_\varphi\circ\Phi(1,\varphi)=&\ \omega\cdot\partial_\varphi+\mathrm{i}(\omega\cdot\partial_\varphi\beta)(\varphi,x)|D|^{\frac12}+\frac14\big(\beta(\omega\cdot\partial_\varphi\beta_x)-\beta_x(\omega\cdot\partial_\varphi\beta)\big)\mathcal{H}+R_{A,\omega\cdot\partial_\varphi}\\&-\sum_{n=3}^{2M+1}\frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}^{n-1}_{A(\varphi)}\big(\omega\cdot\partial_\varphi A(\varphi)\big)\\&-\frac{(-\mathrm{i})^{2M+2}}{(2M+1)!}\int_0^1(1-\tau)^{2M+1}\Phi(\tau,\varphi)^{-1}\big(\mathrm{Ad}^{2M+1}_{A(\varphi)}\big(\omega\cdot\partial_\varphi A(\varphi)\big)\big)\Phi(\tau,\varphi)\,d\tau.\end{aligned} \tag{12.18}$$

This is an expansion in operators with decreasing orders (and size).

In conclusion, by (12.8), (12.9), (12.5), (12.12), (12.18), the term of order $|D|^{\frac12}$ in L_7 in (12.8) is given by

$$\mathrm{i}\big((\omega\cdot\partial_\varphi\beta)+a_7T_{\mathtt{h}}^{\frac12}\big)|D|^{\frac12}. \tag{12.19}$$

Choice of the functions $\beta(\varphi,x)$ and $\alpha(x)$. We choose the function $\beta(\varphi,x)$ such that

$$(\omega\cdot\partial_\varphi\beta)(\varphi,x)+a_7(\varphi,x)=\langle a_7\rangle_\varphi(x),\quad\langle a_7\rangle_\varphi(x):=\frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu}a_7(\varphi,x)\,d\varphi. \tag{12.20}$$

For all $\omega\in\mathtt{DC}(\gamma,\tau)$, the solution of (12.20) is the periodic function

$$\beta(\varphi,x):=-(\omega\cdot\partial_\varphi)^{-1}\big(a_7(\varphi,x)-\langle a_7\rangle_\varphi(x)\big), \tag{12.21}$$

which we extend to the whole parameter space $\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$ by setting $\beta_{ext}:=-(\omega\cdot\partial_\varphi)^{-1}_{ext}(a_7-\langle a_7\rangle_\varphi)$ via the operator $(\omega\cdot\partial_\varphi)^{-1}_{ext}$ defined in Lemma 2.5. For simplicity we still denote by β this extension.

Lemma 12.1. *The real valued function* β *defined in* (12.21) *is odd*(φ)*even*(x). *Moreover there exists* $\sigma(k_0, \tau, \nu) > 0$ *such that, if* (7.7) *holds with* $\mu_0 \geq \sigma$, *then* β *satisfies the following estimates:*

$$\|\beta\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-2}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big), \quad \|\omega \cdot \partial_\varphi \beta\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) \tag{12.22}$$

$$\|\Delta_{12}\beta\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-2}\|\Delta_{12} i\|_{s_1+\sigma}, \quad \|\omega \cdot \partial_\varphi \Delta_{12}\beta\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}. \tag{12.23}$$

Proof. The function a_7 is even(φ)even(x) (see (10.16)), and therefore, by (12.21), β is odd(φ)even(x). Estimates (12.22)-(12.23) follow by (12.20), (12.21), (11.4) and Lemma 2.5. □

By (10.14), (9.7), (9.8) one has

$$a_7 = \sqrt{a_5 a_6} = \sqrt{\mathcal{A}^{-1}(a_2)\mathcal{A}^{-1}(a_3)\mathcal{A}^{-1}(1+\alpha_x)} = \mathcal{A}^{-1}(\sqrt{a_2 a_3})\mathcal{A}^{-1}\big(\sqrt{1+\alpha_x}\big).$$

We now choose the 2π-periodic function $\alpha(x)$ (introduced as a free parameter in (9.1)) so that

$$\langle a_7 \rangle_\varphi(x) = \mathtt{m}_{\frac12} \tag{12.24}$$

is independent of x, for some real constant $\mathtt{m}_{\frac12}$. This is equivalent to solve the equation

$$\langle \sqrt{a_2 a_3}\,\rangle_\varphi(x)\,\sqrt{1+\alpha_x(x)} = \mathtt{m}_{\frac12}$$

whose solution is

$$\mathtt{m}_{\frac12} := \Big(\frac{1}{2\pi}\int_{\mathbb{T}} \frac{dx}{\langle \sqrt{a_2 a_3}\,\rangle_\varphi^2(x)}\Big)^{-\frac12}, \qquad \alpha(x) := \partial_x^{-1}\Big(\frac{\mathtt{m}_{\frac12}^2}{\langle \sqrt{a_2 a_3}\,\rangle_\varphi^2(x)} - 1\Big). \tag{12.25}$$

Lemma 12.2. *The real valued function* $\alpha(x)$ *defined in* (12.25) *is odd*(x) *and* (9.2) *holds. Moreover*

$$|\mathtt{m}_{\frac12} - 1|^{k_0,\gamma} \lesssim \varepsilon\gamma^{-1}, \quad |\Delta_{12}\mathtt{m}_{\frac12}| \lesssim \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1}. \tag{12.26}$$

Proof. Since a_2, a_3 are even(x) by (8.35), the function $\alpha(x)$ defined in (12.25) is odd(x). Estimates (12.26) follow by the definition of $\mathtt{m}_{\frac12}$ in (12.25) and (8.36), (8.38), (7.7), applying also Lemma 2.6 and (2.10). Similarly α satisfies (9.2) by (8.36), (8.38), (12.26), Lemma 2.6 and (2.10). □

By (12.20) and (12.24) the term in (12.19) reduces to

$$\mathrm{i}\big(\omega \cdot \partial_\varphi \beta(\varphi,x) + a_7(\varphi,x) T_{\mathtt{h}}^{\frac12}\big)|D|^{\frac12} = \mathrm{i}\mathtt{m}_{\frac12} T_{\mathtt{h}}^{\frac12}|D|^{\frac12} + \mathtt{R}_\beta \tag{12.27}$$

where $\mathtt{R}_\beta$ is the $OPS^{-\infty}$ operator defined by

$$\mathtt{R}_\beta := \mathrm{i}(\omega \cdot \partial_\varphi \beta)(\mathrm{Id} - T_{\mathtt{h}}^{\frac12})|D|^{\frac12}. \tag{12.28}$$

Finally, the operator L_7 in (12.8) is, in view of (12.9), (12.5), (12.12), (12.18), (12.27),

$$L_7 = \omega \cdot \partial_\varphi + \mathrm{i}\mathtt{m}_{\frac12} T_{\mathtt{h}}^{\frac12}|D|^{\frac12} + a_8 + a_9\mathcal{H} + P_7 + T_7 \tag{12.29}$$

where a_9 is the real valued function

$$a_9 := a_9(\varphi,x) := -\frac12\big(\beta_x\, a_7 - \beta(\partial_x a_7)\big) - \frac14\big(\beta_x\,\omega \cdot \partial_\varphi \beta - \beta\,\omega \cdot \partial_\varphi \beta_x\big), \tag{12.30}$$

P_7 is the operator in $OPS^{-1/2}$ given by

$$P_7 := R_{A,P_6^{(0)}} + R_{A,\omega\cdot\partial_\varphi} - \sum_{n=3}^{2M+1} \frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}_{A(\varphi)}^{n-1}\big(\omega \cdot \partial_\varphi A(\varphi)\big) + \sum_{n=2}^{2M} \frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}_{A(\varphi)}^{n} P_6^{(0)} + P_6 + \mathtt{R}_\beta \tag{12.31}$$

(the operators $R_{A,P_6^{(0)}}, R_{A,\omega\cdot\partial_\varphi}, P_6, \mathtt{R}_\beta$ are defined respectively in (12.13), (12.17), (11.41), (12.28)), and

$$
\begin{aligned}
T_7 := & -\frac{(-\mathrm{i})^{2M+2}}{(2M+1)!}\int_0^1 (1-\tau)^{2M+1}\Phi(\tau,\varphi)^{-1}\big(\mathrm{Ad}_{A(\varphi)}^{2M+1}\big(\omega\cdot\partial_\varphi A(\varphi)\big)\big)\Phi(\tau,\varphi)\,d\tau \\
& + \frac{(-\mathrm{i})^{2M+1}}{(2M)!}\int_0^1 (1-\tau)^{2M}\Phi(\tau,\varphi)^{-1}\mathrm{Ad}_{A(\varphi)}^{2M+1}P_6^{(0)}\Phi(\tau,\varphi)\,d\tau
\end{aligned}
\tag{12.32}
$$

(T_7 stands for "tame remainders", namely remainders satisfying tame estimates together with their derivatives, see (12.41), without controlling their pseudo-differential structure). In conclusion, we have the following lemma.

Lemma 12.3. *Let $\beta(\varphi,x)$ and $\alpha(x)$ be the functions defined in* (12.21) *and* (12.25)*. Then $\mathcal{L}_7 := \mathbf{\Phi}^{-1}\mathcal{L}_6\mathbf{\Phi}$ in* (12.7) *is the real, even and reversible operator*

$$
\mathcal{L}_7 = \omega\cdot\partial_\varphi + \mathrm{i}\mathtt{m}_{\frac12}T_{\mathtt{h}}^{\frac12}|D|^{\frac12}\Sigma + \mathrm{i}\Pi_0 + (a_8 + a_9\mathcal{H})\mathbb{I}_2 + \mathcal{P}_7 + \mathcal{T}_7 \tag{12.33}
$$

where $\mathtt{m}_{\frac12}$ is the real constant defined in (12.25)*, a_8, a_9 are the real valued functions in* (11.1), (12.30),

$$
a_8 = odd(\varphi)even(x)\,, \qquad a_9 = odd(\varphi)odd(x)\,, \tag{12.34}
$$

and $\mathcal{P}_7$, $\mathcal{T}_7$ are the real operators

$$
\mathcal{P}_7 := \begin{pmatrix} P_7 & 0 \\ 0 & \overline{P_7}\end{pmatrix} \in OPS^{-\frac12}\,, \qquad \mathcal{T}_7 := \mathrm{i}\Pi_0(\mathbf{\Phi}-\mathbb{I}_2) + \mathbf{\Phi}^{-1}\mathcal{Q}_6\mathbf{\Phi} + \begin{pmatrix} T_7 & 0 \\ 0 & \overline{T_7}\end{pmatrix}\,, \tag{12.35}
$$

where P_7 is defined in (12.31) *and T_7 in* (12.32)*.*

Proof. Formula (12.33) follows by (12.7) and (12.29). By Lemma 12.1 the real function β is odd(φ)even(x). Thus, by Sections 2.5 and 2.7, the flow map $\mathbf{\Phi}$ in (12.6) is real, even and reversibility preserving and therefore the conjugated operator $\mathcal{L}_7$ is real, even and reversible. Moreover the function a_7 is even(φ)even(x) by (10.16) and a_9 defined in (12.30) is odd(φ)odd(x). ☐

Note that formulas (12.30) and (12.35) (via (12.31), (12.32)) define a_9 and $\mathcal{P}_7, \mathcal{T}_7$ on the whole parameter space $\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$ by means of the extended function β and the corresponding flow Φ. Thus the right hand side of (12.33) defines an extended operator on $\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$, which we still denote by $\mathcal{L}_7$.

In the next lemma we provide some estimates on the operators $\mathcal{P}_7$ and $\mathcal{T}_7$.

Lemma 12.4. *There exists $\sigma(k_0,\tau,\nu)>0$ such that, if* (7.7) *holds with $\mu_0\geq\sigma$, then*

$$
\|a_9\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-2}(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})\,, \quad \forall s\geq s_0\,, \qquad \|\Delta_{12}a_9\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-2}\|\Delta_{12}i\|_{s_1+\sigma}\,. \tag{12.36}
$$

For any $s\geq s_0$ there exists $\delta(s)>0$ small enough such that if $\varepsilon\gamma^{-2}\leq\delta(s)$, then

$$
\|(\Phi^{\pm1}-\mathrm{Id})h\|_s^{k_0,\gamma}, \|(\Phi^*-\mathrm{Id})h\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-2}\big(\|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,, \tag{12.37}
$$

$$
\|\Delta_{12}\Phi^{\pm1}h\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-2}\|\Delta_{12}i\|_{s_1+\sigma}\|h\|_{s_1+\frac12}\,. \tag{12.38}
$$

The pseudo-differential operator $\mathcal{P}_7$ defined in (12.35) *is in $OPS^{-\frac12}$. Moreover for any $M,\alpha>0$, there exists a constant $\aleph_6(M,\alpha)>0$ such that assuming* (7.7) *with $\mu_0\geq\aleph_6(M,\alpha)+\sigma$, the following estimates hold:*

$$
|\!|\!|\mathcal{P}_7|\!|\!|_{-\frac12,s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-2}\big(1+\|\mathfrak{I}_0\|_{s+\aleph_6(M,\alpha)+\sigma}^{k_0,\gamma}\big)\,, \tag{12.39}
$$

$$
|\!|\!|\Delta_{12}\mathcal{P}_7|\!|\!|_{-\frac12,s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-2}\|\Delta_{12}i\|_{s_1+\aleph_6(M,\alpha)+\sigma}\,. \tag{12.40}
$$

Let $S>s_0$, $\beta_0\in\mathbb{N}$, and $M>\frac12(\beta_0+k_0)$. There exists a constant $\aleph_6'(M,\beta_0)>0$ such that, assuming (7.7) *with $\mu_0\geq\aleph_6'(M,\beta_0)+\sigma$, for any $m_1,m_2\geq0$, with $m_1+m_2\leq M-\frac12(\beta_0+k_0)$, for any $\beta\in\mathbb{N}^\nu$, $|\beta|\leq\beta_0$, the operators $\langle D\rangle^{m_1}\partial_\varphi^\beta\mathcal{T}_7\langle D\rangle^{m_2}$, $\langle D\rangle^{m_1}\partial_\varphi^\beta\Delta_{12}\mathcal{T}_7\langle D\rangle^{m_2}$ are $\mathcal{D}^{k_0}$-tame with tame constants satisfying*

$$
\mathfrak{M}_{\langle D\rangle^{m_1}\partial_\varphi^\beta\mathcal{T}_7\langle D\rangle^{m_2}}(s) \lesssim_{M,S} \varepsilon\gamma^{-2}\big(1+\|\mathfrak{I}_0\|_{s+\aleph_6'(M,\beta_0)+\sigma}\big)\,, \qquad \forall s_0\leq s\leq S \tag{12.41}
$$

$$
\|\langle D\rangle^{m_1}\Delta_{12}\partial_\varphi^\beta\mathcal{T}_7\langle D\rangle^{m_2}\|_{\mathcal{L}(H^{s_1})} \lesssim_{M,S} \varepsilon\gamma^{-2}\|\Delta_{12}i\|_{s_1+\aleph_6'(M,\beta_0)+\sigma}\,. \tag{12.42}
$$

Proof. Estimates (12.36) for a_9 defined in (12.30) follow by (11.4), (12.22), (12.23), (2.10) and (7.7).

PROOF OF (12.37)-(12.38). It follows by applying Propositions 2.42, 2.44, estimates (12.22)-(12.23) and using formula $\partial_\lambda^k\big((\Phi^{\pm 1} - \mathrm{Id})h\big) = \sum_{k_1+k_2=k} C(k_1,k_2)\partial_\lambda^{k_1}(\Phi^{\pm 1} - \mathrm{Id})\partial_\lambda^{k_2}h$, for any $k \in \mathbb{N}^{\nu+1}$, $|k| \le k_0$.

PROOF OF (12.39)-(12.40). First we prove (12.39), estimating each term in the definition (12.31) of P_7. The operator $A = \beta(\varphi,x)|D|^{\frac12}$ in (12.2) satisfies, by (2.47) and (12.22),

$$|A|^{k_0,\gamma}_{\frac12,s,\alpha} \lesssim_{s,\alpha} \|\beta\|_s^{k_0,\gamma} \lesssim_{s,\alpha} \varepsilon\gamma^{-2}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma}\big)\,. \tag{12.43}$$

The operator $P_6^{(0)}$ in (12.5) satisfies, by (11.4), (11.5), (2.47), (11.42),

$$|P_6^{(0)}|^{k_0,\gamma}_{\frac12,s,\alpha} \lesssim_{M,s,\alpha} 1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\aleph_5(M,\alpha)+\sigma}\,. \tag{12.44}$$

The estimate of the term $-\sum_{n=3}^{2M+1}\frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}^{n-1}_{A(\varphi)}\big(\omega\cdot\partial_\varphi A(\varphi)\big) + \sum_{n=2}^{2M}\frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}^n_{A(\varphi)}P_6^{(0)}$ in (12.31) then follows by (12.43), (12.44) and by applying Lemmata 2.10, 2.12. The term $\mathtt{R}_\beta \in OPS^{-\infty}$ defined in (12.28) can be estimated by (2.47) (applied with $A := \omega\cdot\partial_\varphi\beta$, $g(D) := (T_{\mathtt{h}}^{\frac12} - \mathrm{Id})|D|^{\frac12} \in OPS^{-\infty}$) and using (12.22), (3.51). The estimate of the terms $R_{A,P_6^{(0)}}, R_{A,\omega\cdot\partial_\varphi}$ in (12.31) follows by their definition given in (12.13), (12.17) and by estimates (11.4), (11.5), (11.42), (12.22), (2.10), (2.47), and Lemmata 2.10, 2.11. Since P_6 satisfies (11.42), estimate (12.39) is proved. Estimate (12.40) can be proved by similar arguments.

PROOF OF (12.41), (12.42). We estimate the term $\mathbf{\Phi}^{-1}\mathcal{Q}_6\mathbf{\Phi}$ in (12.35). For any $k \in \mathbb{N}^{\nu+1}, \beta \in \mathbb{N}^\nu$, $|k| \le k_0$, $|\beta| \le \beta_0$, $\lambda = (\omega,\mathtt{h})$, one has

$$\partial_\lambda^k\partial_\varphi^\beta(\mathbf{\Phi}^{-1}\mathcal{Q}_6\mathbf{\Phi}) = \sum_{\substack{\beta_1+\beta_2+\beta_3=\beta\\ k_1+k_2+k_3=k}} C(\beta_1,\beta_2,\beta_3,k_1,k_2,k_3)(\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\mathbf{\Phi}^{-1})(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}\mathcal{Q}_6)(\partial_\lambda^{k_3}\partial_\varphi^{\beta_3}\mathbf{\Phi})\,. \tag{12.45}$$

For any $m_1, m_2 \ge 0$ satisfying $m_1 + m_2 \le M - \frac12(\beta_0 + k_0)$, we have to provide an estimate for the operator

$$\langle D\rangle^{m_1}(\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\mathbf{\Phi}^{-1})(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}\mathcal{Q}_6)(\partial_\lambda^{k_3}\partial_\varphi^{\beta_3}\mathbf{\Phi})\langle D\rangle^{m_2}\,. \tag{12.46}$$

We write

$$(12.46) = \Big(\langle D\rangle^{m_1}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\mathbf{\Phi}^{-1}\langle D\rangle^{-\frac{|\beta_1|+|k_1|}{2}-m_1}\Big) \tag{12.47}$$

$$\circ\Big(\langle D\rangle^{\frac{|\beta_1|+|k_1|}{2}+m_1}\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}\mathcal{Q}_6\langle D\rangle^{\frac{|\beta_3|+|k_3|}{2}+m_2}\Big) \tag{12.48}$$

$$\circ\Big(\langle D\rangle^{-m_2-\frac{|\beta_3|+|k_3|}{2}}\partial_\lambda^{k_3}\partial_\varphi^{\beta_3}\mathbf{\Phi}\langle D\rangle^{m_2}\Big)\,. \tag{12.49}$$

The terms (12.47)-(12.49) can be estimated separately. To estimate the terms (12.47) and (12.49), we apply (2.101) of Proposition 2.41, (2.111) of Proposition 2.43, and (12.22)-(12.23). The pseudo-differential operator in (12.48) is estimated in $|\ |_{0,s,0}$ norm by using (2.42), (2.45), (2.47), bounds (11.42), (11.43) on $\mathcal{Q}_6$, and the fact that $\frac{|\beta_1|+|k_1|}{2} + m_1 + \frac{|\beta_3|+|k_3|}{2} + m_2 - M \le 0$. Then its action on Sobolev functions is deduced by Lemma 2.29. As a consequence, each operator in (12.46), and hence the whole operator (12.45), satisfies (12.41).

The estimates of the terms in (12.32) can be done arguing similarly, using also Lemma 2.12 and (12.43)-(12.44). The term $\langle D\rangle^{m_1}\partial_\varphi^\beta\Pi_0(\mathbf{\Phi} - \mathbb{I}_2)\langle D\rangle^{m_2}$ can be estimated by applying Lemma 2.39 (with $A = \mathbb{I}_2$, $B = \mathbf{\Phi}$) and (12.37), (12.22), (12.23). □

13 Reduction of the lower orders

In this section we complete the reduction of the operator $\mathcal{L}_7$ in (12.33) to constant coefficients, up to a regularizing remainder of order $|D|^{-M}$. We write

$$\mathcal{L}_7 = \begin{pmatrix} L_7 & 0 \\ 0 & \overline{L_7} \end{pmatrix} + \mathrm{i}\Pi_0 + \mathcal{T}_7\,, \tag{13.1}$$

where

$$L_7 := \omega \cdot \partial_\varphi + \mathtt{im}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}} + a_8 + a_9 \mathcal{H} + P_7 \,, \tag{13.2}$$

the real valued functions a_8, a_9 are introduced in (11.1), (12.30), satisfy (12.34), and the operator $P_7 \in OPS^{-\frac{1}{2}}$ in (12.31) is even and reversible. We first conjugate the operator L_7.

13.1 Reduction of the order 0

In this subsection we reduce to constant coefficients the term $a_8 + a_9\mathcal{H}$ of order zero of L_7 in (13.2). We begin with removing the dependence of $a_8 + a_9\mathcal{H}$ on φ. It turns out that, since a_8, a_9 are odd functions in φ by (12.34), thus with zero average, this step removes completely the terms of order zero. Consider the transformation

$$W_0 := \mathrm{Id} + f_0(\varphi, x) + g_0(\varphi, x)\mathcal{H} \,, \tag{13.3}$$

where f_0, g_0 are real valued functions to be determined. One has

$$\begin{aligned} L_7 W_0 = {} & W_0\big(\omega \cdot \partial_\varphi + \mathtt{im}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}\big) + a_8 + a_9\mathcal{H} + (\omega \cdot \partial_\varphi f_0) + (\omega \cdot \partial_\varphi g_0)\mathcal{H} + \big(a_8 + a_9\mathcal{H}\big)\big(f_0 + g_0\mathcal{H}\big) \\ & + \big[\mathtt{im}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}, W_0\big] + P_7 W_0 \,. \end{aligned} \tag{13.4}$$

Since $\mathcal{H}^2 = -\mathrm{Id} + \pi_0$ on the periodic functions, where π_0 is defined in (2.35), we write

$$\begin{aligned} \big(a_8 + a_9\mathcal{H}\big)\big(f_0 + g_0\mathcal{H}\big) & = a_8 f_0 + a_8 g_0 \mathcal{H} + a_9 \mathcal{H} f_0 + a_9 \mathcal{H} g_0 \mathcal{H} \\ & = (a_8 f_0 - a_9 g_0) + (a_8 g_0 + a_9 f_0)\mathcal{H} + a_9[\mathcal{H}, f_0] + a_9[\mathcal{H}, g_0]\mathcal{H} + a_9 g_0 \pi_0 \,. \end{aligned} \tag{13.5}$$

Then, by (13.4), (13.5), one has

$$\begin{aligned} L_7 W_0 = {} & W_0\big(\omega \cdot \partial_\varphi + \mathtt{im}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}\big) + (\omega \cdot \partial_\varphi f_0 + a_8 + a_8 f_0 - a_9 g_0) \\ & + (\omega \cdot \partial_\varphi g_0 + a_9 + a_8 g_0 + a_9 f_0)\mathcal{H} + \breve{P}_7 \end{aligned} \tag{13.6}$$

where $\breve{P}_7 \in OPS^{-\frac{1}{2}}$ is the operator

$$\breve{P}_7 := a_9[\mathcal{H}, f_0] + a_9[\mathcal{H}, g_0]\mathcal{H} + [\mathtt{im}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}, W_0] + P_7 W_0 + a_9 g_0 \pi_0 \,. \tag{13.7}$$

In order to eliminate the zero order terms in (13.6) we choose the functions f_0, g_0 such that

$$\begin{cases} \omega \cdot \partial_\varphi f_0 + a_8 + a_8 f_0 - a_9 g_0 = 0 \\ \omega \cdot \partial_\varphi g_0 + a_9 + a_8 g_0 + a_9 f_0 = 0 \,. \end{cases} \tag{13.8}$$

Writing $z_0 = 1 + f_0 + \mathrm{i} g_0$, the real system (13.8) is equivalent to the complex scalar equation

$$\omega \cdot \partial_\varphi z_0 + (a_8 + \mathrm{i} a_9) z_0 = 0 \,. \tag{13.9}$$

Since a_8, a_9 are odd functions in φ, we choose, for all $\omega \in \mathtt{DC}(\gamma, \tau)$, the periodic function

$$z_0 := \exp(p_0), \quad p_0 := -(\omega \cdot \partial_\varphi)^{-1}(a_8 + \mathrm{i} a_9), \tag{13.10}$$

which solves (13.9). Thus the real functions

$$\begin{aligned} f_0 & := \mathrm{Re}(z_0) - 1 = \exp(-(\omega \cdot \partial_\varphi)^{-1} a_8) \cos((\omega \cdot \partial_\varphi)^{-1} a_9) - 1, \\ g_0 & := \mathrm{Im}(z_0) = -\exp(-(\omega \cdot \partial_\varphi)^{-1} a_8) \sin((\omega \cdot \partial_\varphi)^{-1} a_9) \end{aligned} \tag{13.11}$$

solve (13.8), and, for $\omega \in \mathtt{DC}(\gamma, \tau)$, equation (13.6) reduces to

$$L_7 W_0 = W_0(\omega \cdot \partial_\varphi + \mathtt{im}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}) + \breve{P}_7 \,. \tag{13.12}$$

We extend the function p_0 in (13.10) to the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ by using the extended operator $(\omega \cdot \partial_\varphi)^{-1}_{ext}$ introduced in Lemma 2.5. Thus the functions z_0, f_0, g_0 in (13.10), (13.11) are defined on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ as well.

Lemma 13.1. *The real valued functions f_0, g_0 in* (13.11) *satisfy*

$$f_0 = even(\varphi) even(x) \,, \quad g_0 = even(\varphi) odd(x) \,. \tag{13.13}$$

Moreover, there exists $\sigma(k_0, \tau, \nu) > 0$ such that, if (7.7) *holds with $\mu_0 \geq \sigma$, then*

$$\|f_0\|_s^{k_0,\gamma} \,, \|g_0\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-3}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) \,, \quad \|\Delta_{12} f_0\|_{s_1}, \|\Delta_{12} g_0\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-3} \|\Delta_{12} i\|_{s_1+\sigma} \,. \tag{13.14}$$

The operator W_0 defined in (13.3) *is even, reversibility preserving, invertible and for any $\alpha > 0$, assuming* (7.7) *with $\mu_0 \geq \alpha + \sigma$, the following estimates hold:*

$$|W_0^{\pm 1} - \mathrm{Id}|_{0,s,\alpha}^{k_0,\gamma} \lesssim_{s,\alpha} \varepsilon\gamma^{-3}\big(1 + \|\mathfrak{I}_0\|_{s+\alpha+\sigma}^{k_0,\gamma}\big) \,, \quad |\Delta_{12} W_0^{\pm 1}|_{0,s_1,\alpha} \lesssim_{s_1,\alpha} \varepsilon\gamma^{-3} \|\Delta_{12} i\|_{s_1+\alpha+\sigma} \,. \tag{13.15}$$

Proof. The parities in (13.13) follow by (13.11) and (12.34). Therefore W_0 in (13.3) is even and reversibility preserving. Estimates (13.14) follow by (13.11), (11.5), (12.36), (2.10), (2.17), (2.19). The operator W_0 defined in (13.3) is invertible by Lemma 2.14, (13.14), (7.7), for $\varepsilon\gamma^{-3}$ small enough. Estimates (13.15) then follow by (13.14), using (2.41), (2.47) and Lemma 2.14. □

For $\omega \in \mathtt{DC}(\gamma, \tau)$, by (13.12) we obtain the even and reversible operator

$$L_7^{(1)} := W_0^{-1} L_7 W_0 = \omega \cdot \partial_\varphi + \mathrm{i} \mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}} + P_7^{(1)} \,, \qquad P_7^{(1)} := W_0^{-1} \breve{P}_7 \,, \tag{13.16}$$

where $\breve{P}_7$ is the operator in $OPS^{-\frac{1}{2}}$ defined in (13.7).

Since the functions f_0, g_0 are defined on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, the operator $\breve{P}_7$ in (13.7) is defined on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, and $\omega \cdot \partial_\varphi + \mathrm{i} \mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}} + P_7^{(1)}$ in (13.16) is an extension of $L_7^{(1)}$ to $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, which we still denote by $L_7^{(1)}$.

Lemma 13.2. *For any $M, \alpha > 0$, there exists a constant $\aleph_7^{(1)}(M, \alpha) > 0$ such that if* (7.7) *holds with $\mu_0 \geq \aleph_7^{(1)}(M, \alpha)$, the remainder $P_7^{(1)} \in OPS^{-\frac{1}{2}}$, defined in* (13.16), *satisfies*

$$\begin{aligned} &|P_7^{(1)}|_{-\frac{1}{2},s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-3}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(1)}(M,\alpha)}^{k_0,\gamma}\big) \,, \\ &|\Delta_{12} P_7^{(1)}|_{-\frac{1}{2},s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-3} \|\Delta_{12} i\|_{s_1+\aleph_7^{(1)}(M,\alpha)} \,. \end{aligned} \tag{13.17}$$

Proof. Estimates (13.17) follow by the definition of $P_7^{(1)}$ given in (13.16), by estimates (13.14), (13.15), (12.26), (12.36), (12.39), (12.40), by applying (2.41), (2.45), (2.47), (2.51) and using also Lemma 2.17. The fact that $P_7^{(1)}$ has size $\varepsilon\gamma^{-3}$ is due to the term $[\mathrm{i} \mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}, W_0] = [\mathrm{i} \mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}, W_0 - \mathrm{Id}]$, because $\mathtt{m}_{\frac{1}{2}} = 1 + O(\varepsilon\gamma^{-1})$ and $W_0 - \mathrm{Id} = O(\varepsilon\gamma^{-3})$. □

We underline that the operator $L_7^{(1)}$ in (13.16) does not contain terms of order zero.

13.2 Reduction at negative orders

In this subsection we define inductively a finite number of transformations to the aim of reducing to constant coefficients all the symbols of orders $> -M$ of the operator $L_7^{(1)}$ in (13.16). The constant M will be fixed in (15.16).

In the rest of the section we prove the following inductive claim:

- Diagonalization of $L_7^{(1)}$ in decreasing orders. For any $m \in \{1, \ldots, 2M\}$, we have an even and reversible operator of the form

$$L_7^{(m)} := \omega \cdot \partial_\varphi + \Lambda_m(D) + P_7^{(m)} \,, \quad P_7^{(m)} \in OPS^{-\frac{m}{2}} \,, \tag{13.18}$$

where

$$\Lambda_m(D) := \mathrm{i} \mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}} + r_m(D) \,, \qquad r_m(D) \in OPS^{-\frac{1}{2}} \,. \tag{13.19}$$

The operator $r_m(D)$ is an even and reversible Fourier multiplier, independent of (φ, x). Also the operator $P_7^{(m)}$ is even and reversible.

For any $M, \alpha > 0$, there exists a constant $\aleph_7^{(m)}(M,\alpha) > 0$ (depending also on τ, k_0, ν) such that, if (7.7) holds with $\mu_0 \geq \aleph_7^{(m)}(M,\alpha)$, then the following estimates hold:

$$|r_m(D)|_{-\frac{1}{2},s,\alpha}^{k_0,\gamma} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+1)}, \quad |\Delta_{12} r_m(D)|_{-\frac{1}{2},s_1,\alpha} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+1)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(m)}(M,\alpha)}, \tag{13.20}$$

$$|P_7^{(m)}|_{-\frac{m}{2},s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+2)}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(m)}(M,\alpha)}^{k_0,\gamma}\big), \tag{13.21}$$

$$|\Delta_{12} P_7^{(m)}|_{-\frac{m}{2},s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+2)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(m)}(M,\alpha)}. \tag{13.22}$$

Note that by (13.19), using (12.26), (13.20) and (2.42) (applied for $g(D) = T_{\mathtt{h}}^{\frac{1}{2}}|D|^{\frac{1}{2}}$) one gets

$$|\Lambda_m(D)|_{\frac{1}{2},s,\alpha}^{k_0,\gamma} \lesssim_{M,\alpha} 1, \quad |\Delta_{12}\Lambda_m(D)|_{\frac{1}{2},s_1,\alpha} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+1)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(m)}(M,\alpha)}. \tag{13.23}$$

For $m \geq 2$ there exist real, even, reversibility preserving, invertible maps $W_{m-1}^{(0)}$, $W_{m-1}^{(1)}$ of the form

$$\begin{aligned} W_{m-1}^{(0)} &:= \mathrm{Id} + w_{m-1}^{(0)}(\varphi, x, D) \qquad \text{with} \qquad w_{m-1}^{(0)}(\varphi, x, \xi) \in S^{-\frac{m-1}{2}}, \\ W_{m-1}^{(1)} &:= \mathrm{Id} + w_{m-1}^{(1)}(x, D) \qquad \text{with} \qquad w_{m-1}^{(1)}(x, \xi) \in S^{-\frac{m-1}{2}+\frac{1}{2}} \end{aligned} \tag{13.24}$$

such that, for all $\omega \in \mathtt{DC}(\gamma,\tau)$,

$$L_7^{(m)} = (W_{m-1}^{(1)})^{-1}(W_{m-1}^{(0)})^{-1} L_7^{(m-1)} W_{m-1}^{(0)} W_{m-1}^{(1)}. \tag{13.25}$$

Initialization. For $m = 1$, the even and reversible operator $L_7^{(1)}$ in (13.16) has the form (13.18)-(13.19) with

$$r_1(D) = 0, \quad \Lambda_1(D) = \mathtt{im}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}. \tag{13.26}$$

Since $\Lambda_1(D)$ is even and reversible, by difference, the operator $P_7^{(1)}$ is even and reversible as well. At $m = 1$, estimate (13.20) is trivial and (13.21)-(13.22) are (13.17).

Inductive step. In the next two subsections, we prove the above inductive claim, see (13.61)-(13.63) and Lemma 13.6. We perform this reduction in two steps:

1. First we look for a transformation $W_m^{(0)}$ to remove the dependence on φ of the terms of order $-m/2$ of the operator $L_7^{(m)}$ in (13.18), see (13.29). The resulting conjugated operator is $L_7^{(m,1)}$ in (13.36).
2. Then we look for a transformation $W_m^{(1)}$ to remove the dependence on x of the terms of order $-m/2$ of the operator $L_7^{(m,1)}$ in (13.36), see (13.49) and (13.53).

13.2.1 Elimination of the dependence on φ

In this subsection we eliminate the dependence on φ from the terms of order $-m/2$ in $P_7^{(m)}$ in (13.18). We conjugate the operator $L_7^{(m)}$ in (13.18) by a transformation of the form (see (13.24))

$$W_m^{(0)} := \mathrm{Id} + w_m^{(0)}(\varphi, x, D), \quad \text{with} \quad w_m^{(0)}(\varphi, x, \xi) \in S^{-\frac{m}{2}}, \tag{13.27}$$

which we shall fix in (13.31). We compute

$$\begin{aligned} L_7^{(m)} W_m^{(0)} = W_m^{(0)}\big(\omega\cdot\partial_\varphi + \Lambda_m(D)\big) + (\omega\cdot\partial_\varphi w_m^{(0)})(\varphi, x, D) + P_7^{(m)} \\ + \big[\Lambda_m(D), w_m^{(0)}(\varphi, x, D)\big] + P_7^{(m)} w_m^{(0)}(\varphi, x, D). \end{aligned} \tag{13.28}$$

Since $\Lambda_m(D) \in OPS^{\frac{1}{2}}$ and the operators $P_7^{(m)}$, $w_m^{(0)}(\varphi, x, D)$ are in $OPS^{-\frac{m}{2}}$, with $m \geq 1$, we have that the commutator $[\Lambda_m(D), w_m^{(0)}(\varphi, x, D)]$ is in $OPS^{-\frac{m}{2}-\frac{1}{2}}$ and $P_7^{(m)} w_m^{(0)}(\varphi, x, D)$ is in $OPS^{-m} \subseteq OPS^{-\frac{m}{2}-\frac{1}{2}}$. Thus the term of order $-m/2$ in (13.28) is

$$(\omega \cdot \partial_\varphi w_m^{(0)})(\varphi, x, D) + P_7^{(m)} .$$

Let $p_7^{(m)}(\varphi, x, \xi) \in S^{-\frac{m}{2}}$ be the symbol of $P_7^{(m)}$. We look for $w_m^{(0)}(\varphi, x, \xi)$ such that

$$\omega \cdot \partial_\varphi w_m^{(0)}(\varphi, x, \xi) + p_7^{(m)}(\varphi, x, \xi) = \langle p_7^{(m)} \rangle_\varphi (x, \xi) \tag{13.29}$$

where

$$\langle p_7^{(m)} \rangle_\varphi (x, \xi) := \frac{1}{(2\pi)^\nu} \int_{\mathbb{T}^\nu} p_7^{(m)}(\varphi, x, \xi)\, d\varphi . \tag{13.30}$$

For all $\omega \in \mathtt{DC}(\gamma, \tau)$, we choose the solution of (13.29) given by the periodic function

$$w_m^{(0)}(\varphi, x, \xi) := (\omega \cdot \partial_\varphi)^{-1} \Big(\langle p_7^{(m)} \rangle_\varphi (x, \xi) - p_7^{(m)}(\varphi, x, \xi) \Big) . \tag{13.31}$$

We extend the symbol $w_m^{(0)}$ in (13.31) to the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ by using the extended operator $(\omega \cdot \partial_\varphi)^{-1}_{ext}$ introduced in Lemma 2.5. As a consequence, the operator $W_m^{(0)}$ in (13.27) is extended accordingly. We still denote by $w_m^{(0)}, W_m^{(0)}$ these extensions.

Lemma 13.3. *The operator $W_m^{(0)}$ defined in* (13.27), (13.31) *is even and reversibility preserving. For any $\alpha, M > 0$ there exists a constant $\aleph_7^{(m,1)}(M, \alpha) > 0$ (depending also on k_0, τ, ν), larger than the constant $\aleph_7^{(m)}(M, \alpha)$ appearing in* (13.20)-(13.23) *such that, if* (7.7) *holds with $\mu_0 \geq \aleph_7^{(m,1)}(M, \alpha)$, then for any $s \geq s_0$*

$$|\mathrm{Op}(w_m^{(0)})|_{-\frac{m}{2}, s, \alpha}^{k_0, \gamma} \lesssim_{M, s, \alpha} \varepsilon \gamma^{-(m+3)} \big(1 + \|\mathfrak{I}_0\|_{s + \aleph_7^{(m,1)}(M, \alpha)}^{k_0, \gamma} \big) \tag{13.32}$$

$$|\Delta_{12} \mathrm{Op}(w_m^{(0)})|_{-\frac{m}{2}, s_1, \alpha} \lesssim_{M, s_1, \alpha} \varepsilon \gamma^{-(m+3)} \|\Delta_{12} i\|_{s_1 + \aleph_7^{(m,1)}(M, \alpha)} . \tag{13.33}$$

As a consequence, the transformation $W_m^{(0)}$ defined in (13.27), (13.31) *is invertible and*

$$|(W_m^{(0)})^{\pm 1} - \mathrm{Id}|_{0, s, \alpha}^{k_0, \gamma} \lesssim_{M, s, \alpha} \varepsilon \gamma^{-(m+3)} \big(1 + \|\mathfrak{I}_0\|_{s + \aleph_7^{(m,1)}(M, \alpha)}^{k_0, \gamma} \big) \tag{13.34}$$

$$|\Delta_{12} (W_m^{(0)})^{\pm 1}|_{0, s_1, \alpha} \lesssim_{M, s_1, \alpha} \varepsilon \gamma^{-(m+3)} \|\Delta_{12} i\|_{s_1 + \aleph_7^{(m,1)}(M, \alpha)} . \tag{13.35}$$

Proof. We begin with proving (13.32). By (2.37)-(2.38) one has

$$|\mathrm{Op}(w_m^{(0)})|_{-\frac{m}{2}, s, \alpha}^{k_0, \gamma} \lesssim_{k_0, \nu} \max_{\beta \in [0, \alpha]} \sup_{\xi \in \mathbb{R}} \langle \xi \rangle^{\frac{m}{2} + \beta} \big\| \partial_\xi^\beta w_m^{(0)}(\cdot, \cdot, \cdot, \xi) \big\|_s^{k_0, \gamma} .$$

By (13.31) and (2.17), for each $\xi \in \mathbb{R}$ one has

$$\|\partial_\xi^\beta w_m^{(0)}(\cdot, \cdot, \cdot, \xi)\|_s^{k_0, \gamma} \lesssim_{k_0, \nu} \gamma^{-1} \big\| \partial_\xi^\beta \big(\langle p_7^{(m)} \rangle_\varphi (\cdot, \xi) - p_7^{(m)}(\cdot, \cdot, \xi) \big) \big\|_{s+\mu}^{k_0, \gamma}$$

where μ is defined in (2.18) with $k + 1 = k_0$. Hence

$$|\mathrm{Op}(w_m^{(0)})|_{-\frac{m}{2}, s, \alpha}^{k_0, \gamma} \lesssim_{k_0, \nu} \gamma^{-1} |P_7^{(m)}|_{-\frac{m}{2}, s+\mu, \alpha}^{k_0, \gamma} ,$$

and (13.32) follows by (13.21). The other bounds are proved similarly, using the explicit formula (13.31), estimates (13.21)-(13.22) and (2.17), (2.45), and Lemma 2.14. □

By (13.28) and (13.29) we get

$$
\begin{aligned}
L_7^{(m)} W_m^{(0)} &= W_m^{(0)}\big(\omega\cdot\partial_\varphi + \Lambda_m(D)\big) + \langle p_7^{(m)}\rangle_\varphi(x,D) + \big[\Lambda_m(D), w_m^{(0)}(\varphi,x,D)\big] + P_7^{(m)} w_m^{(0)}(\varphi,x,D) \\
&= W_m^{(0)}\big(\omega\cdot\partial_\varphi + \Lambda_m(D) + \langle p_7^{(m)}\rangle_\varphi(x,D)\big) - w_m^{(0)}(\varphi,x,D)\langle p_7^{(m)}\rangle_\varphi(x,D) \\
&\quad + \big[\Lambda_m(D), w_m^{(0)}(\varphi,x,D)\big] + P_7^{(m)} w_m^{(0)}(\varphi,x,D)
\end{aligned}
$$

and therefore we obtain the even and reversible operator

$$
L_7^{(m,1)} := (W_m^{(0)})^{-1} L_7^{(m)} W_m^{(0)} = \omega\cdot\partial_\varphi + \Lambda_m(D) + \langle p_7^{(m)}\rangle_\varphi(x,D) + P_7^{(m,1)} \tag{13.36}
$$

where

$$
P_7^{(m,1)} := (W_m^{(0)})^{-1}\Big(\big[\Lambda_m(D), w_m^{(0)}(\varphi,x,D)\big] + P_7^{(m)} w_m^{(0)}(\varphi,x,D) - w_m^{(0)}(\varphi,x,D)\langle p_7^{(m)}\rangle_\varphi(x,D)\Big) \tag{13.37}
$$

is in $OPS^{-\frac{m}{2}-\frac12}$, as we prove in Lemma 13.4 below. Thus the term of order $-\frac{m}{2}$ in (13.36) is $\langle p_7^{(m)}\rangle_\varphi(x,D)$, which does not depend on φ any more.

Lemma 13.4. *The operators $\langle p_7^{(m)}\rangle_\varphi(x,D)$ and $P_7^{(m,1)}$ are even and reversible. The operator $P_7^{(m,1)}$ in* (13.37) *is in $OPS^{-\frac{m}{2}-\frac12}$. For any $\alpha, M > 0$ there exists a constant $\aleph_7^{(m,2)}(M,\alpha) > 0$ (depending also on k_0,τ,ν), larger than the constant $\aleph_7^{(m,1)}(M,\alpha)$ appearing in Lemma 13.3, such that, if* (7.7) *holds with $\mu_0 \geq \aleph_7^{(m,2)}(M,\alpha)$, then for any $s \geq s_0$*

$$
|P_7^{(m,1)}|_{-\frac{m}{2}-\frac12,s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)}(1+\|\mathfrak{I}_0\|_{s+\aleph_7^{(m,2)}(M,\alpha)}^{k_0,\gamma}), \tag{13.38}
$$

$$
|\Delta_{12} P_7^{(m,1)}|_{-\frac{m}{2}-\frac12,s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m,2)}(M,\alpha)}\,. \tag{13.39}
$$

Proof. Since $P_7^{(m)}(x,D)$ is even and reversible by the inductive claim, its φ-average $\langle p_7^{(m)}\rangle_\varphi(x,D)$ defined in (13.30) is even and reversible as well. Since $\Lambda_m(D)$ is reversible and $W_m^{(0)}$ is reversibility preserving we obtain that $P_7^{(m,1)}$ in (13.37) is even and reversible.

Let us prove that $P_7^{(m,1)}$ is in $OPS^{-\frac{m}{2}-\frac12}$. Since $\Lambda_m(D) \in OPS^{\frac12}$ and the operators $P_7^{(m)}$, $w_m^{(0)}(\varphi,x,D)$ are in $OPS^{-\frac{m}{2}}$, with $m \geq 1$, we have that $[\Lambda_m(D), w_m^{(0)}(\varphi,x,D)]$ is in $OPS^{-\frac{m}{2}-\frac12}$ and $P_7^{(m)} w_m^{(0)}(\varphi,x,D)$ is in $OPS^{-m} \subseteq OPS^{-\frac{m}{2}-\frac12}$. Moreover also $w_m^{(0)}(\varphi,x,D)\langle p_7^{(m)}\rangle_\varphi(x,D) \in OPS^{-m} \subseteq OPS^{-\frac{m}{2}-\frac12}$, since $m \geq 1$. Since $(W_m^{(0)})^{-1}$ is in OPS^0, the remainder $P_7^{(m,1)}$ is in $OPS^{-\frac{m}{2}-\frac12}$. Bounds (13.38)-(13.39) follow by the explicit expression in (13.37), Lemma 13.3, estimates (13.20)-(13.23), and (2.43), (2.45), (2.51). ☐

13.2.2 Elimination of the dependence on x

In this subsection we eliminate the dependence on x from $\langle p_7^{(m)}\rangle_\varphi(x,D)$, which is the only term of order $-m/2$ in (13.36). To this aim we conjugate $L_7^{(m,1)}$ in (13.36) by a transformation of the form

$$
W_m^{(1)} := \mathrm{Id} + w_m^{(1)}(x,D), \quad \text{where} \quad w_m^{(1)}(x,\xi) \in S^{-\frac{m}{2}+\frac12} \tag{13.40}
$$

is a φ-independent symbol, which we shall fix in (13.51) (for $m = 1$) and (13.55) (for $m \geq 2$). We denote the space average of the function $\langle p_7^{(m)}\rangle_\varphi(x,\xi)$ defined in (13.30) by

$$
\langle p_7^{(m)}\rangle_{\varphi,x}(\xi) := \frac{1}{2\pi}\int_{\mathbb{T}} \langle p_7^{(m)}\rangle_\varphi(x,\xi)\,dx = \frac{1}{(2\pi)^{\nu+1}}\int_{\mathbb{T}^{\nu+1}} p_7^{(m)}(\varphi,x,\xi)\,d\varphi\,dx\,. \tag{13.41}
$$

By (13.36), we compute

$$
\begin{aligned}
L_7^{(m,1)} W_m^{(1)} &= W_m^{(1)}\Big(\omega\cdot\partial_\varphi + \Lambda_m(D) + \langle p_7^{(m)}\rangle_{\varphi,x}\Big) + \big[\Lambda_m(D), w_m^{(1)}(x,D)\big] + \langle p_7^{(m)}\rangle_\varphi(x,D) - \langle p_7^{(m)}\rangle_{\varphi,x}(D) \\
&\quad + \langle p_7^{(m)}\rangle_\varphi(x,D) w_m^{(1)}(x,D) - w_m^{(1)}(x,D)\langle p_7^{(m)}\rangle_{\varphi,x}(D) + P_7^{(m,1)} W_m^{(1)}\,.
\end{aligned} \tag{13.42}
$$

By formulas (2.30), (2.31) (with $N = 1$) and (2.48), (2.49),

$$\langle p_7^{(m)}\rangle_\varphi(x,D)w_m^{(1)}(x,D) = \mathrm{Op}\Big(\langle p_7^{(m)}\rangle_\varphi(x,\xi)w_m^{(1)}(x,\xi)\Big) + r_{\langle p_7^{(m)}\rangle_\varphi, w_m^{(1)}}(x,D)\,, \tag{13.43}$$

$$w_m^{(1)}(x,D)\langle p_7^{(m)}\rangle_{\varphi,x}(D) = \mathrm{Op}\Big(w_m^{(1)}(x,\xi)\langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\Big) + r_{w_m^{(1)},\langle p_7^{(m)}\rangle_{\varphi,x}}(x,D)\,, \tag{13.44}$$

$$\big[\Lambda_m(D), w_m^{(1)}(x,D)\big] = \mathrm{Op}\Big(-\mathrm{i}\partial_\xi\Lambda_m(\xi)\partial_x w_m^{(1)}(x,\xi)\Big) + \mathtt{r}_2(\Lambda_m, w_m^{(1)})(x,D) \tag{13.45}$$

where

$$r_{\langle p_7^{(m)}\rangle_\varphi, w_m^{(1)}}\ ,\ r_{w_m^{(1)},\langle p_7^{(m)}\rangle_{\varphi,x}} \in S^{-m-\frac12} \subset S^{-\frac m2-\frac12}\,, \qquad \mathtt{r}_2(\Lambda_m, w_m^{(1)})(x,D) \in S^{-\frac m2-1} \subset S^{-\frac m2-\frac12}\,. \tag{13.46}$$

Let $\chi_0 \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ be a cut-off function satisfying

$$\chi_0(\xi) = \chi_0(-\xi) \ \ \forall \xi \in \mathbb{R}\,, \quad \chi_0(\xi) = 0 \ \ \forall |\xi| \le \frac45\,, \quad \chi_0(\xi) = 1 \ \ \forall |\xi| \ge \frac78\,. \tag{13.47}$$

By (13.42)-(13.46), one has

$$\begin{aligned} L_7^{(m,1)} W_m^{(1)} &= W_m^{(1)}\Big(\omega\cdot\partial_\varphi + \Lambda_m(D) + \langle p_7^{(m)}\rangle_{\varphi,x}(D)\Big) \\ &\quad + \mathrm{Op}\Big(-\mathrm{i}\partial_\xi\Lambda_m(\xi)\partial_x w_m^{(1)}(x,\xi) + \chi_0(\xi)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big) \\ &\qquad\quad + \chi_0(\xi)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)w_m^{(1)}(x,\xi)\Big) \\ &\quad + \mathrm{Op}\Big(\big(1-\chi_0(\xi)\big)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)\big(1 + w_m^{(1)}(x,\xi)\big)\Big) \\ &\quad + \mathtt{r}_2(\Lambda_m, w_m^{(1)})(x,D) + r_{\langle p_7^{(m)}\rangle_\varphi, w_m^{(1)}}(x,D) - r_{w_m^{(1)},\langle p_7^{(m)}\rangle_{\varphi,x}}(x,D) + P_7^{(m,1)}W_m^{(1)}\,. \end{aligned} \tag{13.48}$$

The terms containing $1-\chi_0(\xi)$ are in $S^{-\infty}$, by definition (13.47). The term

$$\mathrm{Op}\Big(-\mathrm{i}\partial_\xi\Lambda_m(\xi)\partial_x w_m^{(1)}(x,\xi) + \chi_0(\xi)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)\Big)$$

is of order $-\frac m2$. The term

$$\mathrm{Op}\,\Big(\chi_0(\xi)\big\{\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big\}w_m^{(1)}(x,\xi)\Big)$$

is of order $-m+\frac12$, which equals $-\frac m2$ for $m=1$, and is strictly less than $-\frac m2$ for $m \ge 2$. Hence we split the two cases $m=1$ and $m \ge 2$.

First case: $m = 1$. We look for $w_m^{(1)}(x,\xi) = w_1^{(1)}(x,\xi)$ such that

$$-\mathrm{i}\partial_\xi\Lambda_1(\xi)\partial_x w_1^{(1)}(x,\xi) + \chi_0(\xi)\Big(\langle p_7^{(1)}\rangle_\varphi(x,\xi) - \langle p_7^{(1)}\rangle_{\varphi,x}(\xi)\Big)(1 + w_1^{(1)}(x,\xi)) = 0\,. \tag{13.49}$$

By (13.26) and recalling (2.33), (2.16), for $|\xi| \ge 4/5$ one has $\Lambda_1(\xi) = \mathrm{i}\mathtt{m}_{\frac12}\tanh^{\frac12}(\mathtt{h}|\xi|)|\xi|^{\frac12}$. Since, by (12.26), $|\mathtt{m}_{\frac12}| \ge 1/2$ for $\varepsilon\gamma^{-1}$ small enough, we have

$$\inf_{|\xi|\ge\frac45} |\xi|^{\frac12}|\partial_\xi\Lambda_1(\xi)| \ge \delta > 0\,, \tag{13.50}$$

where δ depends only on $\mathtt{h}_1$. Using that $\langle p_7^{(1)}\rangle_\varphi - \langle p_7^{(1)}\rangle_{\varphi,x}$ has zero average in x, we choose the solution of (13.49) given by the periodic function

$$w_1^{(1)}(x,\xi) := \exp\big(g_1(x,\xi)\big) - 1, \quad g_1(x,\xi) := \begin{cases} \dfrac{\chi_0(\xi)\partial_x^{-1}\big(\langle p_7^{(1)}\rangle_\varphi(x,\xi) - \langle p_7^{(1)}\rangle_{\varphi,x}(\xi)\big)}{\mathrm{i}\partial_\xi\Lambda_1(\xi)} & \text{if } |\xi| \ge \frac45 \\ 0 & \text{if } |\xi| \le \frac45\,. \end{cases} \tag{13.51}$$

Note that, by the definition of the cut-off function χ_0 given in (13.47), recalling (13.26), (13.50) and applying estimates (2.42), (12.26), the Fourier multiplier $\frac{\chi_0(\xi)}{\partial_\xi \Lambda_1(\xi)}$ is a symbol in $S^{\frac12}$ and satisfies

$$\Big|\mathrm{Op}\Big(\frac{\chi_0(\xi)}{\partial_\xi \Lambda_1(\xi)}\Big)\Big|^{k_0,\gamma}_{\frac12,s,\alpha} \lesssim_\alpha 1\,, \quad \Big|\Delta_{12}\mathrm{Op}\Big(\frac{\chi_0(\xi)}{\partial_\xi \Lambda_1(\xi)}\Big)\Big|_{\frac12,s_1,\alpha} \lesssim_\alpha \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1}\,. \tag{13.52}$$

Therefore the function $g_1(x,\xi)$ is a well-defined symbol in S^0.

Second case: $m \geq 2$. We look for $w_m^{(1)}(x,\xi)$ such that

$$-\,\mathrm{i}\partial_\xi\Lambda_m(\xi)\partial_x w_m^{(1)}(x,\xi) + \chi_0(\xi)\Big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\Big) = 0\,. \tag{13.53}$$

Recalling (13.19)-(13.20) and (13.50), one has that

$$\begin{aligned}\inf_{|\xi|\geq\frac45}|\xi|^{\frac12}|\partial_\xi\Lambda_m(\xi)| &\geq \inf_{|\xi|\geq\frac45}|\xi|^{\frac12}|\partial_\xi\Lambda_1(\xi)| - \sup_{\xi\in\mathbb{R}}|\xi|^{\frac12}|\partial_\xi r_m(\xi)| \geq \delta - |r_m(D)|_{-\frac12,0,1}\\ &\geq \delta - C\varepsilon\gamma^{-(m+1)} \geq \frac{\delta}{2}\end{aligned} \tag{13.54}$$

for $\varepsilon\gamma^{-(m+1)}$ small enough. Since $\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)$ has zero average in x, we choose the solution of (13.53) given by the periodic function

$$w_m^{(1)}(x,\xi) := \begin{cases} \dfrac{\chi_0(\xi)\partial_x^{-1}\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)}{\mathrm{i}\partial_\xi\Lambda_m(\xi)} & \text{if } |\xi|\geq\frac45\\ 0 & \text{if } |\xi|\leq\frac45\,. \end{cases} \tag{13.55}$$

By the definition of the cut-off function χ_0 in (13.47), recalling (13.26), (13.19), (13.54), and applying estimates (2.42), (12.26), (13.20), the Fourier multiplier $\frac{\chi_0(\xi)}{\partial_\xi\Lambda_m(\xi)}$ is a symbol in $S^{\frac12}$ and satisfies

$$\Big|\mathrm{Op}\Big(\frac{\chi_0(\xi)}{\partial_\xi \Lambda_m(\xi)}\Big)\Big|^{k_0,\gamma}_{\frac12,s,\alpha} \lesssim_{M,\alpha} 1\,, \quad \Big|\Delta_{12}\mathrm{Op}\Big(\frac{\chi_0(\xi)}{\partial_\xi \Lambda_m(\xi)}\Big)\Big|_{\frac12,s_1,\alpha} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+1)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m)}(M,\alpha)}\,. \tag{13.56}$$

By (13.54), the function $w_m^{(1)}(x,\xi)$ is a well-defined symbol in $S^{-\frac{m}{2}+\frac12}$.

In both cases $m=1$ and $m\geq 2$, we have eliminated the terms of order $-\frac{m}{2}$ from the right hand side of (13.48).

Lemma 13.5. *The operators $W_m^{(1)}$ defined in* (13.40)*,* (13.51) *for $m=1$, and* (13.55) *for $m\geq2$, are even and reversibility preserving. For any $M,\alpha>0$ there exists a constant $\aleph_7^{(m,3)}(M,\alpha)>0$ (depending also on k_0,τ,ν), larger than the constant $\aleph_7^{(m,2)}(M,\alpha)$ appearing in Lemma 13.4, such that, if* (7.7) *holds with $\mu_0\geq\aleph_7^{(m,3)}(M,\alpha)$, then for any $s\geq s_0$*

$$|\mathrm{Op}(w_m^{(1)})|^{k_0,\gamma}_{-\frac{m}{2}+\frac12,s,\alpha} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\aleph_7^{(m,3)}(M,\alpha)}\big) \tag{13.57}$$

$$|\Delta_{12}\mathrm{Op}(w_m^{(1)})|_{-\frac{m}{2}+\frac12,s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m,3)}(M,\alpha)}\,. \tag{13.58}$$

As a consequence, the transformation $W_m^{(1)}$ is invertible and

$$|(W_m^{(1)})^{\pm1} - \mathrm{Id}|^{k_0,\gamma}_{0,s,\alpha} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)}\big(1+\|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\aleph_7^{(m,3)}(M,\alpha)}\big) \tag{13.59}$$

$$|\Delta_{12}(W_m^{(1)})^{\pm1}|_{0,s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m,3)}(M,\alpha)}\,. \tag{13.60}$$

Proof. The lemma follows by the explicit expressions in (13.40), (13.51), (13.55), (13.41), by estimates (2.42), (2.44), (2.47), Lemmata 2.10, 2.11, 2.14 and estimates (13.21), (13.22), (13.52), (13.56). ∎

In conclusion, by (13.48), (13.49) and (13.53), we obtain the even and reversible operator

$$L_7^{(m+1)} := (W_m^{(1)})^{-1} L_7^{(m,1)} W_m^{(1)} = \omega \cdot \partial_\varphi + \Lambda_{m+1}(D) + P_7^{(m+1)} \tag{13.61}$$

where

$$\begin{aligned} \Lambda_{m+1}(D) &:= \Lambda_m(D) + \langle p_7^{(m)} \rangle_{\varphi,x}(D) = \mathtt{im}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}} + r_{m+1}(D)\,, \\ r_{m+1}(D) &:= r_m(D) + \langle p_7^{(m)} \rangle_{\varphi,x}(D)\,, \end{aligned} \tag{13.62}$$

and

$$\begin{aligned} P_7^{(m+1)} := (W_m^{(1)})^{-1} \Big\{ & \mathtt{r}_2(\Lambda_m, w_m^{(1)})(x,D) + r_{\langle p_7^{(m)} \rangle_\varphi, w_m^{(1)}}(x,D) - r_{w_m^{(1)}, \langle p_7^{(m)} \rangle_{\varphi,x}}(x,D) + P_7^{(m,1)} W_m^{(1)} \\ & + \chi_{(m \geq 2)} \mathrm{Op}\Big(\chi_0(\xi) \big(\langle p_7^{(m)} \rangle_\varphi(x,\xi) - \langle p_7^{(m)} \rangle_{\varphi,x}(\xi) \big) w_m^{(1)}(x,\xi) \Big) \\ & + \mathrm{Op}\Big((1-\chi_0(\xi)) \big(\langle p_7^{(m)} \rangle_\varphi(x,\xi) - \langle p_7^{(m)} \rangle_{\varphi,x}(\xi) \big) \big(1 + w_m^{(1)}(x,\xi) \big) \Big) \Big\} \end{aligned} \tag{13.63}$$

with $\chi_{(m\geq 2)}$ equal to 1 if $m \geq 2$, and zero otherwise.

Lemma 13.6. *The operators $\Lambda_{m+1}(D)$, $r_{m+1}(D)$, $P_7^{(m+1)}$ are even and reversible. For any $M, \alpha > 0$ there exists a constant $\aleph_7^{(m+1)}(M,\alpha) > 0$ (depending also on k_0, τ, ν), larger than the constant $\aleph_7^{(m,3)}(M,\alpha)$ appearing in Lemma 13.5, such that, if (7.7) holds with $\mu_0 \geq \aleph_7^{(m+1)}(M,\alpha)$, then for any $s \geq s_0$*

$$|r_{m+1}(D)|_{-\frac{1}{2},s,\alpha}^{k_0,\gamma} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+2)}\,, \quad |\Delta_{12} r_{m+1}(D)|_{-\frac{1}{2},s_1,\alpha} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+2)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(m+1)}(M,\alpha)} \tag{13.64}$$

$$|P_7^{(m+1)}|_{-\frac{m}{2}-\frac{1}{2},s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)} \big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(m+1)}(M,\alpha)}^{k_0,\gamma}\big)\,, \tag{13.65}$$

$$|\Delta_{12} P_7^{(m+1)}|_{-\frac{m}{2}-\frac{1}{2},s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(m+1)}(M,\alpha)}\,. \tag{13.66}$$

Proof. Since the operator $\langle p_7^{(m)} \rangle_\varphi(x,D)$ is even and reversible by Lemma 13.4, the average $\langle p_7^{(m)} \rangle_{\varphi,x}(D)$ defined in (13.41) is even and reversible as well (we use Remark 2.23). Since $r_m(D)$, $\Lambda_m(D)$ are even and reversible by the inductive claim, then also $r_{m+1}(D)$, $\Lambda_{m+1}(D)$ defined in (13.62) are even and reversible.

Estimates (13.64)-(13.66) for $r_{m+1}(D)$ and $P_7^{(m+1)}$ defined respectively in (13.62) and (13.63) follow by the explicit expressions of $\langle p_7^{(m)} \rangle_{\varphi,x}(\xi)$ in (13.41) and $w_m^{(1)}$ in (13.51) and (13.55) (for $m=1$ and $m \geq 2$ respectively), by applying (2.44), (2.42), (13.59)-(13.60), (13.38)-(13.39), (2.47), Lemmata 2.10, 2.11, and the inductive estimates (13.20)-(13.23). □

Thus, the proof of the inductive claims (13.20)-(13.25) is complete.

13.2.3 Conclusion of the reduction of $L_7^{(1)}$

Composing all the previous transformations, we obtain the even and reversibility preserving map

$$W := W_0 \circ W_1^{(0)} \circ W_1^{(1)} \circ \ldots \circ W_{2M-1}^{(0)} \circ W_{2M-1}^{(1)}\,, \tag{13.67}$$

where W_0 is defined in (13.3) and for $m = 1, \ldots, 2M-1$, $W_m^{(0)}, W_m^{(1)}$ are defined in (13.27), (13.40). The order M will be fixed in (15.16). By (13.18), (13.19), (13.25) at $m = 2M$, the operator L_7 in (13.2) is conjugated, for all $\omega \in \mathtt{DC}(\gamma,\tau)$, to the even and reversible operator

$$L_8 := L_7^{(2M)} = W^{-1} L_7 W = \omega \cdot \partial_\varphi + \Lambda_{2M}(D) + P_{2M} \tag{13.68}$$

where $P_{2M} := P_7^{(2M)} \in OPS^{-M}$ and

$$\Lambda_{2M}(D) = \mathtt{im}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}} + r_{2M}(D)\,, \quad r_{2M}(D) \in OPS^{-\frac{1}{2}}\,. \tag{13.69}$$

Lemma 13.7. *Assume* (7.7) *with* $\mu_0 \geq \aleph_7^{(2M)}(M,0)$. *Then, for any* $s \geq s_0$, *the following estimates hold:*

$$|r_{2M}(D)|_{-\frac{1}{2},s,0}^{k_0,\gamma} \lesssim_M \varepsilon\gamma^{-(2M+1)}, \quad |\Delta_{12} r_{2M}(D)|_{-\frac{1}{2},s_1,0} \lesssim_M \varepsilon\gamma^{-(2M+1)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(2M)}(M,0)}, \tag{13.70}$$

$$|P_{2M}|_{-M,s,0}^{k_0,\gamma} \lesssim_{M,s} \varepsilon\gamma^{-2(M+1)}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(2M)}(M,0)}^{k_0,\gamma}\big), \tag{13.71}$$

$$|\Delta_{12} P_{2M}|_{-M,s_1,0} \lesssim_{M,s_1} \varepsilon\gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(2M)}(M,0)}, \tag{13.72}$$

$$|W^{\pm 1} - \mathrm{Id}|_{0,s,0}^{k_0,\gamma} \lesssim_{M,s} \varepsilon\gamma^{-2(M+1)}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(2M)}(M,0)}^{k_0,\gamma}\big), \tag{13.73}$$

$$|\Delta_{12} W^{\pm 1}|_{0,s_1,0} \lesssim_{M,s_1} \varepsilon\gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(2M)}(M,0)}. \tag{13.74}$$

Proof. Estimates (13.70), (13.71), (13.72) follow by (13.20), (13.21), (13.22) applied for $m = 2M$. Estimates (13.73)-(13.74) for the map W defined in (13.67), and its inverse W^{-1}, follow by (13.15), (13.34), (13.35), (13.59), (13.60), applying the composition estimate (2.45) (with $m = m' = \alpha = 0$). ☐

Since $\Lambda_{2M}(D)$ is even and reversible, we have that

$$\Lambda_{2M}(\xi), r_{2M}(\xi) \in \mathrm{i}\mathbb{R} \qquad \text{and} \qquad \Lambda_{2M}(\xi) = \Lambda_{2M}(-\xi), \ r_{2M}(\xi) = r_{2M}(-\xi). \tag{13.75}$$

In conclusion, we write the even and reversible operator L_8 in (13.68) as

$$L_8 = \omega \cdot \partial_\varphi + \mathrm{i} D_8 + P_{2M} \tag{13.76}$$

where D_8 is the diagonal operator

$$D_8 := -\mathrm{i}\Lambda_{2M}(D) := \mathrm{diag}_{j\in\mathbb{Z}}(\mu_j), \qquad \mu_j := \mathtt{m}_{\frac{1}{2}} |j|^{\frac{1}{2}} \tanh(\mathtt{h}|j|)^{\frac{1}{2}} + r_j, \quad r_j := -\mathrm{i}\, r_{2M}(j), \tag{13.77}$$

$$\mu_j, r_j \in \mathbb{R}, \qquad \mu_j = \mu_{-j}, \ r_j = r_{-j}, \quad \forall j \in \mathbb{Z}, \tag{13.78}$$

with $r_j \in \mathbb{R}$ satisfying, by (13.70),

$$\sup_{j\in\mathbb{Z}} |j|^{\frac{1}{2}} |r_j|^{k_0,\gamma} \lesssim_M \varepsilon\gamma^{-(2M+1)}, \quad \sup_{j\in\mathbb{Z}} |j|^{\frac{1}{2}} |\Delta_{12} r_j| \lesssim_M \varepsilon\gamma^{-(2M+1)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(2M)}(M,0)} \tag{13.79}$$

and $P_{2M} \in OPS^{-M}$ satisfies (13.71)-(13.72).

From now on, we do not need to expand further the operators in decreasing orders and we will only estimate the tame constants of the operators acting on periodic functions (see Definitions 2.25 and 2.30).

Remark 13.8. *In view of Lemma 2.29, the tame constants of* P_{2M} *can be deduced by estimates* (13.71)-(13.72) *of the pseudo-differential norm* $|P_{2M}|_{-M,s,\alpha}$ *with* $\alpha = 0$. *The iterative reduction in decreasing orders performed in the previous sections cannot be set in* $|\ |_{-M,s,0}$ *norms, because each step of the procedure requires some derivatives of symbols with respect to* ξ *(in the remainder of commutators, in the Poisson brackets of symbols, and also in* (13.55)*), and* α *keeps track of the regularity of symbols with respect to* ξ.

13.3 Conjugation of $\mathcal{L}_7$

In the previous subsections 13.1-13.2 we have conjugated the operator L_7 defined in (13.2) to L_8 in (13.68), whose symbol is constant in (φ, x), up to smoothing remainders of order $-M$. Now we conjugate the whole operator $\mathcal{L}_7$ in (13.1) by the real, even and reversibility preserving map

$$\mathcal{W} := \begin{pmatrix} W & 0 \\ 0 & \overline{W} \end{pmatrix} \tag{13.80}$$

where W is defined in (13.67). By (13.68), (13.76) we obtain, for all $\omega \in \mathtt{DC}(\gamma, \tau)$, the real, even and reversible operator

$$\mathcal{L}_8 := \mathcal{W}^{-1} \mathcal{L}_7 \mathcal{W} = \omega \cdot \partial_\varphi + \mathrm{i}\mathcal{D}_8 + \mathrm{i}\Pi_0 + \mathcal{T}_8 \tag{13.81}$$

where $\mathcal{D}_8$ is the diagonal operator

$$\mathcal{D}_8 := \begin{pmatrix} D_8 & 0 \\ 0 & -D_8 \end{pmatrix}, \tag{13.82}$$

with D_8 defined in (13.77), and the remainder $\mathcal{T}_8$ is

$$\mathcal{T}_8 := \mathrm{i}\mathcal{W}^{-1}\Pi_0\mathcal{W} - \mathrm{i}\Pi_0 + \mathcal{W}^{-1}\mathcal{T}_7\mathcal{W} + \mathcal{P}_{2M}, \quad \mathcal{P}_{2M} := \begin{pmatrix} P_{2M} & 0 \\ 0 & \overline{P_{2M}} \end{pmatrix} \tag{13.83}$$

with P_{2M} defined in (13.68). Note that $\mathcal{T}_8$ is defined on the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$. Therefore the operator in the right hand side in (13.81) is defined on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ as well. This defines the extended operator $\mathcal{L}_8$ on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$.

Lemma 13.9. *For any $M > 0$, there exists a constant $\aleph_8(M) > 0$ (depending also on τ, ν, k_0) such that, if (7.7) holds with $\mu_0 \geq \aleph_8(M)$, then for any $s \geq s_0$*

$$|\mathcal{W}^{\pm 1} - \mathrm{Id}|_{0,s,0}^{k_0,\gamma}, \ |\mathcal{W}^* - \mathrm{Id}|_{0,s,0}^{k_0,\gamma} \lesssim_{M,s} \varepsilon\gamma^{-2(M+1)}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_8(M)}^{k_0,\gamma}\big), \tag{13.84}$$

$$|\Delta_{12}\mathcal{W}^{\pm 1}|_{0,s_1,0}, \ |\Delta_{12}\mathcal{W}^*|_{0,s_1,0} \lesssim_{M,s_1} \varepsilon\gamma^{-2(M+1)}\|\Delta_{12}i\|_{s_1+\aleph_8(M)}. \tag{13.85}$$

Let $S > s_0$, $\beta_0 \in \mathbb{N}$, and $M > \frac{1}{2}(\beta_0 + k_0)$. There exists a constant $\aleph_8'(M, \beta_0) > 0$ such that, assuming (7.7) with $\mu_0 \geq \aleph_8'(M, \beta_0)$, for any $m_1, m_2 \geq 0$, with $m_1 + m_2 \leq M - \frac{1}{2}(\beta_0 + k_0)$, for any $\beta \in \mathbb{N}^\nu$, $|\beta| \leq \beta_0$, the operators $\langle D\rangle^{m_1}(\partial_\varphi^\beta \mathcal{T}_8)\langle D\rangle^{m_2}$, $\langle D\rangle^{m_1}\Delta_{12}(\partial_\varphi^\beta \mathcal{T}_8)\langle D\rangle^{m_2}$ are $\mathcal{D}^{k_0}$-tame with tame constants satisfying

$$\mathfrak{M}_{\langle D\rangle^{m_1}(\partial_\varphi^\beta \mathcal{T}_8)\langle D\rangle^{m_2}}(s) \lesssim_{M,S} \varepsilon\gamma^{-2(M+1)}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_8'(M,\beta_0)}\big), \qquad \forall s_0 \leq s \leq S \tag{13.86}$$

$$\|\langle D\rangle^{m_1}\Delta_{12}(\partial_\varphi^\beta \mathcal{T}_8)\langle D\rangle^{m_2}\|_{\mathcal{L}(H^{s_1})} \lesssim_{M,S} \varepsilon\gamma^{-2(M+1)}\|\Delta_{12}i\|_{s_1+\aleph_8'(M,\beta_0)}. \tag{13.87}$$

Proof. Estimates (13.84), (13.85) follow by definition (13.80), by estimates (13.73), (13.74) and using also Lemma 2.13 to estimate the adjoint operator. Let us prove (13.86) (the proof of (13.87) follows by similar arguments). First we analyze the term $\mathcal{W}^{-1}\mathcal{T}_7\mathcal{W}$. Let $m_1, m_2 \geq 0$, with $m_1 + m_2 \leq M - \frac{1}{2}(\beta_0 + k_0)$ and $\beta \in \mathbb{N}^\nu$ with $|\beta| \leq \beta_0$. Arguing as in the proof of Lemma 12.4, we have to analyze, for any $\beta_1, \beta_2, \beta_3 \in \mathbb{N}^\nu$ with $\beta_1 + \beta_2 + \beta_3 = \beta$, the operator

$$(\partial_\varphi^{\beta_1}\mathcal{W}^{-1})(\partial_\varphi^{\beta_2}\mathcal{T}_7)(\partial_\varphi^{\beta_3}\mathcal{W}).$$

We write

$$\begin{aligned} &\langle D\rangle^{m_1}(\partial_\varphi^{\beta_1}\mathcal{W}^{-1})(\partial_\varphi^{\beta_2}\mathcal{T}_7)(\partial_\varphi^{\beta_3}\mathcal{W})\langle D\rangle^{m_2} \\ &= \Big(\langle D\rangle^{m_1}\partial_\varphi^{\beta_1}\mathcal{W}\langle D\rangle^{-m_1}\Big) \circ \Big(\langle D\rangle^{m_1}\partial_\varphi^{\beta_2}\mathcal{T}_7\langle D\rangle^{m_2}\Big) \circ \Big(\langle D\rangle^{-m_2}\partial_\varphi^{\beta_3}\mathcal{W}\langle D\rangle^{m_2}\Big). \end{aligned} \tag{13.88}$$

For any $m \geq 0$, $\beta \in \mathbb{N}^\nu$, $|\beta| \leq \beta_0$, by (2.76), (2.42), (2.47), (2.45), one has

$$\mathfrak{M}_{\langle D\rangle^m(\partial_\varphi^\beta \mathcal{W}^{\pm 1})\langle D\rangle^{-m}}(s) \lesssim_s |\langle D\rangle^m(\partial_\varphi^\beta \mathcal{W}^{\pm 1})\langle D\rangle^{-m}|_{0,s,0}^{k_0,\gamma} \lesssim_s |\partial_\varphi^\beta \mathcal{W}^{\pm 1}|_{0,s+m,0}^{k_0,\gamma} \lesssim_s |\mathcal{W}^{\pm 1}|_{0,s+\beta_0+m,0}^{k_0,\gamma}$$

and $|\mathcal{W}^{\pm 1}|_{0,s+\beta_0+m,0}^{k_0,\gamma}$ can be estimated by using (13.84). The estimate of (13.88) then follows by (12.41) and Lemma 2.27. The tame estimate of $\langle D\rangle^{m_1}\partial_\varphi^\beta \mathcal{P}_{2M}\langle D\rangle^{m_2}$ follows by (2.76), (13.71), (13.72). The tame estimate of the term $\mathrm{i}\langle D\rangle^{m_1}\partial_\varphi^\beta\big(\mathcal{W}^{-1}\Pi_0\mathcal{W} - \Pi_0\big)\langle D\rangle^{m_2}$ follows by Lemma 2.39 (applied with $A = \mathcal{W}^{-1}$ and $B = \mathcal{W}$) and (2.76), (13.84), (13.85). □

14 Conclusion: reduction of $\mathcal{L}_\omega$ up to smoothing operators

By Sections 7-13, for all $\lambda = (\omega, \mathtt{h}) \in \mathtt{DC}(\gamma, \tau) \times [\mathtt{h}_1, \mathtt{h}_2]$ the real, even and reversible operator $\mathcal{L}$ in (7.6) is conjugated to the real, even and reversible operator $\mathcal{L}_8$ defined in (13.81), namely

$$\mathcal{P}^{-1}\mathcal{L}\mathcal{P} = \mathcal{L}_8 = \omega \cdot \partial_\varphi + \mathrm{i}\mathcal{D}_8 + \mathrm{i}\Pi_0 + \mathcal{T}_8, \tag{14.1}$$

where $\mathcal{P}$ is the real, even and reversibility preserving map

$$\mathcal{P} := \mathcal{Z}\mathcal{B}\mathcal{A}\mathcal{M}_2\mathcal{M}_3\mathcal{C}\mathbf{\Phi}_M\mathbf{\Phi}\mathcal{W}\,. \tag{14.2}$$

Moreover, as already noticed below (13.83), the operator $\mathcal{L}_8$ is defined on the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$.

Now we deduce a similar conjugation result for the projected linearized operator $\mathcal{L}_\omega$ defined in (6.27), which acts on the normal subspace $H^\perp_{\mathbb{S}^+}$, whose relation with $\mathcal{L}$ is stated in (7.5). The operator $\mathcal{L}_\omega$ is even and reversible as stated in Lemma 7.1.

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 $\Pi^\perp_{\mathbb{S}_0} := \mathrm{Id} - \Pi_{\mathbb{S}_0}$. We also denote $H^\perp_{\mathbb{S}_0} := \Pi^\perp_{\mathbb{S}_0} L^2(\mathbb{T})$ and $H^s_\perp := H^s(\mathbb{T}^{\nu+1}) \cap H^\perp_{\mathbb{S}_0}$.

Lemma 14.1. *Let $M > 0$. There exists a constant $\sigma_M > 0$ (depending also on k_0, τ, ν) such that, assuming (7.7) with $\mu_0 \geq \sigma_M$, the following holds: for any $s > s_0$ there exists a constant $\delta(s) > 0$ such that, if $\varepsilon\gamma^{-2(M+1)} \leq \delta(s)$, then the operator*

$$\mathcal{P}_\perp := \Pi^\perp_{\mathbb{S}_0} \mathcal{P} \Pi^\perp_{\mathbb{S}_0} \tag{14.3}$$

is invertible and for each family of functions $h := h(\lambda) \in H^{s+\sigma_M}_\perp \times H^{s+\sigma_M}_\perp$ it satisfies

$$\|\mathcal{P}^{\pm 1}_\perp h\|^{k_0,\gamma}_s \lesssim_{M,s} \|h\|^{k_0,\gamma}_{s+\sigma_M} + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma_M} \|h\|^{k_0,\gamma}_{s_0+\sigma_M}\,, \tag{14.4}$$

$$\|(\Delta_{12}\mathcal{P}^{\pm 1}_\perp) h\|_{s_1} \lesssim_{M,s_1} \varepsilon\gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\sigma_M} \|h\|_{s_1+1}\,. \tag{14.5}$$

The operator $\mathcal{P}_\perp$ is real, even and reversibility preserving. The operators $\mathcal{P}, \mathcal{P}^{-1}$ also satisfy (14.4), (14.5).

Proof. By applying (2.77) together with (7.15), (8.29), (9.10), (10.7), (10.15), (2.67), (11.38), (12.37), (13.84) one has that

$$\|Ah\|^{k_0,\gamma}_s \lesssim_s \|h\|^{k_0,\gamma}_{s+\mu_M} + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu_M} \|h\|^{k_0,\gamma}_{s_0+\mu_M}\,, \quad A \in \{\mathcal{Z}^{\pm 1}, \mathcal{B}^{\pm 1}, \mathcal{A}^{\pm 1}, \mathcal{M}_2^{\pm 1}, \mathcal{M}_3^{\pm 1}, \mathcal{C}^{\pm 1}, \mathbf{\Phi}_M^{\pm 1}, \mathbf{\Phi}^{\pm 1}, \mathcal{W}^{\pm 1}\}\,,$$

for some $\mu_M > 0$. Then by the definition (14.2) of $\mathcal{P}$, by composition, one gets that $\|\mathcal{P}^{\pm 1} h\|^{k_0,\gamma}_s \lesssim_{M,s} \|h\|^{k_0,\gamma}_{s+\sigma_M} + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma_M} \|h\|^{k_0,\gamma}_{s_0+\sigma_M}$ for some constant $\sigma_M > 0$ larger than $\mu_M > 0$, thus $\mathcal{P}^{\pm 1}$ satisfy (14.4). In order to prove that $\mathcal{P}_\perp$ is invertible, it is sufficient to prove that $\Pi_{\mathbb{S}_0} \mathcal{P} \Pi_{\mathbb{S}_0}$ is invertible, and argue as in the proof of Lemma 9.4 in [1], or Section 8.1 of [8]. This follows by a perturbative argument, for $\varepsilon\gamma^{-2(M+1)}$ small, using that $\Pi_{\mathbb{S}_0}$ is a finite dimensional projector. The proof of (14.5) follows similarly by using (7.18), (8.31), (9.10), (10.19), (11.39), (12.38), (13.85). ☐

Finally, for all $\lambda = (\omega, \mathtt{h}) \in \mathtt{DC}(\gamma, \tau) \times [\mathtt{h}_1, \mathtt{h}_2]$, the operator $\mathcal{L}_\omega$ defined in (6.27) is conjugated to

$$\mathcal{L}_\perp := \mathcal{P}^{-1}_\perp \mathcal{L}_\omega \mathcal{P}_\perp = \Pi^\perp_{\mathbb{S}_0} \mathcal{L}_8 \Pi^\perp_{\mathbb{S}_0} + R_M \tag{14.6}$$

where

$$R_M := \mathcal{P}^{-1}_\perp \Pi^\perp_{\mathbb{S}_0} \big(\mathcal{P} \Pi_{\mathbb{S}_0} \mathcal{L}_8 \Pi^\perp_{\mathbb{S}_0} - \mathcal{L} \Pi_{\mathbb{S}_0} \mathcal{P} \Pi^\perp_{\mathbb{S}_0} + \varepsilon R \mathcal{P}_\perp \big) \tag{14.7}$$

$$= \mathcal{P}^{-1}_\perp \Pi^\perp_{\mathbb{S}_0} \mathcal{P} \Pi_{\mathbb{S}_0} \mathcal{T}_8 \Pi^\perp_{\mathbb{S}_0} + \mathcal{P}^{-1}_\perp \Pi^\perp_{\mathbb{S}_0} J \partial_u \nabla_u H(T_\delta(\varphi)) \Pi_{\mathbb{S}_0} \mathcal{P} \Pi^\perp_{\mathbb{S}_0} + \varepsilon \mathcal{P}^{-1}_\perp \Pi^\perp_{\mathbb{S}_0} R \mathcal{P}_\perp \tag{14.8}$$

is a finite dimensional operator. To prove (14.6)-(14.7) we first use (7.5) and (14.3) to get $\mathcal{L}_\omega \mathcal{P}_\perp = \Pi^\perp_{\mathbb{S}_0}(\mathcal{L} + \varepsilon R) \Pi^\perp_{\mathbb{S}_0} \mathcal{P} \Pi^\perp_{\mathbb{S}_0}$, then we use (14.1) to get $\Pi^\perp_{\mathbb{S}_0} \mathcal{L} \mathcal{P} \Pi^\perp_{\mathbb{S}_0} = \Pi^\perp_{\mathbb{S}_0} \mathcal{P} \mathcal{L}_8 \Pi^\perp_{\mathbb{S}_0}$, and we also use the decomposition $\mathbb{I}_2 = \Pi_{\mathbb{S}_0} + \Pi^\perp_{\mathbb{S}_0}$. To get (14.8), we use (14.1), (7.5), and we note that $\Pi_{\mathbb{S}_0}\, \omega \cdot \partial_\varphi\, \Pi^\perp_{\mathbb{S}_0} = 0$, $\Pi^\perp_{\mathbb{S}_0}\, \omega \cdot \partial_\varphi\, \Pi_{\mathbb{S}_0} = 0$, and $\Pi_{\mathbb{S}_0} \mathrm{i} \mathcal{D}_8 \Pi^\perp_{\mathbb{S}_0} = 0$, by (13.82) and (13.77).

Lemma 14.2. *The operator R_M has the finite dimensional form (7.3). Moreover, let $S > s_0$ and $M > \frac{1}{2}(\beta_0 + k_0)$. For any $\beta \in \mathbb{N}^\nu$, $|\beta| \leq \beta_0$, there exists a constant $\aleph_9(M, \beta_0) > 0$ (depending also on k_0, τ, ν) such that, if (7.7) holds with $\mu_0 \geq \aleph_9(M, \beta_0)$, then for any $m_1, m_2 \geq 0$, with $m_1 + m_2 \leq M - \frac{1}{2}(\beta_0 + k_0)$, one has that the operators $\langle D \rangle^{m_1} \partial^\beta_\varphi R_M \langle D \rangle^{m_2}$, $\langle D \rangle^{m_1} \partial^\beta_\varphi \Delta_{12} R_M \langle D \rangle^{m_2}$ are $\mathcal{D}^{k_0}$-tame with tame constants*

$$\mathfrak{M}_{\langle D \rangle^{m_1} \partial^\beta_\varphi R_M \langle D \rangle^{m_2}}(s) \lesssim_{M,S} \varepsilon\gamma^{-2(M+1)} \big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\aleph_9(M,\beta_0)}\big)\,, \qquad \forall s_0 \leq s \leq S \tag{14.9}$$

$$\|\langle D \rangle^{m_1} \Delta_{12} \partial^\beta_\varphi R_M \langle D \rangle^{m_2}\|_{\mathcal{L}(H^{s_1})} \lesssim_{M,S} \varepsilon\gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph_9(M,\beta_0)}\,. \tag{14.10}$$

Proof. To prove that the operator R_M has the finite dimensional form (7.3), notice that in the first two terms in (14.8) there is the finite dimensional projector $\Pi_{\mathbb{S}_0}$, that the operator R in the third term in (14.8) already has the finite dimensional form (7.3), and use the property that $\mathcal{P}_\perp(a(\varphi)h) = a(\varphi)\mathcal{P}_\perp h$ for all $h = h(\varphi, x)$ and all $a(\varphi)$ independent of x, see also the proof of Lemma 2.39 (and Lemma 6.30 in [21] and Lemma 8.3 in [8]). To estimate R_M, use (14.4), (14.5) for $\mathcal{P}$, (13.86), (13.87) for $\mathcal{T}_8$, (7.5), (7.6), (7.16), (7.17), (3.5) for $J\partial_u\nabla_u H(T_\delta(\varphi))$, (7.3), (7.4) for R. The term $\Pi_{\mathbb{S}_0}^\perp J\partial_u\nabla_u H(T_\delta(\varphi))\Pi_{\mathbb{S}_0}$ is small because $\Pi_{\mathbb{S}_0}^\perp \begin{pmatrix} 0 & -D\tanh(\mathtt{h}D) \\ 1 & 0 \end{pmatrix} \Pi_{\mathbb{S}_0}$ is zero. □

By (14.6) and (13.81) we get

$$\mathcal{L}_\perp = \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_\perp + \mathcal{R}_\perp \tag{14.11}$$

where $\mathbb{I}_\perp$ denotes the identity map of $H_{\mathbb{S}_0}^\perp$ (acting on scalar functions u, as well as on pairs $(u, \bar{u})$ in a diagonal manner),

$$\mathcal{D}_\perp := \begin{pmatrix} D_\perp & 0 \\ 0 & -D_\perp \end{pmatrix}, \qquad D_\perp := \Pi_{\mathbb{S}_0}^\perp D_8 \Pi_{\mathbb{S}_0}^\perp\,, \tag{14.12}$$

and $\mathcal{R}_\perp$ is the operator

$$\mathcal{R}_\perp := \Pi_{\mathbb{S}_0}^\perp \mathcal{T}_8 \Pi_{\mathbb{S}_0}^\perp + R_M\,, \quad \mathcal{R}_\perp = \begin{pmatrix} \mathcal{R}_{\perp,1} & \mathcal{R}_{\perp,2} \\ \overline{\mathcal{R}_{\perp,2}} & \overline{\mathcal{R}_{\perp,1}} \end{pmatrix}. \tag{14.13}$$

The operator $\mathcal{R}_\perp$ in (14.13) is defined for all $\lambda = (\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, because $\mathcal{T}_8$ in (13.83) and the operator in the right hand side of (14.8) are defined on the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$. As a consequence, the right hand side of (14.11) extends the definition of $\mathcal{L}_\perp$ to $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$. We still denote the extended operator by $\mathcal{L}_\perp$.

In conclusion, we have obtained the following proposition.

Proposition 14.3. (Reduction of $\mathcal{L}_\omega$ up to smoothing remainders) *For all $\lambda = (\omega, \mathtt{h}) \in \mathtt{DC}(\gamma, \tau) \times [\mathtt{h}_1, \mathtt{h}_2]$, the operator $\mathcal{L}_\omega$ in* (7.5) *is conjugated via* (14.6) *to the real, even and reversible operator $\mathcal{L}_\perp$. For all $\lambda \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, the extended operator $\mathcal{L}_\perp$ defined by the right hand side of* (14.11) *has the form*

$$\mathcal{L}_\perp = \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_\perp + \mathcal{R}_\perp \tag{14.14}$$

where $\mathcal{D}_\perp$ is the diagonal operator

$$\mathcal{D}_\perp := \begin{pmatrix} D_\perp & 0 \\ 0 & -D_\perp \end{pmatrix}, \qquad D_\perp = \mathrm{diag}_{j \in \mathbb{S}_0^c}\, \mu_j\,, \qquad \mu_{-j} = \mu_j\,, \tag{14.15}$$

with eigenvalues μ_j, defined in (13.77)*, given by*

$$\mu_j = \mathtt{m}_{\frac{1}{2}} |j|^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h}|j|) + r_j \in \mathbb{R}\,, \qquad r_{-j} = r_j\,, \tag{14.16}$$

where $\mathtt{m}_{\frac{1}{2}}, r_j \in \mathbb{R}$ satisfy (12.26), (13.79)*. The operator $\mathcal{R}_\perp$ defined in* (14.13) *is real, even and reversible.*

Let $S > s_0$, $\beta_0 \in \mathbb{N}$, and $M > \frac{1}{2}(\beta_0 + k_0)$. There exists a constant $\aleph(M, \beta_0) > 0$ (depending also on k_0, τ, ν) such that, assuming (7.7) *with $\mu_0 \geq \aleph(M, \beta_0)$, for any $m_1, m_2 \geq 0$, with $m_1 + m_2 \leq M - \frac{1}{2}(\beta_0 + k_0)$, for any $\beta \in \mathbb{N}^\nu$, $|\beta| \leq \beta_0$, the operators $\langle D\rangle^{m_1} \partial_\varphi^\beta \mathcal{R}_\perp \langle D\rangle^{m_2}$, $\langle D\rangle^{m_1} \partial_\varphi^\beta \Delta_{12} \mathcal{R}_\perp \langle D\rangle^{m_2}$ are $\mathcal{D}^{k_0}$-tame with tame constants satisfying*

$$\mathfrak{M}_{\langle D\rangle^{m_1} \partial_\varphi^\beta \mathcal{R}_\perp \langle D\rangle^{m_2}}(s) \lesssim_{M,S} \varepsilon \gamma^{-2(M+1)} \big(1 + \|\mathfrak{I}_0\|_{s+\aleph(M,\beta_0)}^{k_0,\gamma}\big)\,, \qquad \forall s_0 \leq s \leq S \tag{14.17}$$

$$\|\langle D\rangle^{m_1} \Delta_{12} \partial_\varphi^\beta \mathcal{R}_\perp \langle D\rangle^{m_2}\|_{\mathcal{L}(H^{s_1})} \lesssim_{M,S} \varepsilon \gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph(M,\beta_0)}\,. \tag{14.18}$$

Proof. Estimates (14.17)-(14.18) for the term $\Pi_{\mathbb{S}_0}^\perp \mathcal{T}_8 \Pi_{\mathbb{S}_0}^\perp$ in (14.13) follow directly by (13.86), (13.87). Estimates (14.17)-(14.18) for R_M are (14.9)-(14.10). □

15 Almost-diagonalization and invertibility of $\mathcal{L}_\omega$

In Proposition 14.3 we obtained the operator $\mathcal{L}_\perp = \mathcal{L}_\perp(\varphi)$ in (14.14) which is diagonal up to the smoothing operator $\mathcal{R}_\perp$. In this section we implement a diagonalization KAM iterative scheme to reduce the size of the non-diagonal term $\mathcal{R}_\perp$.

We first replace the operator $\mathcal{L}_\perp$ in (14.14) with the operator $\mathcal{L}_\perp^{sym}$ defined in (15.1) below, which coincides with $\mathcal{L}_\perp$ on the subspace of functions even in x, see Lemma 15.1. We define the linear operator $\mathcal{L}_\perp^{sym}$, acting on $H_{\mathbb{S}_0}^\perp$, as

$$\mathcal{L}_\perp^{sym} := \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_\perp + \mathcal{R}_\perp^{sym}, \quad \mathcal{R}_\perp^{sym} := \begin{pmatrix} \mathcal{R}_{\perp,1}^{sym} & \mathcal{R}_{\perp,2}^{sym} \\ \overline{\mathcal{R}_{\perp,2}^{sym}} & \overline{\mathcal{R}_{\perp,1}^{sym}} \end{pmatrix}, \tag{15.1}$$

where $\mathcal{R}_{\perp,i}^{sym}$, $i = 1,2$, are defined by their matrix entries

$$(\mathcal{R}_{\perp,i}^{sym})_j^{j'}(\ell) := \begin{cases} (\mathcal{R}_{\perp,i})_j^{j'}(\ell) + (\mathcal{R}_{\perp,i})_j^{-j'}(\ell) & \text{if } jj' > 0, \\ 0 & \text{if } jj' < 0, \end{cases} \qquad j,j' \in \mathbb{S}_0^c, \qquad i = 1,2, \tag{15.2}$$

and $\mathcal{R}_{\perp,i}$, $i = 1,2$ are introduced in (14.13). Note that, in particular, $(\mathcal{R}_{\perp,i}^{sym})_j^{j'} = 0$, $i = 1,2$ on the anti-diagonal $j' = -j$.

Lemma 15.1. *The operator $\mathcal{R}_\perp^{sym}$ coincides with $\mathcal{R}_\perp$ on the subspace of functions even(x) in $H_{\mathbb{S}_0}^\perp \times H_{\mathbb{S}_0}^\perp$, namely*

$$\mathcal{R}_\perp h = \mathcal{R}_\perp^{sym} h, \quad \forall h \in H_{\mathbb{S}_0}^\perp \times H_{\mathbb{S}_0}^\perp, \quad h = h(\varphi,x) = \text{even}(x). \tag{15.3}$$

$\mathcal{R}_\perp^{sym}$ is real, even and reversible, and it satisfies the same bounds (14.17), (14.18) as $\mathcal{R}_\perp$.

Proof. For any function $h \in H_{\mathbb{S}_0}^\perp$ that is even(x), for $i = 1,2$, by (15.2) one has

$$\begin{aligned} \mathcal{R}_{\perp,i}^{sym} h(x) &= \sum_{j,j' \in \mathbb{S}_0^c} (\mathcal{R}_{\perp,i}^{sym})_j^{j'} h_{j'} e^{\mathrm{i}jx} = \sum_{jj'>0} [(\mathcal{R}_{\perp,i})_j^{j'} + (\mathcal{R}_{\perp,i})_j^{-j'}] h_{j'} e^{\mathrm{i}jx} \\ &= \sum_{\substack{j>0 \\ j'>0}} (\mathcal{R}_{\perp,i})_j^{j'} h_{j'} e^{\mathrm{i}jx} + \sum_{\substack{j>0 \\ j'>0}} (\mathcal{R}_{\perp,i})_j^{-j'} h_{j'} e^{\mathrm{i}jx} + \sum_{\substack{j<0 \\ j'<0}} (\mathcal{R}_{\perp,i})_j^{j'} h_{j'} e^{\mathrm{i}jx} + \sum_{\substack{j<0 \\ j'<0}} (\mathcal{R}_{\perp,i})_j^{-j'} h_{j'} e^{\mathrm{i}jx} \\ &= \sum_{\substack{j>0 \\ j'>0}} (\mathcal{R}_{\perp,i})_j^{j'} h_{j'} e^{\mathrm{i}jx} + \sum_{\substack{j>0 \\ j'<0}} (\mathcal{R}_{\perp,i})_j^{j'} h_{j'} e^{\mathrm{i}jx} + \sum_{\substack{j<0 \\ j'<0}} (\mathcal{R}_{\perp,i})_j^{j'} h_{j'} e^{\mathrm{i}jx} + \sum_{\substack{j<0 \\ j'>0}} (\mathcal{R}_{\perp,i})_j^{j'} h_{j'} e^{\mathrm{i}jx} \qquad (15.4) \\ &= \sum_{j,j' \in \mathbb{S}_0^c} (\mathcal{R}_{\perp,i})_j^{j'} h_{j'} e^{\mathrm{i}jx} = \mathcal{R}_{\perp,i} h(x) \end{aligned}$$

where to get (15.4) we have used that $h_{-j'} = h_{j'}$ in the second and fourth sum.

The operator $\mathcal{R}_\perp^{sym}$ is real by (15.1), it is even by (15.3) because $\mathcal{R}_\perp$ is even, and it is reversible by (15.2) and (2.70). Using definition (15.2), the fact that $\mathcal{R}_\perp$ is an even operator, and (2.65), we deduce that

$$(\mathcal{R}_{\perp,i}^{sym})_{-j}^{-j'} = (\mathcal{R}_{\perp,i}^{sym})_j^{j'} \quad \forall j,j' \in \mathbb{S}_0^c. \tag{15.5}$$

Moreover, using (15.2) and (15.5), one proves that for all $n \in \mathbb{S}_0^c$, $n > 0$,

$$\mathcal{R}_{\perp,i}^{sym}[\cos(nx)] = \sum_{j>0} (\mathcal{R}_{\perp,i}^{sym})_j^n \cos(jx) = \mathcal{R}_{\perp,i}[\cos(nx)], \qquad \mathcal{R}_{\perp,i}^{sym}[\sin(nx)] = \sum_{j>0} (\mathcal{R}_{\perp,i}^{sym})_j^n \sin(jx) \tag{15.6}$$

where $(\mathcal{R}_{\perp,i}^{sym})_j^n$ are the matrix elements defined in (15.2).

Finally we consider the decomposition $L^2(\mathbb{T}) := L^2_{\text{even}} \oplus L^2_{\text{odd}}$, $h = h_{\text{even}} + h_{\text{odd}}$, where h_{even} is even(x) and h_{odd} is odd(x), and we define the isometry $\mathcal{M} : L^2_{\text{odd}} \to L^2_{\text{even}}$, $\mathcal{M}[\sin(nx)] := \cos(nx)$, $n \geq 1$, which preserves all Sobolev norms. Hence, by (15.6),

$$\mathcal{R}_{\perp,i}^{sym} h = \mathcal{R}_{\perp,i}^{sym} h_{\text{even}} + \mathcal{R}_{\perp,i}^{sym} h_{\text{odd}} = \mathcal{R}_{\perp,i} h_{\text{even}} + \mathcal{M}^{-1} \mathcal{R}_{\perp,i} \mathcal{M} h_{\text{odd}} \quad \forall h \in H_{\mathbb{S}_0}^\perp.$$

We deduce that $\|\mathcal{R}_{\perp,i}^{sym} h\|_s \lesssim \|\mathcal{R}_{\perp,i} h\|_s$, and similarly with Whitney norms $\| \ \|_s^{k_0,\gamma}$. ◻

As a starting point of the recursive scheme, we consider the real, even, reversible linear operator $\mathcal{L}_{\perp}^{sym}$ in (15.1), acting on $H_{\mathbb{S}_0}^{\perp}$, defined for all $(\omega, \mathtt{h}) \in \mathbb{R}^{\nu} \times [\mathtt{h}_1, \mathtt{h}_2]$, which we rename

$$\mathcal{L}_0 := \mathcal{L}_{\perp}^{sym} := \omega \cdot \partial_{\varphi} \mathbb{I}_{\perp} + \mathrm{i}\mathcal{D}_0 + \mathcal{R}_0 \,, \qquad \mathcal{D}_0 := \mathcal{D}_{\perp}, \quad \mathcal{R}_0 := \mathcal{R}_{\perp}^{sym} \,, \tag{15.7}$$

with

$$\mathcal{D}_0 := \begin{pmatrix} D_0 & 0 \\ 0 & -D_0 \end{pmatrix} , \qquad D_0 := \mathrm{diag}_{j \in \mathbb{S}_0^c} \, \mu_j^0 \,, \qquad \mu_j^0 := \mathtt{m}_{\frac{1}{2}} |j|^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h}|j|) + r_j \,, \tag{15.8}$$

where $\mathtt{m}_{\frac{1}{2}} := \mathtt{m}_{\frac{1}{2}}(\omega, \mathtt{h}) \in \mathbb{R}$ satisfies (12.26), $r_j := r_j(\omega, \mathtt{h}) \in \mathbb{R}$, $r_j = r_{-j}$ satisfy (13.79), and

$$\mathcal{R}_0 := \begin{pmatrix} R_1^{(0)} & R_2^{(0)} \\ \overline{R}_2^{(0)} & \overline{R}_1^{(0)} \end{pmatrix} , \quad R_i^{(0)} : H_{\mathbb{S}_0}^{\perp} \to H_{\mathbb{S}_0}^{\perp} \,, \qquad i = 1, 2 \,. \tag{15.9}$$

Notation. In this section we shall use the following notation:

1. Given an operator R, the expression $\partial_{\varphi_i}^s \langle D \rangle^{\mathtt{m}} R \langle D \rangle^{\mathtt{m}}$ denotes the operator $\langle D \rangle^{\mathtt{m}} \circ \big(\partial_{\varphi_i}^s R(\varphi) \big) \circ \langle D \rangle^{\mathtt{m}}$. Similarly, $\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} R \langle D \rangle^{\mathtt{m}}$ denotes the operator $\langle D \rangle^{\mathtt{m}} \circ \big(\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} R \big) \circ \langle D \rangle^{\mathtt{m}}$, where $\langle \partial_{\varphi,x} \rangle^{\mathtt{b}}$ is introduced in Definition 2.7.
2. To avoid confusion with the induction index $\nu = 0, 1, 2, \dots$ appearing in Theorem 15.4, we shall denote the cardinality of the set $\mathbb{S}^+$ of tangential sites by $|\mathbb{S}^+|$ (instead of ν, as it was denoted in the previous sections).

The operator $\mathcal{R}_0$ in (15.9) satisfies the following tame estimates, which we verify in Lemma 15.3 below. Define the constants

$$\begin{aligned} &\mathtt{b} := [\mathtt{a}] + 2 \in \mathbb{N} \,, \quad \mathtt{a} := \max\{3\tau_1, \chi(\tau+1)(4\mathtt{d}+1) + 1\} \,, \quad \chi := 3/2 \,, \\ &\tau_1 := \tau(k_0+1) + k_0 + \mathtt{m} \,, \quad \mathtt{m} := \mathtt{d}(k_0+1) + \frac{k_0}{2} \,, \end{aligned} \tag{15.10}$$

where $\mathtt{d} > \frac{3}{4} k_0^*$, by (5.25).

- **(Smallness of $\mathcal{R}_0$).** *The operators*

$$\langle D \rangle^{\mathtt{m}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+1}, \qquad \partial_{\varphi_i}^{s_0} \langle D \rangle^{\mathtt{m}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+1} \,, \qquad \forall i = 1, \dots, |\mathbb{S}^+| \,, \tag{15.11}$$

$$\langle D \rangle^{\mathtt{m}+\mathtt{b}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}, \quad \partial_{\varphi_i}^{s_0+\mathtt{b}} \langle D \rangle^{\mathtt{m}+\mathtt{b}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+\mathtt{b}+1} \,, \tag{15.12}$$

where $\mathtt{m}, \mathtt{b}$ *are defined in* (15.10), *are* $\mathcal{D}^{k_0}$*-tame with tame constants, defined for all* $s_0 \le s \le S$,

$$\mathbb{M}_0(s) := \max_{i=1,\dots,|\mathbb{S}^+|} \big\{ \mathfrak{M}_{\langle D \rangle^{\mathtt{m}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+1}}(s), \mathfrak{M}_{\partial_{\varphi_i}^{s_0} \langle D \rangle^{\mathtt{m}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+1}}(s) \big\} \tag{15.13}$$

$$\mathbb{M}_0(s, \mathtt{b}) := \max_{i=1,\dots,|\mathbb{S}^+|} \big\{ \mathfrak{M}_{\langle D \rangle^{\mathtt{m}+\mathtt{b}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}}(s), \mathfrak{M}_{\partial_{\varphi_i}^{s_0+\mathtt{b}} \langle D \rangle^{\mathtt{m}+\mathtt{b}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}}(s) \big\} \tag{15.14}$$

satisfying

$$\mathfrak{M}_0(s_0, \mathtt{b}) := \max\{\mathbb{M}_0(s_0), \mathbb{M}_0(s_0, \mathtt{b})\} \le C(S) \varepsilon \gamma^{-2(M+1)} \,. \tag{15.15}$$

Remark 15.2. *The condition* $\mathtt{a} \ge \chi(\tau+1)(4\mathtt{d}+1) + 1$ *in* (15.10) *will be used in Section 16 in order to verify inequality* (16.5).

Proposition 14.3 implies that the operator $\mathcal{R}_0 = \mathcal{R}_{\perp}^{sym}$ satisfies the above tame estimates by fixing the constant M large enough (which means performing sufficiently many regularizing steps in Sections 11 and 13), namely

$$M := \Big[2\mathtt{m} + 2\mathtt{b} + 1 + \frac{\mathtt{b} + s_0 + k_0}{2} \Big] + 1 \in \mathbb{N} \tag{15.16}$$

where $[\,\cdot\,]$ denotes the integer part, and $\mathtt{m}$ and $\mathtt{b}$ are defined in (15.10). We also set

$$\mu(\mathtt{b}) := \aleph(M, s_0 + \mathtt{b}) \,, \tag{15.17}$$

where the constant $\aleph(M, s_0 + \mathtt{b})$ is given in Proposition 14.3.

Lemma 15.3. (Tame estimates of $\mathcal{R}_0 := \mathcal{R}_\perp^{sym}$) *Assume* (7.7) *with* $\mu_0 \geq \mu(\mathtt{b})$. *Then the operator* $\mathcal{R}_0 := \mathcal{R}_\perp^{sym}$ *defined in* (15.1), (15.2) *satisfies, for all* $s_0 \leq s \leq S$,

$$\mathfrak{M}_0(s,\mathtt{b}) := \max\{\mathbb{M}_0(s), \mathbb{M}_0(s,\mathtt{b})\} \lesssim_S \varepsilon\gamma^{-2(M+1)}\big(1 + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})}^{k_0,\gamma}\big) \tag{15.18}$$

where $\mathbb{M}_0(s)$, $\mathbb{M}_0(s,\mathtt{b})$ *are defined in* (15.13), (15.14). *In particular* (15.15) *holds. Moreover, for all* $i = 1,\ldots,|\mathbb{S}^+|$, $\beta \in \mathbb{N}$, $\beta \leq s_0 + \mathtt{b}$, *the operators* $\partial_{\varphi_i}^\beta \langle D\rangle^{\mathtt{m}} \Delta_{12}\mathcal{R}_0 \langle D\rangle^{\mathtt{m}+1}$, $\partial_{\varphi_i}^\beta \langle D\rangle^{\mathtt{m}+\mathtt{b}} \Delta_{12}\mathcal{R}_0 \langle D\rangle^{\mathtt{m}+\mathtt{b}+1}$ *satisfy the bounds*

$$\|\partial_{\varphi_i}^\beta \langle D\rangle^{\mathtt{m}} \Delta_{12}\mathcal{R}_0 \langle D\rangle^{\mathtt{m}+1}\|_{\mathcal{L}(H^{s_0})}, \|\partial_{\varphi_i}^\beta \langle D\rangle^{\mathtt{m}+\mathtt{b}} \Delta_{12}\mathcal{R}_0 \langle D\rangle^{\mathtt{m}+\mathtt{b}+1}\|_{\mathcal{L}(H^{s_0})} \lesssim_S \varepsilon\gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_0+\mu(\mathtt{b})} . \tag{15.19}$$

Proof. Estimate (15.18) follows by Lemma 15.1, by (14.17) with $m_1 = \mathtt{m}$, $m_2 = \mathtt{m}+1$ for $\mathbb{M}_0(s)$, with $m_1 = \mathtt{m}+\mathtt{b}$, $m_2 = \mathtt{m}+\mathtt{b}+1$ for $\mathbb{M}_0(s,\mathtt{b})$, and by definitions (15.10), (15.16), (15.17). Estimates (15.19) follow similarly, applying (14.18) with the same choices of m_1, m_2 and with $s_1 = s_0$. □

We perform the almost-reducibility of $\mathcal{L}_0$ along the scale

$$N_{-1} := 1 , \quad N_\nu := N_0^{\chi^\nu} \quad \forall \nu \geq 0 , \quad \chi = 3/2 , \tag{15.20}$$

requiring inductively at each step the second order Melnikov non-resonance conditions in (15.29).

Theorem 15.4. (Almost-reducibility of $\mathcal{L}_0$: KAM iteration) *There exists* $\tau_2 := \tau_2(\tau, |\mathbb{S}^+|) > \tau_1 + \mathtt{a}$ *(where* $\tau_1, \mathtt{a}$ *are defined in* (15.10)*) such that, for all* $S > s_0$, *there are* $N_0 := N_0(S,\mathtt{b}) \in \mathbb{N}$, $\delta_0 := \delta_0(S,\mathtt{b}) \in (0,1)$ *such that, if*

$$\varepsilon\gamma^{-2(M+1)} \leq \delta_0, \quad N_0^{\tau_2}\mathfrak{M}_0(s_0,\mathtt{b})\gamma^{-1} \leq 1 \tag{15.21}$$

(see (15.15)*), then, for all* $n \in \mathbb{N}$, $\nu = 0, 1, \ldots, n$:

$(\mathbf{S1})_\nu$ *There exists a real, even and reversible operator*

$$\mathcal{L}_\nu := \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_\nu + \mathcal{R}_\nu , \quad \mathcal{D}_\nu := \begin{pmatrix} D_\nu & 0 \\ 0 & -D_\nu \end{pmatrix} , \quad D_\nu := \mathrm{diag}_{j\in\mathbb{S}_0^c}\mu_j^\nu , \tag{15.22}$$

defined for all $(\omega,\mathtt{h})$ *in* $\mathbb{R}^{|\mathbb{S}^+|} \times [\mathtt{h}_1,\mathtt{h}_2]$ *where* μ_j^ν *are* k_0 *times differentiable functions of the form*

$$\mu_j^\nu(\omega,\mathtt{h}) := \mu_j^0(\omega,\mathtt{h}) + r_j^\nu(\omega,\mathtt{h}) \in \mathbb{R} \tag{15.23}$$

where μ_j^0 *are defined in* (15.8), *satisfying*

$$\mu_j^\nu = \mu_{-j}^\nu , \quad \textit{i.e. } r_j^\nu = r_{-j}^\nu , \quad |r_j^\nu|^{k_0,\gamma} \leq C(S,\mathtt{b})\varepsilon\gamma^{-2(M+1)}|j|^{-2\mathtt{m}} , \quad \forall j \in \mathbb{S}_0^c \tag{15.24}$$

and, for $\nu \geq 1$,

$$|\mu_j^\nu - \mu_j^{\nu-1}|^{k_0,\gamma} \leq C|j|^{-2\mathtt{m}}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_{\nu-1}\langle D\rangle^{\mathtt{m}}}(s_0) \leq C(S,\mathtt{b})\varepsilon\gamma^{-2(M+1)}|j|^{-2\mathtt{m}}N_{\nu-2}^{-\mathtt{a}} . \tag{15.25}$$

The remainder

$$\mathcal{R}_\nu := \begin{pmatrix} R_1^{(\nu)} & R_2^{(\nu)} \\ \overline{R}_2^{(\nu)} & \overline{R}_1^{(\nu)} \end{pmatrix} \tag{15.26}$$

satisfies

$$(R_1^{(\nu)})_j^{j'}(\ell) = (R_2^{(\nu)})_j^{j'}(\ell) = 0 \quad \forall (\ell, j, j'), \ jj' < 0, \tag{15.27}$$

and it is $\mathcal{D}^{k_0}$*-modulo-tame: more precisely, the operators* $\langle D\rangle^{\mathtt{m}}\mathcal{R}_\nu\langle D\rangle^{\mathtt{m}}$ *and* $\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_\nu\langle D\rangle^{\mathtt{m}}$ *are* $\mathcal{D}^{k_0}$*-modulo-tame and there exists a constant* $C_* := C_*(s_0,\mathtt{b}) > 0$ *such that, for any* $s \in [s_0, S]$,

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_\nu\langle D\rangle^{\mathtt{m}}}(s) \leq \frac{C_*\mathfrak{M}_0(s,\mathtt{b})}{N_{\nu-1}^{\mathtt{a}}} , \quad \mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_\nu\langle D\rangle^{\mathtt{m}}}(s) \leq C_*\mathfrak{M}_0(s,\mathtt{b})N_{\nu-1} . \tag{15.28}$$

Define the sets $\mathtt{\Lambda}_\nu^\gamma$ *by* $\mathtt{\Lambda}_0^\gamma := \mathtt{DC}(2\gamma,\tau)\times[\mathtt{h}_1,\mathtt{h}_2]$, *and, for all* $\nu \geq 1$,

$$\begin{aligned}\mathtt{\Lambda}_\nu^\gamma := \mathtt{\Lambda}_\nu^\gamma(i) := \Big\{ & \lambda = (\omega,\mathtt{h}) \in \mathtt{\Lambda}_{\nu-1}^\gamma \ : \\ & |\omega\cdot\ell + \mu_j^{\nu-1} - \mu_{j'}^{\nu-1}| \geq \gamma j^{-\mathtt{d}} j'^{-\mathtt{d}} \langle\ell\rangle^{-\tau} \quad \forall |\ell|, |j-j'| \leq N_{\nu-1}, \quad j,j' \in \mathbb{N}^+\setminus\mathbb{S}^+, \quad (\ell,j,j') \neq (0,j,j), \\ & |\omega\cdot\ell + \mu_j^{\nu-1} + \mu_{j'}^{\nu-1}| \geq \gamma(\sqrt{j}+\sqrt{j'})\langle\ell\rangle^{-\tau} \quad \forall |\ell|, |j-j'| \leq N_{\nu-1}\,, \quad j,j' \in \mathbb{N}^+\setminus\mathbb{S}^+ \Big\}\,. \end{aligned} \tag{15.29}$$

For $\nu \geq 1$, *there exists a real, even and reversibility preserving map, defined for all* $(\omega,\mathtt{h})$ *in* $\mathbb{R}^{|\mathbb{S}^+|}\times[\mathtt{h}_1,\mathtt{h}_2]$, *of the form*

$$\Phi_{\nu-1} := \mathbb{I}_\perp + \Psi_{\nu-1}\,, \quad \Psi_{\nu-1} := \begin{pmatrix} \Psi_{\nu-1,1} & \Psi_{\nu-1,2} \\ \overline{\Psi}_{\nu-1,2} & \overline{\Psi}_{\nu-1,1} \end{pmatrix} \tag{15.30}$$

such that for all $\lambda = (\omega,\mathtt{h}) \in \mathtt{\Lambda}_\nu^\gamma$ *the following conjugation formula holds:*

$$\mathcal{L}_\nu = \Phi_{\nu-1}^{-1}\mathcal{L}_{\nu-1}\Phi_{\nu-1}\,. \tag{15.31}$$

The operators $\langle D\rangle^{\pm\mathtt{m}}\Psi_{\nu-1}\langle D\rangle^{\mp\mathtt{m}}$ *and* $\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\Psi_{\nu-1}\langle D\rangle^{\mp\mathtt{m}}$ *are* $\mathcal{D}^{k_0}$*-modulo-tame on* $\mathbb{R}^{|\mathbb{S}^+|}\times[\mathtt{h}_1,\mathtt{h}_2]$ *with modulo-tame constants satisfying, for all* $s \in [s_0,S]$, ($\tau_1, \mathtt{a}$ *are defined in* (15.10))

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\pm\mathtt{m}}\Psi_{\nu-1}\langle D\rangle^{\mp\mathtt{m}}}(s) \leq C(s_0,\mathtt{b})\gamma^{-1}N_{\nu-1}^{\tau_1}N_{\nu-2}^{-\mathtt{a}}\mathfrak{M}_0(s,\mathtt{b})\,, \tag{15.32}$$

$$\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\Psi_{\nu-1}\langle D\rangle^{\mp\mathtt{m}}}(s) \leq C(s_0,\mathtt{b})\gamma^{-1}N_{\nu-1}^{\tau_1}N_{\nu-2}\mathfrak{M}_0(s,\mathtt{b})\,, \tag{15.33}$$

$$\mathfrak{M}^\sharp_{\Psi_{\nu-1}}(s) \leq C(s_0,\mathtt{b})\gamma^{-1}N_{\nu-1}^{\tau_1}N_{\nu-2}^{-\mathtt{a}}\mathfrak{M}_0(s,\mathtt{b})\,. \tag{15.34}$$

$(\mathbf{S2})_\nu$ *Let* $i_1(\omega,\mathtt{h})$, $i_2(\omega,\mathtt{h})$ *be such that* $\mathcal{R}_0(i_1)$, $\mathcal{R}_0(i_2)$ *satisfy* (15.15). *Then for all* $(\omega,\mathtt{h}) \in \mathtt{\Lambda}_\nu^{\gamma_1}(i_1)\cap\mathtt{\Lambda}_\nu^{\gamma_2}(i_2)$ *with* $\gamma_1,\gamma_2 \in [\gamma/2, 2\gamma]$, *the following estimates hold*

$$|||\langle D\rangle^{\mathtt{m}}\Delta_{12}\mathcal{R}_\nu\langle D\rangle^{\mathtt{m}}|||_{\mathcal{L}(H^{s_0})} \lesssim_{S,\mathtt{b}} \varepsilon\gamma^{-2(M+1)}N_{\nu-1}^{-\mathtt{a}}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})}, \tag{15.35}$$

$$|||\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\Delta_{12}\mathcal{R}_\nu\langle D\rangle^{\mathtt{m}}|||_{\mathcal{L}(H^{s_0})} \lesssim_{S,\mathtt{b}} \varepsilon\gamma^{-2(M+1)}N_{\nu-1}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})}\,. \tag{15.36}$$

Moreover for $\nu \geq 1$, *for all* $j \in \mathbb{S}_0^c$,

$$\big|\Delta_{12}(r_j^\nu - r_j^{\nu-1})\big| \lesssim_{S,\mathtt{b}} \varepsilon\gamma^{-2(M+1)}|j|^{-2\mathtt{m}}N_{\nu-2}^{-\mathtt{a}}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})}\,, \tag{15.37}$$

$$|\Delta_{12}r_j^\nu| \lesssim_{S,\mathtt{b}} \varepsilon\gamma^{-2(M+1)}|j|^{-2\mathtt{m}}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})}\,. \tag{15.38}$$

$(\mathbf{S3})_\nu$ *Let* i_1, i_2 *be like in* $(\mathbf{S2})_\nu$ *and* $0 < \rho \leq \gamma/2$. *Then*

$$C(S)N_{\nu-1}^{(\tau+1)(4\mathtt{d}+1)}\gamma^{-4\mathtt{d}}\|i_2-i_1\|_{s_0+\mu(\mathtt{b})} \leq \rho \quad \Longrightarrow \quad \mathtt{\Lambda}_\nu^\gamma(i_1) \subseteq \mathtt{\Lambda}_\nu^{\gamma-\rho}(i_2)\,. \tag{15.39}$$

We make some comments:

1. Note that in (15.37)-(15.38) we do not need norms $|\ |^{k_0,\gamma}$. This is the reason why we did not estimate the derivatives with respect to $(\omega,\mathtt{h})$ of the operators $\Delta_{12}\mathcal{R}$ in the previous sections.

2. Since the second Melnikov conditions $|\omega\cdot\ell+\mu_j^{\nu-1}-\mu_{j'}^{\nu-1}| \geq \gamma|j|^{-\mathtt{d}}|j'|^{-\mathtt{d}}\langle\ell\rangle^{-\tau}$ lose regularity both in φ and in x, for the convergence of the reducibility scheme we use the smoothing operators Π_N, defined in (2.24), which regularize in both φ and x. As a consequence, the natural smallness condition to impose at the zero step of the recursion is the one we verify in Lemma 15.6. Thanks to (15.52), to verify such a smallness condition it is sufficient to control the tame constants of the operators (15.12).

3. An important point of Theorem 15.4 is to require bound (15.21) for $\mathfrak{M}_0(s_0,\mathtt{b})$ only in low norm, which is verified in Lemma 15.3. On the other hand, Theorem 15.4 provides the smallness (15.28) of the tame constants $\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_\nu\langle D\rangle^{\mathtt{m}}}(s)$ and proves that $\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_\nu\langle D\rangle^{\mathtt{m}}}(s,\mathtt{b})$, $\nu \geq 0$, do not diverge too much.

Theorem 15.4 implies that the invertible operator

$$\mathcal{U}_n := \Phi_0 \circ \ldots \circ \Phi_{n-1}\,, \quad n \geq 1\,, \tag{15.40}$$

has almost-diagonalized $\mathcal{L}_0$, i.e. (15.45) below holds. As a corollary, we deduce the following theorem.

Theorem 15.5. (Almost-reducibility of $\mathcal{L}_0$) *Assume* (7.7) *with* $\mu_0 \geq \mu(\mathtt{b})$. *Let* $\mathcal{R}_0 = \mathcal{R}_\perp^{sym}$, $\mathcal{L}_0 = \mathcal{L}_\perp^{sym}$ *in* (15.1)-(15.2). *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*

$$N_0^{\tau_2} \varepsilon \gamma^{-(2M+3)} \leq \delta_0 \tag{15.41}$$

holds, where the constant $\tau_2 := \tau_2(\tau, |\mathbb{S}^+|)$ *is defined in Theorem 15.4 and* M *is defined in* (15.16), *then, for all* $n \in \mathbb{N}$, *for all* $\lambda = (\omega, \mathtt{h}) \in \mathbb{R}^{|\mathbb{S}^+|} \times [\mathtt{h}_1, \mathtt{h}_2]$, *the operator* $\mathcal{U}_n$ *in* (15.40) *and its inverse* $\mathcal{U}_n^{-1}$ *are real, even, reversibility preserving, and* $\mathcal{D}^{k_0}$*-modulo-tame, with*

$$\mathfrak{M}^\sharp_{\mathcal{U}_n^{\pm 1} - \mathbb{I}_\perp}(s) \lesssim_S \varepsilon \gamma^{-(2M+3)} N_0^{\tau_1} \big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu(\mathtt{b})}\big) \quad \forall s_0 \leq s \leq S\,, \tag{15.42}$$

where τ_1 *is defined in* (15.10).

The operator $\mathcal{L}_n = \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_n + \mathcal{R}_n$ *defined in* (15.22) *(with* $\nu = n$*) is real, even and reversible. The operator* $\langle D\rangle^{\mathtt{m}} \mathcal{R}_n \langle D\rangle^{\mathtt{m}}$ *is* $\mathcal{D}^{k_0}$*-modulo-tame, with*

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}} \mathcal{R}_n \langle D\rangle^{\mathtt{m}}}(s) \lesssim_S \varepsilon \gamma^{-2(M+1)} N_{n-1}^{-\mathtt{a}} \big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu(\mathtt{b})}\big) \quad \forall s_0 \leq s \leq S\,. \tag{15.43}$$

Moreover, for all $\lambda = (\omega, \mathtt{h})$ *in the set*

$$\Lambda_n^\gamma = \bigcap_{\nu=0}^{n} \Lambda_\nu^\gamma \tag{15.44}$$

defined in (15.29), *the following conjugation formula holds:*

$$\mathcal{L}_n = \mathcal{U}_n^{-1} \mathcal{L}_0 \mathcal{U}_n\,. \tag{15.45}$$

Proof. Assumption (15.21) of Theorem 15.4 holds by (15.18), (7.7) with $\mu_0 \geq \mu(\mathtt{b})$, and (15.41). Estimate (15.43) follows by (15.28) (for $\nu = n$) and (15.18). It remains to prove (15.42). The estimates of $\mathfrak{M}^\sharp_{\Phi_\nu^{\pm1} - \mathbb{I}_\perp}(s)$, $\nu = 0, \ldots, n-1$, are obtained by using (15.34), (15.21) and Lemma 2.33. Then the estimate of $\mathcal{U}_n^{\pm 1} - \mathbb{I}_\perp$ follows as in the proof of Theorem 7.5 in [21], using Lemma 2.32. □

15.1 Proof of Theorem 15.4

PROOF OF $(\mathbf{S1})_0$. The real, even and reversible operator $\mathcal{L}_0$ defined in (15.7)-(15.9) has the form (15.22)-(15.23) for $\nu = 0$ with $r_j^0(\omega, \mathtt{h}) = 0$, and (15.24) holds trivially. Moreover (15.27) is satisfied for $\nu = 0$ by the definition of $\mathcal{R}_0 := \mathcal{R}_\perp^{sym}$ in (15.2). We now prove that also (15.28) for $\nu = 0$ holds:

Lemma 15.6. $\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}} \mathcal{R}_0 \langle D\rangle^{\mathtt{m}}}(s)$, $\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^{\mathtt{m}} \mathcal{R}_0 \langle D\rangle^{\mathtt{m}}}(s) \lesssim_{s_0,\mathtt{b}} \mathfrak{M}_0(s, \mathtt{b})$.

Proof. Let $R \in \{R_1^{(0)}, R_2^{(0)}\}$ and set $\lambda := (\omega, \mathtt{h})$. For any $\alpha, \beta \in \mathbb{N}$, the matrix elements of the operator $\partial_{\varphi_i}^\alpha \langle D\rangle^\beta R \langle D\rangle^{\beta+1}$, $i = 1, \ldots, |\mathbb{S}^+|$, are

$$\mathrm{i}^\alpha (\ell_i - \ell_i')^\alpha \langle j\rangle^\beta R_j^{j'}(\ell - \ell') \langle j'\rangle^{\beta+1}\,.$$

Then, by (2.75) with $\sigma = 0$, and (15.13), (15.14) we have that $\forall |k| \leq k_0$, $s_0 \leq s \leq S$, $\ell' \in \mathbb{Z}^{|\mathbb{S}^+|}$, $j' \in \mathbb{S}_0^c$,

$$\gamma^{2|k|} \sum\nolimits_{\ell,j} \langle \ell, j\rangle^{2s} \langle j\rangle^{2\mathtt{m}} |\partial_\lambda^k R_j^{j'}(\ell - \ell')|^2 \langle j'\rangle^{2(\mathtt{m}+1)} \leq 2\mathbb{M}_0^2(s_0) \langle \ell', j'\rangle^{2s} + 2\mathbb{M}_0^2(s) \langle \ell', j'\rangle^{2s_0}\,, \tag{15.46}$$

$$\gamma^{2|k|} \sum\nolimits_{\ell,j} \langle \ell, j\rangle^{2s} |\ell_i - \ell_i'|^{2s_0} \langle j\rangle^{2\mathtt{m}} |\partial_\lambda^k R_j^{j'}(\ell - \ell')|^2 \langle j'\rangle^{2(\mathtt{m}+1)} \leq 2\mathbb{M}_0^2(s_0) \langle \ell', j'\rangle^{2s} + 2\mathbb{M}_0^2(s) \langle \ell', j'\rangle^{2s_0}\,, \tag{15.47}$$

$$\gamma^{2|k|} \sum\nolimits_{\ell,j} \langle \ell, j\rangle^{2s} \langle j\rangle^{2(\mathtt{m}+\mathtt{b})} |\partial_\lambda^k R_j^{j'}(\ell - \ell')|^2 \langle j'\rangle^{2(\mathtt{m}+\mathtt{b}+1)} \leq 2\mathbb{M}_0^2(s_0, \mathtt{b}) \langle \ell', j'\rangle^{2s} + 2\mathbb{M}_0^2(s, \mathtt{b}) \langle \ell', j'\rangle^{2s_0}\,, \tag{15.48}$$

$$\begin{aligned} &\gamma^{2|k|} \sum\nolimits_{\ell,j} \langle \ell, j\rangle^{2s} |\ell_i - \ell_i'|^{2(s_0+\mathtt{b})} \langle j\rangle^{2(\mathtt{m}+\mathtt{b})} |\partial_\lambda^k R_j^{j'}(\ell - \ell')|^2 \langle j'\rangle^{2(\mathtt{m}+\mathtt{b}+1)} \\ &\qquad \leq 2\mathbb{M}_0^2(s_0, \mathtt{b}) \langle \ell', j'\rangle^{2s} + 2\mathbb{M}_0^2(s, \mathtt{b}) \langle \ell', j'\rangle^{2s_0}\,. \end{aligned} \tag{15.49}$$

Using the inequality $\langle \ell - \ell' \rangle^{2\alpha} \lesssim_\alpha 1 + \max_{i=1,\dots,|\mathbb{S}^+|} |\ell_i - \ell'_i|^{2\alpha}$ for $\alpha = s_0$ and $\alpha = s_0 + \mathtt{b}$, and recalling the definition of $\mathfrak{M}_0(s, \mathtt{b})$ in (15.18), estimates (15.46)-(15.49) imply

$$\gamma^{2|k|} \sum_{\ell,j} \langle \ell, j \rangle^{2s} \langle \ell - \ell' \rangle^{2s_0} \langle j \rangle^{2\mathtt{m}} |\partial_\lambda^k R_j^{j'}(\ell - \ell')|^2 \langle j' \rangle^{2(\mathtt{m}+1)} \lesssim_{\mathtt{b}} \mathfrak{M}_0^2(s_0, \mathtt{b}) \langle \ell', j' \rangle^{2s} + \mathfrak{M}_0^2(s, \mathtt{b}) \langle \ell', j' \rangle^{2s_0} \quad (15.50)$$

$$\gamma^{2|k|} \sum_{\ell,j} \langle \ell, j \rangle^{2s} \langle \ell - \ell' \rangle^{2(s_0+\mathtt{b})} \langle j \rangle^{2(\mathtt{m}+\mathtt{b})} |\partial_\lambda^k R_j^{j'}(\ell - \ell')|^2 \langle j' \rangle^{2(\mathtt{m}+\mathtt{b}+1)}$$
$$\lesssim_{\mathtt{b}} \mathfrak{M}_0^2(s_0, \mathtt{b}) \langle \ell', j' \rangle^{2s} + \mathfrak{M}_0^2(s, \mathtt{b}) \langle \ell', j' \rangle^{2s_0} . \quad (15.51)$$

We now prove that $\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} R \langle D \rangle^{\mathtt{m}}$ is $\mathcal{D}^{k_0}$-modulo-tame. For all $|k| \leq k_0$, by Cauchy-Schwarz inequality and using that

$$\langle \ell - \ell', j - j' \rangle^{\mathtt{b}} \lesssim_{\mathtt{b}} \langle \ell - \ell' \rangle^{\mathtt{b}} \langle j - j' \rangle^{\mathtt{b}} \lesssim_{\mathtt{b}} \langle \ell - \ell' \rangle^{\mathtt{b}} (\langle j \rangle^{\mathtt{b}} + \langle j' \rangle^{\mathtt{b}}) \lesssim_{\mathtt{b}} \langle \ell - \ell' \rangle^{\mathtt{b}} \langle j \rangle^{\mathtt{b}} \langle j' \rangle^{\mathtt{b}} \quad (15.52)$$

we get

$$\begin{aligned}
\| |\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} \partial_\lambda^k R \langle D \rangle^{\mathtt{m}} | h \|_s^2 &\lesssim_{\mathtt{b}} \sum_{\ell,j} \langle \ell, j \rangle^{2s} \Big(\sum_{\ell',j'} |\langle \ell - \ell' \rangle^{\mathtt{b}} \langle j \rangle^{\mathtt{m}+\mathtt{b}} \partial_\lambda^k R_j^{j'}(\ell - \ell') \langle j' \rangle^{\mathtt{m}+\mathtt{b}} | |h_{\ell',j'}| \Big)^2 \\
&\lesssim_{\mathtt{b}} \sum_{\ell,j} \langle \ell, j \rangle^{2s} \Big(\sum_{\ell',j'} \langle \ell - \ell' \rangle^{s_0+\mathtt{b}} \langle j \rangle^{\mathtt{m}+\mathtt{b}} |\partial_\lambda^k R_j^{j'}(\ell - \ell')| \langle j' \rangle^{\mathtt{m}+\mathtt{b}+1} |h_{\ell',j'}| \frac{1}{\langle \ell - \ell' \rangle^{s_0} \langle j' \rangle} \Big)^2 \\
&\lesssim_{s_0,\mathtt{b}} \sum_{\ell,j} \langle \ell, j \rangle^{2s} \sum_{\ell',j'} \langle \ell - \ell' \rangle^{2(s_0+\mathtt{b})} \langle j \rangle^{2(\mathtt{m}+\mathtt{b})} |\partial_\lambda^k R_j^{j'}(\ell - \ell')|^2 \langle j' \rangle^{2(\mathtt{m}+\mathtt{b}+1)} |h_{\ell',j'}|^2 \\
&\lesssim_{s_0,\mathtt{b}} \sum_{\ell',j'} |h_{\ell',j'}|^2 \sum_{\ell,j} \langle \ell, j \rangle^{2s} \langle \ell - \ell' \rangle^{2(s_0+\mathtt{b})} \langle j \rangle^{2(\mathtt{m}+\mathtt{b})} |\partial_\lambda^k R_j^{j'}(\ell - \ell')|^2 \langle j' \rangle^{2(\mathtt{m}+\mathtt{b}+1)} \\
&\overset{(15.51)}{\lesssim_{s_0,\mathtt{b}}} \gamma^{-2|k|} \sum_{\ell',j'} |h_{\ell',j'}|^2 \big(\mathfrak{M}_0^2(s_0, \mathtt{b}) \langle \ell', j' \rangle^{2s} + \mathfrak{M}_0^2(s, \mathtt{b}) \langle \ell', j' \rangle^{2s_0} \big) \\
&\lesssim_{s_0,\mathtt{b}} \gamma^{-2|k|} \big(\mathfrak{M}_0^2(s_0, \mathtt{b}) \|h\|_s^2 + \mathfrak{M}_0^2(s, \mathtt{b}) \|h\|_{s_0}^2 \big) \qquad (15.53)
\end{aligned}$$

using (2.27). Therefore (recall (2.78)) the modulo-tame constant $\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} R \langle D \rangle^{\mathtt{m}}}(s) \lesssim_{s_0,\mathtt{b}} \mathfrak{M}_0(s, \mathtt{b})$. Since (15.53) holds for both $R = R_1^{(0)}$ and $R = R_2^{(0)}$, we have proved that

$$\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}}}(s) \lesssim_{s_0,\mathtt{b}} \mathfrak{M}_0(s, \mathtt{b}) .$$

The inequality $\mathfrak{M}^\sharp_{\langle D \rangle^{\mathtt{m}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}}}(s) \lesssim \mathfrak{M}_0(s, \mathtt{b})$ follows similarly by (15.50). □

Proof of $(\mathbf{S2})_0$. We prove (15.36) at $\nu = 0$, namely we prove that, for $R = R_1^{(0)}$ or $R = R_2^{(0)}$,

$$\| |\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} \Delta_{12} R \langle D \rangle^{\mathtt{m}} | h \|_{s_0}^2 \lesssim_{S,\mathtt{b}} (\varepsilon \gamma^{-2(M+1)})^2 \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})}^2 \|h\|_{s_0}^2 , \quad \forall h \in H^{s_0} . \quad (15.54)$$

By (15.19) we get

$$\|\langle D \rangle^{\mathtt{m}+\mathtt{b}} \Delta_{12} R \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}\|_{\mathcal{L}(H^{s_0})}, \|\partial_{\varphi_l}^{s_0+\mathtt{b}} \langle D \rangle^{\mathtt{m}+\mathtt{b}} \Delta_{12} R \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}\|_{\mathcal{L}(H^{s_0})} \lesssim_{S,\mathtt{b}} \varepsilon \gamma^{-2(M+1)} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})}$$

for all $l = 1, \dots, |\mathbb{S}^+|$. Using (2.23) we deduce that, for all $\ell' \in \mathbb{Z}^{|\mathbb{S}^+|}$, $j' \in \mathbb{S}_0^c$,

$$\sum_{\ell,j} \langle \ell, j \rangle^{2s_0} \langle \ell - \ell' \rangle^{2(s_0+\mathtt{b})} \langle j \rangle^{2(\mathtt{m}+\mathtt{b})} |(\Delta_{12} R)_j^{j'}(\ell - \ell')|^2 \langle j' \rangle^{2(\mathtt{m}+\mathtt{b}+1)} \lesssim_{S,\mathtt{b}} (\varepsilon \gamma^{-2(M+1)})^2 \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})}^2 \langle \ell', j' \rangle^{2s_0} .$$

Using (15.52), and arguing as in (15.53), we get (15.54). The proof of (15.35) at $\nu = 0$ is analogous.

Proof of $(\mathbf{S3})_0$. It is trivial because, by definition, $\Lambda_0^\gamma = \mathtt{DC}(2\gamma, \tau) \times [\mathtt{h}_1, \mathtt{h}_2] \subseteq \mathtt{DC}(2\gamma - 2\rho, \tau) \times [\mathtt{h}_1, \mathtt{h}_2] = \Lambda_0^{\gamma - \rho}$.

15.1.1 The reducibility step

In this section we describe the inductive step and show how to define $\mathcal{L}_{\nu+1}$ (and Ψ_ν, Φ_ν, etc). To simplify the notation we drop the index ν and write $+$ instead of $\nu+1$, so that we write $\mathcal{L} := \mathcal{L}_\nu$, $\mathcal{D} := \mathcal{D}_\nu$, $D := D_\nu$, $\mu_j = \mu_j^\nu$, $\mathcal{R} := \mathcal{R}_\nu$, $R_1 := R_1^{(\nu)}$, $R_2 := R_2^{(\nu)}$, and $\mathcal{L}_+ := \mathcal{L}_{\nu+1}$, $\mathcal{D}_+ := \mathcal{D}_{\nu+1}$, and so on.

We conjugate the operator $\mathcal{L}$ in (15.22) by a transformation of the form (see (15.30))

$$\Phi := \mathbb{I}_\perp + \Psi\,, \qquad \Psi := \begin{pmatrix} \Psi_1 & \Psi_2 \\ \overline{\Psi}_2 & \overline{\Psi}_1 \end{pmatrix}. \tag{15.55}$$

We have

$$\mathcal{L}\Phi = \Phi(\omega\cdot\partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}) + (\omega\cdot\partial_\varphi\Psi + \mathrm{i}[\mathcal{D},\Psi] + \Pi_N\mathcal{R}) + \Pi_N^\perp\mathcal{R} + \mathcal{R}\Psi \tag{15.56}$$

where the projector Π_N is defined in (2.24), $\Pi_N^\perp := \mathbb{I}_2 - \Pi_N$, and $\omega\cdot\partial_\varphi\Psi$ is the commutator $[\omega\cdot\partial_\varphi,\Psi]$. We want to solve the homological equation

$$\omega\cdot\partial_\varphi\Psi + \mathrm{i}[\mathcal{D},\Psi] + \Pi_N\mathcal{R} = [\mathcal{R}] \tag{15.57}$$

where

$$[\mathcal{R}] := \begin{pmatrix} [R_1] & 0 \\ 0 & [\overline{R}_1] \end{pmatrix}, \quad [R_1] := \mathrm{diag}_{j\in\mathbb{S}_0^c}(R_1)_j^j(0)\,. \tag{15.58}$$

By (15.22), (15.26), (15.55), equation (15.57) is equivalent to the two scalar homological equations

$$\omega\cdot\partial_\varphi\Psi_1 + \mathrm{i}[D,\Psi_1] + \Pi_N R_1 = [R_1]\,, \qquad \omega\cdot\partial_\varphi\Psi_2 + \mathrm{i}(D\Psi_2 + \Psi_2 D) + \Pi_N R_2 = 0 \tag{15.59}$$

(note that $[R_1] = [\Pi_N R_1]$). We choose the solution of (15.59) given by

$$(\Psi_1)_j^{j'}(\ell) := \begin{cases} -\dfrac{(R_1)_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + \mu_j - \mu_{j'})} & \forall(\ell,j,j') \neq (0,j,\pm j)\,,\ |\ell|, |j-j'| \le N, \\ 0 & \text{otherwise;} \end{cases} \tag{15.60}$$

$$(\Psi_2)_j^{j'}(\ell) := \begin{cases} -\dfrac{(R_2)_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + \mu_j + \mu_{j'})} & \forall(\ell,j,j') \in \mathbb{Z}^{|\mathbb{S}^+|}\times\mathbb{S}_0^c\times\mathbb{S}_0^c\,,\ |\ell|, |j-j'| \le N, \\ 0 & \text{otherwise.} \end{cases} \tag{15.61}$$

Note that, since $\mu_j = \mu_{-j}$ for all $j\in\mathbb{S}_0^c$ (see (15.24)), the denominators in (15.60), (15.61) are different from zero for $(\omega,\mathtt{h}) \in \Lambda_{\nu+1}^\gamma$ (see (15.29) with $\nu \rightsquigarrow \nu+1$) and the maps Ψ_1, Ψ_2 are well defined on $\Lambda_{\nu+1}^\gamma$. Also note that the term $[R_1]$ in (15.58) (which is the term we are not able to remove by conjugation with Ψ_1 in (15.59)) contains only the diagonal entries $j'=j$ and not the anti-diagonal ones $j'=-j$, because $\mathcal{R}$ is zero on $j'=-j$ by (15.27). Thus, by construction,

$$(\Psi_1)_j^{j'}(\ell) = (\Psi_2)_j^{j'}(\ell) = 0 \quad \forall(\ell,j,j'),\ jj' < 0\,. \tag{15.62}$$

Lemma 15.7. (Homological equations) *The operators Ψ_1, Ψ_2 defined in* (15.60), (15.61) *(which, for all $\lambda\in\Lambda_{\nu+1}^\gamma$, solve the homological equations* (15.59)*) admit an extension to the whole parameter space $\mathbb{R}^{|\mathbb{S}^+|}\times[\mathtt{h}_1,\mathtt{h}_2]$. Such extended operators are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constants satisfying*

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\pm\mathtt{m}}\Psi\langle D\rangle^{\mp\mathtt{m}}}(s) \lesssim_{k_0} N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}\langle D\rangle^{\mathtt{m}}}(s), \tag{15.63}$$

$$\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\Psi\langle D\rangle^{\mp\mathtt{m}}}(s) \lesssim_{k_0} N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}\langle D\rangle^{\mathtt{m}}}(s) \tag{15.64}$$

$$\mathfrak{M}^\sharp_\Psi(s) \lesssim_{k_0} N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\mathcal{R}}(s) \tag{15.65}$$

where $\tau_1, \mathtt{b}, \mathtt{m}$ are defined in (15.10).

Given i_1, i_2, let $\Delta_{12}\Psi := \Psi(i_2) - \Psi(i_1)$. If $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$, then, for all $(\omega, \mathtt{h}) \in \Lambda^{\gamma_1}_{\nu+1}(i_1) \cap \Lambda^{\gamma_2}_{\nu+1}(i_2)$,

$$
\begin{aligned}
\| |\langle D\rangle^{\pm\mathtt{m}}\Delta_{12}\Psi\langle D\rangle^{\mp\mathtt{m}}| \|_{\mathcal{L}(H^{s_0})} &\lesssim N^{2\tau+2\mathtt{d}+\frac12}\gamma^{-1}\big(\| |\langle D\rangle^{\mathtt{m}}\mathcal{R}(i_2)\langle D\rangle^{\mathtt{m}}| \|_{\mathcal{L}(H^{s_0})}\|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} \\
&\quad + \| |\langle D\rangle^{\mathtt{m}}\Delta_{12}\mathcal{R}\langle D\rangle^{\mathtt{m}}| \|_{\mathcal{L}(H^{s_0})}\big)\,, \qquad (15.66)\\
\| |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\Delta_{12}\Psi\langle D\rangle^{\mp\mathtt{m}}| \|_{\mathcal{L}(H^{s_0})} &\lesssim N^{2\tau+2\mathtt{d}+\frac12}\gamma^{-1}\big(\| |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}(i_2)\langle D\rangle^{\mathtt{m}}| \|_{\mathcal{L}(H^{s_0})}\|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} \\
&\quad + \| |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\Delta_{12}\mathcal{R}\langle D\rangle^{\mathtt{m}}| \|_{\mathcal{L}(H^{s_0})}\big)\,. \qquad (15.67)
\end{aligned}
$$

Moreover Ψ is real, even and reversibility preserving.

Proof. For all $\lambda \in \Lambda^{\gamma}_{\nu+1}$, $(\ell, j, j') \neq (0, j, \pm j)$, $j, j' \in \mathbb{S}_0^c$ $|\ell|, |j - j'| \le N$, we have the small divisor estimate

$$|\omega\cdot\ell + \mu_j - \mu_{j'}| = |\omega\cdot\ell + \mu_{|j|} - \mu_{|j'|}| \ge \gamma|j|^{-\mathtt{d}}|j'|^{-\mathtt{d}}\langle\ell\rangle^{-\tau}$$

by (15.29), because $||j| - |j'|| \le |j - j'| \le N$. As in Lemma A.4, we extend the restriction to $F = \Lambda^{\gamma}_{\nu+1}$ of the function $(\omega\cdot\ell + \mu_j - \mu_{j'})^{-1}$ to the whole parameter space $\mathbb{R}^{|\mathbb{S}^+|}\times[\mathtt{h}_1, \mathtt{h}_2]$ by setting

$$g_{\ell,j,j'}(\lambda) := \frac{\chi\big(f(\lambda)\rho^{-1}\big)}{f(\lambda)}\,, \quad f(\lambda) := \omega\cdot\ell + \mu_j - \mu_{j'}\,, \quad \rho := \gamma\langle\ell\rangle^{-\tau}|j|^{-\mathtt{d}}|j'|^{-\mathtt{d}}\,,$$

where χ is the cut-off function in (2.16). We now estimate the corresponding constant M in (A.14). For $n \ge 1$, $x > 0$, the n-th derivative of the function $\tanh^{\frac12}(x)$ is $P_n(\tanh(x))\tanh^{\frac12 - n}(x)(1 - \tanh^2(x))$, where P_n is a polynomial of degree $\le 2n - 2$. Hence $|\partial^n_{\mathtt{h}}\{\tanh^{\frac12}(\mathtt{h}|j|)\}| \le C$ for all $n = 0, \ldots, k_0$, for all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, for all $j \in \mathbb{Z}$, for some $C = C(k_0, \mathtt{h}_1)$ independent of $n, \mathtt{h}, j$. By (15.23), (15.24), (15.8), (12.26), (13.79) (and recalling that μ_j here denotes μ^{ν}_j), since $\varepsilon\gamma^{-2(M+1)} \le \gamma$, we deduce that

$$\gamma^{|\alpha|}|\partial^{\alpha}_{\lambda}\mu_j| \lesssim \gamma|j|^{\frac12} \quad \forall\alpha \in \mathbb{N}^{|\mathbb{S}^+|+1},\ 1 \le |\alpha| \le k_0\,. \qquad (15.68)$$

Since $\gamma^{|\alpha|}|\partial^{\alpha}_{\lambda}(\omega\cdot\ell)| \le \gamma|\ell|$ for all $|\alpha| \ge 1$, we conclude that

$$\gamma^{|\alpha|}|\partial^{\alpha}_{\lambda}(\omega\cdot\ell + \mu_j - \mu_{j'})| \lesssim \gamma(|\ell| + |j|^{\frac12} + |j'|^{\frac12}) \lesssim \gamma\langle\ell\rangle|j|^{\frac12}|j'|^{\frac12}\,, \quad \forall\, 1 \le |\alpha| \le k_0\,. \qquad (15.69)$$

Thus (A.14) holds with $M = C\gamma\langle\ell\rangle|j|^{\frac12}|j'|^{\frac12}$ (which is $\ge \rho$) and (A.15) implies that

$$|g_{\ell,j,j'}|^{k_0,\gamma} \lesssim \gamma^{-1}\langle\ell\rangle^{\tau(k_0+1)+k_0}|j|^{\mathtt{m}}|j'|^{\mathtt{m}} \quad \text{with} \quad \mathtt{m} = (k_0+1)\mathtt{d} + \frac{k_0}{2} \qquad (15.70)$$

defined in (15.10). Formula (15.60) with $(\omega\cdot\ell + \mu_j - \mu_{j'})^{-1}$ replaced by $g_{\ell,j,j'}(\lambda)$ defines the extended operator Ψ_1 to $\mathbb{R}^{|\mathbb{S}^+|}\times[\mathtt{h}_1, \mathtt{h}_2]$. Analogously, we construct an extension of the function $(\omega\cdot\ell + \mu_j + \mu_{j'})^{-1}$ to the whole $\mathbb{R}^{|\mathbb{S}^+|}\times[\mathtt{h}_1, \mathtt{h}_2]$, and we obtain an extension of the operator Ψ_2 in (15.61).

Proof of (15.63), (15.64), (15.65). We prove (15.64) for Ψ_1, then the estimate for Ψ_2 follows in the same way, as well as (15.63), (15.65). Furthermore, we analyze the operator $\langle D\rangle^{\mathtt{m}}\partial^k_{\lambda}\Psi_1\langle D\rangle^{-\mathtt{m}}$, since $\langle D\rangle^{-\mathtt{m}}\partial^k_{\lambda}\Psi_1\langle D\rangle^{\mathtt{m}}$ can be treated in the same way. Differentiating $(\Psi_1)^{j'}_j(\ell) = g_{\ell,j,j'}(R_1)^{j'}_j(\ell)$, one has that, for any $|k| \le k_0$,

$$
\begin{aligned}
|\partial^k_{\lambda}(\Psi_1)^{j'}_j(\ell)| &\lesssim \sum_{k_1+k_2=k}|\partial^{k_1}_{\lambda}g_{\ell,j,j'}||\partial^{k_2}_{\lambda}(R_1)^{j'}_j(\ell)| \lesssim \sum_{k_1+k_2=k}\gamma^{-|k_1|}|g_{\ell,j,j'}|^{k_0,\gamma}|\partial^{k_2}_{\lambda}(R_1)^{j'}_j(\ell)| \\
&\overset{(15.70)}{\lesssim} \langle\ell\rangle^{\tau(k_0+1)+k_0}|j|^{\mathtt{m}}|j'|^{\mathtt{m}}\gamma^{-1-|k|}\sum_{|k_2|\le|k|}\gamma^{|k_2|}|\partial^{k_2}_{\lambda}(R_1)^{j'}_j(\ell)|\,. \qquad (15.71)
\end{aligned}
$$

For $|j - j'| \le N$, $j, j' \neq 0$, one has

$$|j|^{2\mathtt{m}} \lesssim |j|^{\mathtt{m}}(|j'|^{\mathtt{m}} + |j - j'|^{\mathtt{m}}) \lesssim |j|^{\mathtt{m}}(|j'|^{\mathtt{m}} + N^{\mathtt{m}}) \lesssim |j|^{\mathtt{m}}|j'|^{\mathtt{m}}N^{\mathtt{m}}. \qquad (15.72)$$

Hence, by (15.71) and (15.72), for all $|k| \le k_0$, $j, j' \in \mathbb{S}_0^c$, $\ell \in \mathbb{Z}^{|\mathbb{S}^+|}$, $|\ell| \le N$, $|j-j'| \le N$, one has

$$|j|^{\mathtt{m}} |\partial_\lambda^k (\Psi_1)_j^{j'}(\ell)| |j'|^{-\mathtt{m}} \lesssim N^{\tau_1} \gamma^{-1-|k|} \sum_{|k_2| \le |k|} \gamma^{|k_2|} |j|^{\mathtt{m}} |\partial_\lambda^{k_2} (R_1)_j^{j'}(\ell)| |j'|^{\mathtt{m}} \tag{15.73}$$

where $\tau_1 = \tau(k_0+1) + k_0 + \mathfrak{m}$ is defined in (15.10). Therefore, for all $0 \le |k| \le k_0$, we get

$$\begin{aligned}
&\| \, |\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} \partial_\lambda^k \Psi_1 \langle D \rangle^{-\mathtt{m}} | \, h \|_s^2 \\
&\le \sum_{\ell,j} \langle \ell, j \rangle^{2s} \Big(\sum_{|\ell'-\ell|, |j'-j| \le N} \langle \ell-\ell', j-j' \rangle^{\mathtt{b}} \langle j \rangle^{\mathtt{m}} |\partial_\lambda^k (\Psi_1)_j^{j'}(\ell-\ell')| \langle j' \rangle^{-\mathtt{m}} |h_{\ell',j'}| \Big)^2 \\
&\overset{(15.73)}{\lesssim_{k_0}} N^{2\tau_1} \gamma^{-2(1+|k|)} \sum_{|k_2| \le |k|} \gamma^{2|k_2|} \sum_{\ell,j} \langle \ell, j \rangle^{2s} \Big(\sum_{\ell',j'} |\langle \ell-\ell', j-j' \rangle^{\mathtt{b}} \langle j \rangle^{\mathtt{m}} \partial_\lambda^{k_2} (R_1)_j^{j'}(\ell-\ell') \langle j' \rangle^{\mathtt{m}}| |h_{\ell',j'}| \Big)^2 \\
&\lesssim_{k_0} N^{2\tau_1} \gamma^{-2(1+|k|)} \sum_{|k_2| \le |k|} \gamma^{2|k_2|} \big\| \, |\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} \partial_\lambda^{k_2} (R_1) \langle D \rangle^{\mathtt{m}} | \, [\, |h| \,] \, \big\|_s^2 \\
&\overset{(2.78),(2.27)}{\lesssim_{k_0}} N^{2\tau_1} \gamma^{-2(1+|k|)} \Big(\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} R_1 \langle D \rangle^{\mathtt{m}}}(s) \|h\|_{s_0} + \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} R_1 \langle D \rangle^{\mathtt{m}}}(s_0) \|h\|_s \Big)^2
\end{aligned} \tag{15.74}$$

and, recalling Definition 2.30, inequality (15.64) follows.

Proof of (15.66)-(15.67). By (15.60), for all $(\omega, \mathtt{h}) \in \Lambda_{\nu+1}^{\gamma_1}(i_1) \cap \Lambda_{\nu+1}^{\gamma_2}(i_2)$, one has

$$\Delta_{12}(\Psi_1)_j^{j'}(\ell) = \frac{\Delta_{12}(R_1)_j^{j'}(\ell)}{\delta_{\ell j j'}(i_1)} - (R_1)_j^{j'}(\ell)(i_2) \frac{\Delta_{12}\delta_{\ell j j'}}{\delta_{\ell j j'}(i_1)\delta_{\ell j j'}(i_2)}, \quad \delta_{\ell j j'} := \mathrm{i}(\omega \cdot \ell + \mu_j - \mu_{j'})$$

where we highlight the dependence of the operators and divisors on i_1, i_2. By (15.23), (15.8), (12.26), (13.79), (15.38), and $s_1 + \aleph_7^{(2M)}(M,0) \le s_0 + \mu(\mathtt{b})$, we get

$$\begin{aligned}
|\Delta_{12}\delta_{\ell j j'}| &= |\Delta_{12}(\mu_j - \mu_{j'})| \le |\Delta_{12}\mu_j| + |\Delta_{12}\mu_{j'}| \\
&\le \big|\Delta_{12}\mathtt{m}_{\frac12}\big| \big(|j|^{\frac12} \tanh^{\frac12}(\mathtt{h}|j|) + |j'|^{\frac12} \tanh^{\frac12}(\mathtt{h}|j'|) \big) + |\Delta_{12} r_j| + |\Delta_{12} r_{j'}| + |\Delta_{12} r_j^\nu| + |\Delta_{12} r_{j'}^\nu| \\
&\lesssim \varepsilon \gamma^{-2(M+1)} |j|^{\frac12} |j'|^{\frac12} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} .
\end{aligned}$$

Since $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$, for $\varepsilon\gamma^{-2M-3}$ small enough, one has

$$|\Delta_{12}(\Psi_1)_j^{j'}(\ell)| \lesssim N^{2\tau} \gamma^{-1} |j|^{2\mathtt{d}+\frac12} |j'|^{2\mathtt{d}+\frac12} \big(|\Delta_{12}(R_1)_j^{j'}(\ell)| + |(R_1)_j^{j'}(\ell)(i_2)| \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} \big) . \tag{15.75}$$

For $|j-j'| \le N$, recalling that $\mathtt{m} > 2\mathtt{d} + \frac12$ by (15.10), we have

$$\begin{aligned}
|j|^{\mathtt{m}+2\mathtt{d}+\frac12} |j'|^{2\mathtt{d}+\frac12-\mathtt{m}} \le |j|^{\mathtt{m}+2\mathtt{d}+\frac12} &\lesssim |j|^{\mathtt{m}} \big(|j-j'|^{2\mathtt{d}+\frac12} + |j'|^{2\mathtt{d}+\frac12} \big) \\
&\lesssim |j|^{\mathtt{m}} \big(N^{2\mathtt{d}+\frac12} + |j'|^{\mathtt{m}} \big) \lesssim N^{2\mathtt{d}+\frac12} |j|^{\mathtt{m}} |j'|^{\mathtt{m}}
\end{aligned}$$

and, by (15.75), we deduce

$$\langle j \rangle^{\mathtt{m}} |\Delta_{12}(\Psi_1)_j^{j'}(\ell)| \langle j' \rangle^{-\mathtt{m}} \lesssim N^{2\tau+2\mathtt{d}+\frac12} \gamma^{-1} \langle j \rangle^{\mathtt{m}} \langle j' \rangle^{\mathtt{m}} \big(|(R_1)_j^{j'}(\ell)(i_2)| \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} + |\Delta_{12}(R_1)_j^{j'}(\ell)| \big) .$$

The operator $\Delta_{12}\Psi_2$ satisfies a similar estimate and (15.66), (15.67) follow arguing as in (15.74).

Finally, since $\mathcal{R}$ is even and reversible, (15.60), (15.61) and (2.70)-(2.71) imply that Ψ is even and reversibility preserving. □

If Ψ, with Ψ_1, Ψ_2 defined in (15.60)-(15.61), satisfies the smallness condition

$$4C(\mathtt{b})C(k_0)\mathfrak{M}^\sharp_\Psi(s_0) \le 1/2 , \tag{15.76}$$

then, by Lemma 2.33, Φ is invertible, and (15.56), (15.57) imply that, for all $\lambda \in \Lambda^{\gamma}_{\nu+1}$,

$$\mathcal{L}_+ = \Phi^{-1}\mathcal{L}\Phi = \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_+ + \mathcal{R}_+ \tag{15.77}$$

which proves (15.31) and (15.22) at the step $\nu + 1$, with

$$\mathrm{i}\mathcal{D}_+ := \mathrm{i}\mathcal{D} + [\mathcal{R}] \,, \quad \mathcal{R}_+ := \Phi^{-1}\big(\Pi_N^\perp \mathcal{R} + \mathcal{R}\Psi - \Psi[\mathcal{R}]\big) \,. \tag{15.78}$$

We note that $\mathcal{R}_+$ satisfies

$$\mathcal{R}_+ = \begin{pmatrix} (R_+)_1 & (R_+)_2 \\ (\overline{R}_+)_2 & (\overline{R}_+)_1 \end{pmatrix}, \quad [(R_+)_1]_j^{j'}(\ell) = [(R_+)_2]_j^{j'}(\ell) = 0 \quad \forall (\ell, j, j'),\ jj' < 0 \,, \tag{15.79}$$

similarly as $\mathcal{R}_\nu$ in (15.27), because the property of having zero matrix entries for $jj' < 0$ is preserved by matrix product, and $\mathcal{R}, \Psi, [\mathcal{R}]$ satisfy such a property (see (15.27), (15.62), (15.58)), and therefore, by Neumann series, also Φ^{-1} does.

The right hand sides of (15.77)-(15.78) define an extension of $\mathcal{L}_+$ to the whole parameter space $\mathbb{R}^{|\mathbb{S}^+|} \times [\mathtt{h}_1, \mathtt{h}_2]$, since $\mathcal{R}$ and Ψ are defined on $\mathbb{R}^{|\mathbb{S}^+|} \times [\mathtt{h}_1, \mathtt{h}_2]$.

The new operator $\mathcal{L}_+$ in (15.77) has the same form as $\mathcal{L}$ in (15.22), with the non-diagonal remainder $\mathcal{R}_+$ defined in (15.78) which is the sum of a quadratic function of Ψ, $\mathcal{R}$ and a term $\Pi_N^\perp \mathcal{R}$ supported on high frequencies. The new normal form $\mathcal{D}_+$ in (15.78) is diagonal:

Lemma 15.8. (New diagonal part). *For all $(\omega, \mathtt{h}) \in \mathbb{R}^{|\mathbb{S}^+|} \times [\mathtt{h}_1, \mathtt{h}_2]$ we have*

$$\mathrm{i}\mathcal{D}_+ = \mathrm{i}\mathcal{D} + [\mathcal{R}] = \mathrm{i}\begin{pmatrix} D_+ & 0 \\ 0 & -D_+ \end{pmatrix}, \quad D_+ := \mathrm{diag}_{j \in \mathbb{S}_0^c} \mu_j^+ \,, \quad \mu_j^+ := \mu_j + \mathtt{r}_j \in \mathbb{R} \,, \tag{15.80}$$

with $\mathtt{r}_j = \mathtt{r}_{-j}$, $\mu_j^+ = \mu_{-j}^+$ for all $j \in \mathbb{S}_0^c$, and, on $\mathbb{R}^{|\mathbb{S}^+|} \times [\mathtt{h}_1, \mathtt{h}_2]$,

$$|\mathtt{r}_j|^{k_0,\gamma} = |\mu_j^+ - \mu_j|^{k_0,\gamma} \lesssim |j|^{-2\mathtt{m}} \mathfrak{M}^\sharp_{\langle D \rangle^{\mathtt{m}} \mathcal{R} \langle D \rangle^{\mathtt{m}}}(s_0). \tag{15.81}$$

Moreover, given tori $i_1(\omega, \mathtt{h})$, $i_2(\omega, \mathtt{h})$, the difference

$$|\mathtt{r}_j(i_1) - \mathtt{r}_j(i_2)| \lesssim |j|^{-2\mathtt{m}} \| |\langle D \rangle^{\mathtt{m}} \Delta_{12} \mathcal{R} \langle D \rangle^{\mathtt{m}} |\|_{\mathcal{L}(H^{s_0})} \,. \tag{15.82}$$

Proof. Identity (15.80) follows by (15.22) and (15.58) with $\mathtt{r}_j := -\mathrm{i}(R_1)_j^j(0)$. Since R_1 satisfies (15.27) and it is even, we deduce, by (2.65), that $\mathtt{r}_{-j} = \mathtt{r}_j$. Since $\mathcal{R}$ is reversible, (2.70) implies that $\mathtt{r}_j := -\mathrm{i}(R_1)_j^j(0)$ satisfies $\mathtt{r}_j = \overline{\mathtt{r}_{-j}}$. Therefore $\mathtt{r}_j = \overline{\mathtt{r}_{-j}} = \overline{\mathtt{r}_j}$ and each $\mathtt{r}_j \in \mathbb{R}$.

Recalling Definition 2.30, we have $\| |\partial_\lambda^k (\langle D \rangle^{\mathtt{m}} R_1 \langle D \rangle^{\mathtt{m}})| h\|_{s_0} \leq 2\gamma^{-|k|} \mathfrak{M}^\sharp_{\langle D \rangle^{\mathtt{m}} R_1 \langle D \rangle^{\mathtt{m}}}(s_0) \|h\|_{s_0}$, for all $\lambda = (\omega, \mathtt{h})$, $0 \leq |k| \leq k_0$, and therefore (see (2.75))

$$|\partial_\lambda^k (R_1)_j^j(0)| \lesssim |j|^{-2\mathtt{m}} \gamma^{-|k|} \mathfrak{M}^\sharp_{\langle D \rangle^{\mathtt{m}} R_1 \langle D \rangle^{\mathtt{m}}}(s_0) \lesssim |j|^{-2\mathtt{m}} \gamma^{-|k|} \mathfrak{M}^\sharp_{\langle D \rangle^{\mathtt{m}} \mathcal{R} \langle D \rangle^{\mathtt{m}}}(s_0)$$

which implies (15.81). Estimate (15.82) follows by $|\Delta_{12}(R_1)_j^j(0)| \lesssim |j|^{-2\mathtt{m}} \| |\langle D \rangle^{\mathtt{m}} \Delta_{12} \mathcal{R} \langle D \rangle^{\mathtt{m}} |\|_{\mathcal{L}(H^{s_0})}$. $\square$

15.1.2 The iteration

Let $n \geq 0$ and suppose that $(\mathbf{S1})_\nu$-$(\mathbf{S3})_\nu$ are true for all $\nu = 0, \ldots, n$. We prove $(\mathbf{S1})_{n+1}$-$(\mathbf{S3})_{n+1}$. For simplicity of notation (as in other parts of the paper) we omit to write the dependence on k_0 which is considered as a fixed constant.

PROOF OF $(\mathbf{S1})_{n+1}$. By (15.63)-(15.65), (15.28), and using that $\mathfrak{M}^\sharp_{\mathcal{R}_n}(s) \lesssim \mathfrak{M}^\sharp_{\langle D \rangle^{\mathtt{m}} \mathcal{R}_n \langle D \rangle^{\mathtt{m}}}(s)$, the operator Ψ_n defined in Lemma 15.7 satisfies estimates (15.32)-(15.34) with $\nu = n+1$. In particular at $s = s_0$ we have

$$\mathfrak{M}^\sharp_{\langle D \rangle^{\pm \mathtt{m}} \Psi_n \langle D \rangle^{\mp \mathtt{m}}}(s_0) \,,\ \mathfrak{M}^\sharp_{\Psi_n}(s_0) \leq C(s_0, \mathtt{b}) N_n^{\tau_1} N_{n-1}^{-\mathtt{a}} \gamma^{-1} \mathfrak{M}_0(s_0, \mathtt{b}) \,. \tag{15.83}$$

Therefore, by (15.83), (15.10), (15.21), choosing $\tau_2 > \tau_1$, the smallness condition (15.76) holds for $N_0 := N_0(S, \mathtt{b})$ large enough (for any $n \geq 0$), and the map $\Phi_n = \mathbb{I}_\perp + \Psi_n$ is invertible, with inverse

$$\Phi_n^{-1} = \mathbb{I}_\perp + \check{\Psi}_n\,, \qquad \check{\Psi}_n := \begin{pmatrix} \check{\Psi}_{n,1} & \check{\Psi}_{n,2} \\ \overline{\check{\Psi}}_{n,2} & \overline{\check{\Psi}}_{n,1} \end{pmatrix}. \tag{15.84}$$

Moreover also the smallness condition (2.88) (of Corollary 2.34) with $A = \Psi_n$, holds, and Lemma 2.33, Corollary 2.34 and Lemma 15.7 imply that the maps $\check{\Psi}_n$, $\langle D\rangle^{\pm\mathtt{m}}\check{\Psi}_n\langle D\rangle^{\mp\mathtt{m}}$ and $\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\check{\Psi}_n\langle D\rangle^{\mp\mathtt{m}}$ are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constants satisfying

$$\begin{aligned} \mathfrak{M}^\sharp_{\check{\Psi}_n}(s),\ \mathfrak{M}^\sharp_{\langle D\rangle^{\pm\mathtt{m}}\check{\Psi}_n\langle D\rangle^{\mp\mathtt{m}}}(s) &\lesssim_{s_0,\mathtt{b}} N_n^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s) && (15.85)\\ &\overset{(15.28)_{|n}}{\lesssim_{s_0,\mathtt{b}}} N_n^{\tau_1}N_{n-1}^{-\mathtt{a}}\gamma^{-1}\mathfrak{M}_0(s,\mathtt{b})\,, && (15.86) \end{aligned}$$

and

$$\begin{aligned} \mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\check{\Psi}_n\langle D\rangle^{\mp\mathtt{m}}}(s) &\lesssim_{s_0,\mathtt{b}} N_n^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s)\\ &\quad + N_n^{2\tau_1}\gamma^{-2}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s_0)\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s) && (15.87)\\ &\overset{(15.28)_{|n},(15.10),(15.21)}{\lesssim_{s_0,\mathtt{b}}} N_n^{\tau_1}N_{n-1}\gamma^{-1}\mathfrak{M}_0(s,\mathtt{b})\,. && (15.88) \end{aligned}$$

Conjugating $\mathcal{L}_n$ by Φ_n, we obtain, by (15.77)-(15.78), for all $\lambda \in \Lambda^\gamma_{n+1}$,

$$\mathcal{L}_{n+1} = \Phi_n^{-1}\mathcal{L}_n\Phi_n = \omega\cdot\partial_\varphi\mathbb{I}_\perp + \mathrm{i}\mathcal{D}_{n+1} + \mathcal{R}_{n+1}\,, \tag{15.89}$$

namely (15.31) at $\nu = n+1$, where

$$\mathrm{i}\mathcal{D}_{n+1} := \mathrm{i}\mathcal{D}_n + [\mathcal{R}_n]\,, \qquad \mathcal{R}_{n+1} := \Phi_n^{-1}\big(\Pi^\perp_{N_n}\mathcal{R}_n + \mathcal{R}_n\Psi_n - \Psi_n[\mathcal{R}_n]\big)\,. \tag{15.90}$$

The operator $\mathcal{L}_{n+1}$ is real, even and reversible because Φ_n is real, even and reversibility preserving (Lemma 15.7) and $\mathcal{L}_n$ is real, even and reversible. Note that the operators $\mathcal{D}_{n+1}, \mathcal{R}_{n+1}$ are defined on $\mathbb{R}^{|\mathbb{S}^+|}\times[\mathtt{h}_1,\mathtt{h}_2]$, and the identity (15.89) holds on Λ^γ_{n+1}.

By Lemma 15.8 the operator $\mathcal{D}_{n+1}$ is diagonal and, by (15.15), (15.28), (15.18), its eigenvalues $\mu_j^{n+1} : \mathbb{R}^{|\mathbb{S}^+|}\times[\mathtt{h}_1,\mathtt{h}_2]\to\mathbb{R}$ satisfy

$$|\mathtt{r}_j^n|^{k_0,\gamma} = |\mu_j^{n+1} - \mu_j^n|^{k_0,\gamma} \lesssim |j|^{-2\mathtt{m}}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s_0) \leq C(S,\mathtt{b})\varepsilon\gamma^{-2(M+1)}|j|^{-2\mathtt{m}}N_{n-1}^{-\mathtt{a}}\,,$$

which is (15.25) with $\nu = n+1$. Thus also (15.24) at $\nu = n+1$ holds, by a telescoping sum. In addition, by (15.79) the operator $\mathcal{R}_{n+1}$ satisfies (15.27) with $\nu = n+1$. In order to prove that (15.28) holds with $\nu = n+1$, we first provide the following inductive estimates on the new remainder $\mathcal{R}_{n+1}$.

Lemma 15.9. *The operators $\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}$ and $\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}$ are $\mathcal{D}^{k_0}$-modulo-tame, with*

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) \lesssim_{s_0,\mathtt{b}} N_n^{-\mathtt{b}}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s) + \frac{N_n^{\tau_1}}{\gamma}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s)\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s_0)\,, \tag{15.91}$$

$$\begin{aligned} \mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) &\lesssim_{s_0,\mathtt{b}} \mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s)\\ &\quad + N_n^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s_0)\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s)\,. && (15.92) \end{aligned}$$

Proof. By (15.90) and (15.84), we write

$$\begin{aligned} \langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}} &= \langle D\rangle^{\mathtt{m}}\Pi^\perp_{N_n}\mathcal{R}_n\langle D\rangle^{\mathtt{m}} + (\langle D\rangle^{\mathtt{m}}\check{\Psi}_n\langle D\rangle^{-\mathtt{m}})(\langle D\rangle^{\mathtt{m}}\Pi^\perp_{N_n}\mathcal{R}_n\langle D\rangle^{\mathtt{m}})\\ &\quad + \Big(\mathbb{I}_\perp + \langle D\rangle^{\mathtt{m}}\check{\Psi}_n\langle D\rangle^{-\mathtt{m}}\Big)\Big((\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}})(\langle D\rangle^{-\mathtt{m}}\Psi_n\langle D\rangle^{\mathtt{m}})\Big)\\ &\quad - \Big(\mathbb{I}_\perp + \langle D\rangle^{\mathtt{m}}\check{\Psi}_n\langle D\rangle^{-\mathtt{m}}\Big)\Big((\langle D\rangle^{\mathtt{m}}\Psi_n\langle D\rangle^{-\mathtt{m}})(\langle D\rangle^{\mathtt{m}}[\mathcal{R}_n]\langle D\rangle^{\mathtt{m}})\Big)\,. && (15.93) \end{aligned}$$

The proof of (15.91) follows by estimating separately all the terms in (15.93), applying Lemmata 2.35, 2.32, and (15.63), (15.85), $(15.28)_{|n}$, (15.10), (15.21). The proof of (15.92) follows by formula (15.93), Lemmata 2.32, 2.35 and estimates (15.63), (15.64), (15.85), $(15.28)_{|n}$, (15.10), (15.21). $\square$

In the next lemma we prove that (15.28) holds at $\nu = n+1$, concluding the proof of $(\mathbf{S1})_{n+1}$.

Lemma 15.10. *For $N_0 = N_0(S, \mathtt{b}) > 0$ large enough we have*

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) \leq C_*(s_0,\mathtt{b}) N_n^{-\mathtt{a}}\mathfrak{M}_0(s,\mathtt{b})$$
$$\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) \leq C_*(s_0,\mathtt{b}) N_n\mathfrak{M}_0(s,\mathtt{b}) .$$

Proof. By (15.91) and (15.28) we get

$$\begin{aligned}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) &\lesssim_{s_0,\mathtt{b}} N_n^{-\mathtt{b}}N_{n-1}\mathfrak{M}_0(s,\mathtt{b}) + N_n^{\tau_1}\gamma^{-1}\mathfrak{M}_0(s,\mathtt{b})\mathfrak{M}_0(s_0,\mathtt{b})N_{n-1}^{-2\mathtt{a}}\\ &\leq C_*(s_0,\mathtt{b})N_n^{-\mathtt{a}}\mathfrak{M}_0(s,\mathtt{b})\end{aligned}$$

by (15.10), (15.21), taking $N_0(S, \mathtt{b}) > 0$ large enough and $\tau_2 > \tau_1 + \mathtt{a}$. Then by (15.92), (15.28) we get that

$$\begin{aligned}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) &\lesssim_{s_0,\mathtt{b}} N_{n-1}\mathfrak{M}_0(s,\mathtt{b}) + N_n^{\tau_1}N_{n-1}^{1-\mathtt{a}}\gamma^{-1}\mathfrak{M}_0(s,\mathtt{b})\mathfrak{M}_0(s_0,\mathtt{b})\\ &\leq C_*(s_0,\mathtt{b})N_n\mathfrak{M}_0(s,\mathtt{b})\end{aligned}$$

by (15.10), (15.21) and taking $N_0(S, \mathtt{b}) > 0$ large enough. $\square$

Proof of $(\mathbf{S2})_{n+1}$. At the n-th step we have already constructed the operators

$$\mathcal{R}_n(i_1)\,,\ \Psi_{n-1}(i_1)\,,\ \mathcal{R}_n(i_2)\,,\ \Psi_{n-1}(i_2)\,,$$

which are defined for any $(\omega, \mathtt{h}) \in \mathbb{R}^{|\mathbb{S}^+|} \times [\mathtt{h}_1, \mathtt{h}_2]$ and satisfy estimates (15.28), (15.32), (15.33). We now estimate the operator $\Delta_{12}\mathcal{R}_{n+1}$ for any $(\omega, \mathtt{h}) \in \Lambda^{\gamma_1}_{n+1}(i_1) \cap \Lambda^{\gamma_2}_{n+1}(i_2)$. For $(\omega, \mathtt{h}) \in \Lambda^{\gamma_1}_{n+1}(i_1) \cap \Lambda^{\gamma_2}_{n+1}(i_2)$, by (15.66), (2.74), (15.28), (15.15), (15.35) we get

$$\| \,|\langle D\rangle^{\pm\mathtt{m}}\Delta_{12}\Psi_n\langle D\rangle^{\mp\mathtt{m}}|\,\|_{\mathcal{L}(H^{s_0})} \lesssim_{S,\mathtt{b}} N_n^{2\tau+2\mathtt{d}+\frac12}N_{n-1}^{-\mathtt{a}}\varepsilon\gamma^{-2M-3}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})} \,. \tag{15.94}$$

Moreover, using (15.67), (2.74), (15.28), (15.15), (15.36), we get

$$\| \,|\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\Delta_{12}\Psi_n\langle D\rangle^{\mp\mathtt{m}}|\,\|_{\mathcal{L}(H^{s_0})} \lesssim_{S,\mathtt{b}} N_n^{2\tau+2\mathtt{d}+\frac12}N_{n-1}\varepsilon\gamma^{-2M-3}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})} \,. \tag{15.95}$$

By (15.83), (15.10), (15.20), (15.21), using that $\tau_2 > \tau_1$ (and taking N_0 large enough), the smallness condition (2.91) is verified. Therefore, applying Lemma 2.38 together with estimates (15.94), (15.95), (15.86), (15.88), (2.74) and using (15.10), (15.21), we get

$$\| \,|\Delta_{12}\langle D\rangle^{\pm\mathtt{m}}\check\Psi_n\langle D\rangle^{\mp\mathtt{m}}|\,\|_{\mathcal{L}(H^{s_0})} \lesssim_{S,\mathtt{b}} N_n^{2\tau+2\mathtt{d}+\frac12}N_{n-1}^{-\mathtt{a}}\varepsilon\gamma^{-2M-3}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})} \,, \tag{15.96}$$

$$\| \,|\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\Delta_{12}\check\Psi_n\langle D\rangle^{\mp\mathtt{m}}|\,\|_{\mathcal{L}(H^{s_0})} \lesssim_{S,\mathtt{b}} N_n^{2\tau+2\mathtt{d}+\frac12}N_{n-1}\varepsilon\gamma^{-2M-3}\|i_1-i_2\|_{s_0+\mu(\mathtt{b})} \,. \tag{15.97}$$

Estimates (15.35), (15.36) for $\nu = n+1$ for the term $\Delta_{12}\mathcal{R}_{n+1}$ (where $\mathcal{R}_{n+1}$ is defined in (15.90)) follow by recalling (15.93), by a repeated application of triangular inequality, by Lemma 2.36, using estimates (15.96), (15.97), (15.86), (15.88), (15.94), (15.95), (2.74), (15.28), (15.32), (15.33), (15.35), (15.36), (15.15), taking $N_0(S, \mathtt{b}) > 0$ large enough, recalling (15.10) and using the smallness condition (15.21).

The proof of (15.37) for $\nu = n+1$ follows estimating $\Delta_{12}(r_j^{n+1} - r_j^n) = \Delta_{12}\mathtt{r}_j^n$ by (15.82) of Lemma 15.8 and by (15.35) for $\nu = n$. Estimate (15.38) for $\nu = n+1$ follows by a telescoping argument using (15.37) and (15.35).

Proof of $(\mathbf{S3})_{n+1}$. First we note that the non-resonance conditions imposed in (15.29) are actually finitely many. We prove the following

- CLAIM: Let $\omega \in \mathtt{DC}(2\gamma, \tau)$ and $\varepsilon\gamma^{-2(M+1)} \leq 1$. Then there exists $C_0 > 0$ such that, for any $\nu = 0, \ldots, n$, for all $|\ell|, |j - j'| \leq N_\nu$, $j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+$, if

$$\min\{j, j'\} \geq C_0 N_\nu^{2(\tau+1)} \gamma^{-2}, \tag{15.98}$$

then $|\omega \cdot \ell + \mu_j^\nu - \mu_{j'}^\nu| \geq \gamma \langle \ell \rangle^{-\tau}$.

PROOF OF THE CLAIM. By (15.23), (15.24) and recalling also (13.79), one has

$$\mu_j^\nu = \mathtt{m}_{\frac{1}{2}} j^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h} j) + \mathfrak{r}_j^\nu, \quad \mathfrak{r}_j^\nu := r_j + r_j^\nu, \quad \sup_{j \in \mathbb{S}^c} j^{\frac{1}{2}} |\mathfrak{r}_j^\nu|^{k_0,\gamma} \lesssim_S \varepsilon\gamma^{-2(M+1)}. \tag{15.99}$$

For all $j, j' \in \mathbb{N} \setminus \{0\}$, one has

$$|\sqrt{j \tanh(\mathtt{h} j)} - \sqrt{j' \tanh(\mathtt{h} j')}| \leq \frac{C(\mathtt{h})}{\min\{\sqrt{j}, \sqrt{j'}\}} |j - j'|. \tag{15.100}$$

Then, using (15.100) and that $\omega \in \mathtt{DC}(2\gamma, \tau)$, we have, for $|j - j'| \leq N_\nu$, $|\ell| \leq N_\nu$,

$$\begin{aligned} |\omega \cdot \ell + \mu_j^\nu - \mu_{j'}^\nu| &\geq |\omega \cdot \ell| - |\mathtt{m}_{\frac{1}{2}}| \frac{C(\mathtt{h})}{\min\{\sqrt{j}, \sqrt{j'}\}} |j - j'| - |\mathfrak{r}_j^\nu| - |\mathfrak{r}_{j'}^\nu| \\ &\overset{(12.26),(15.99)}{\geq} \frac{2\gamma}{\langle \ell \rangle^\tau} - \frac{2C(\mathtt{h}) N_\nu}{\min\{\sqrt{j}, \sqrt{j'}\}} - \frac{C(S)\varepsilon\gamma^{-2(M+1)}}{\min\{\sqrt{j}, \sqrt{j'}\}} \overset{(15.98)}{\geq} \frac{\gamma}{\langle \ell \rangle^\tau}, \end{aligned}$$

where the last inequality holds for C_0 large enough. This proves the claim.

Now we prove $(\mathbf{S3})_{n+1}$, namely that

$$C(S) N_n^{(\tau+1)(4\mathtt{d}+1)} \gamma^{-4\mathtt{d}} \|i_2 - i_1\|_{s_0+\mu(\mathtt{b})} \leq \rho \quad \Longrightarrow \quad \Lambda_{n+1}^\gamma(i_1) \subseteq \Lambda_{n+1}^{\gamma-\rho}(i_2). \tag{15.101}$$

Let $\lambda \in \Lambda_{n+1}^\gamma(i_1)$. Definition (15.29) and (15.39) with $\nu = n$ (i.e. $(\mathbf{S3})_n$) imply that $\Lambda_{n+1}^\gamma(i_1) \subseteq \Lambda_n^\gamma(i_1) \subseteq \Lambda_n^{\gamma-\rho}(i_2)$. Moreover $\lambda \in \Lambda_n^{\gamma-\rho}(i_2) \subseteq \Lambda_n^{\gamma/2}(i_2)$ because $\rho \leq \gamma/2$. Thus $\Lambda_{n+1}^\gamma(i_1) \subseteq \Lambda_n^{\gamma-\rho}(i_2) \subseteq \Lambda_n^{\gamma/2}(i_2)$. Hence $\Lambda_{n+1}^\gamma(i_1) \subseteq \Lambda_n^\gamma(i_1) \cap \Lambda_n^{\gamma/2}(i_2)$, and estimate (15.38) on $|\Delta_{12} r_j^n| = |r_j^n(\lambda, i_2(\lambda)) - r_j^n(\lambda, i_1(\lambda))|$ holds for any $\lambda \in \Lambda_{n+1}^\gamma(i_1)$. By the previous claim, since $\omega \in \mathtt{DC}(2\gamma, \tau)$, for all $|\ell|, |j - j'| \leq N_n$ satisfying (15.98) with $\nu = n$ we have

$$|\omega \cdot \ell + \mu_j^n(\lambda, i_2(\lambda)) - \mu_{j'}^n(\lambda, i_2(\lambda))| \geq \frac{\gamma}{\langle \ell \rangle^\tau} \geq \frac{\gamma}{\langle \ell \rangle^\tau j^{\mathtt{d}} j'^{\mathtt{d}}} \geq \frac{\gamma - \rho}{\langle \ell \rangle^\tau j^{\mathtt{d}} j'^{\mathtt{d}}}.$$

It remains to prove that the second Melnikov conditions in (15.29) with $\nu = n+1$ also hold for j, j' violating $(15.98)|_{\nu=n}$, namely that

$$|\omega \cdot \ell + \mu_j^n(\lambda, i_2(\lambda)) - \mu_{j'}^n(\lambda, i_2(\lambda))| \geq \frac{\gamma - \rho}{\langle \ell \rangle^\tau j^{\mathtt{d}} j'^{\mathtt{d}}}, \ \forall |\ell|, |j - j'| \leq N_n, \ \min\{j, j'\} \leq C_0 N_n^{2(\tau+1)} \gamma^{-2}. \tag{15.102}$$

The conditions on j, j' in (15.102) imply that

$$\max\{j, j'\} = \min\{j, j'\} + |j - j'| \leq C_0 N_n^{2(\tau+1)} \gamma^{-2} + N_n \leq 2 C_0 N_n^{2(\tau+1)} \gamma^{-2}. \tag{15.103}$$

Now by (15.23), (15.24), (15.100), recalling (12.26), (13.79), (15.38) and the bound $\varepsilon\gamma^{-2(M+1)} \leq 1$, we get

$$\begin{aligned} |(\mu_j^n - \mu_{j'}^n)(\lambda, i_2(\lambda)) - (\mu_j^n - \mu_{j'}^n)(\lambda, i_1(\lambda))| &\leq |(\mu_j^0 - \mu_{j'}^0)(\lambda, i_2(\lambda)) - (\mu_j^0 - \mu_{j'}^0)(\lambda, i_1(\lambda))| \\ &\quad + |r_j^n(\lambda, i_2(\lambda)) - r_j^n(\lambda, i_1(\lambda))| + |r_{j'}^n(\lambda, i_2(\lambda)) - r_{j'}^n(\lambda, i_1(\lambda))| \\ &\leq \frac{C(S) N_n}{\min\{\sqrt{j}, \sqrt{j'}\}} \|i_2 - i_1\|_{s_0+\mu(\mathtt{b})}. \end{aligned} \tag{15.104}$$

Since $\lambda \in \Lambda^{\gamma}_{n+1}(i_1)$, by (15.104) we have, for all $|\ell| \le N_n$, $|j-j'| \le N_n$,

$$\begin{aligned}|\omega\cdot\ell+\mu^n_j(i_2)-\mu^n_{j'}(i_2)| &\ge |\omega\cdot\ell+\mu^n_j(i_1)-\mu^n_{j'}(i_1)| - |(\mu^n_j-\mu^n_{j'})(i_2)-(\mu^n_j-\mu^n_{j'})(i_1)|\\ &\ge \frac{\gamma}{\langle\ell\rangle^{\tau} j^{\mathtt{d}} j'^{\mathtt{d}}} - \frac{C(S)N_n}{\min\{\sqrt{j},\sqrt{j'}\}}\|i_2-i_1\|_{s_0+\mu(\mathtt{b})}\\ &\ge \frac{\gamma}{\langle\ell\rangle^{\tau} j^{\mathtt{d}} j'^{\mathtt{d}}} - C(S)N_n\|i_2-i_1\|_{s_0+\mu(\mathtt{b})} \ge \frac{\gamma-\rho}{\langle\ell\rangle^{\tau} j^{\mathtt{d}} j'^{\mathtt{d}}}\end{aligned}$$

provided $C(S)N_n\langle\ell\rangle^{\tau} j^{\mathtt{d}} j'^{\mathtt{d}}\|i_2-i_1\|_{s_0+\mu(\mathtt{b})} \le \rho$. Using that $|\ell| \le N_n$ and (15.103), the above inequality is implied by the inequality assumed in (15.101). The proof for the second Melnikov conditions for $\omega\cdot\ell+\mu^n_j+\mu^n_{j'}$ can be carried out similarly (in fact, it is simpler). This completes the proof of (15.39) with $\nu = n+1$. □

15.2 Almost-invertibility of $\mathcal{L}_\omega$

By (14.6), $\mathcal{L}_\omega = \mathcal{P}_\perp \mathcal{L}_\perp \mathcal{P}_\perp^{-1}$, where $\mathcal{P}_\perp$ is defined in (14.2), (14.3). By (15.45), for any $\lambda \in \Lambda^{\gamma}_n$, we have that $\mathcal{L}_0 = \mathcal{U}_n \mathcal{L}_n \mathcal{U}_n^{-1}$, where $\mathcal{U}_n$ is defined in (15.40), $\mathcal{L}_0 = \mathcal{L}_\perp^{sym}$, and $\mathcal{L}_\perp^{sym} = \mathcal{L}_\perp$ on the subspace of functions even in x (see (15.3)). Thus

$$\mathcal{L}_\omega = \mathcal{V}_n \mathcal{L}_n \mathcal{V}_n^{-1}, \quad \mathcal{V}_n := \mathcal{P}_\perp \mathcal{U}_n. \tag{15.105}$$

By Lemmata 2.28, 2.31, by estimate (15.42), using the smallness condition (15.41) and $\tau_2 > \tau_1$ (see Theorem 15.4), the operators $\mathcal{U}_n^{\pm 1}$ satisfy, for all $s_0 \le s \le S$,

$$\|\mathcal{U}_n^{\pm1} h\|_s^{k_0,\gamma} \lesssim_S \|h\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu(\mathtt{b})}\|h\|^{k_0,\gamma}_{s_0}. \tag{15.106}$$

Therefore, by definition (15.105) and recalling (14.4), (15.106), (15.16), (15.17), the operators $\mathcal{V}_n^{\pm1}$ satisfy, for all $s_0 \le s \le S$, the estimate

$$\|\mathcal{V}_n^{\pm1} h\|_s^{k_0,\gamma} \lesssim_S \|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu(\mathtt{b})}\|h\|^{k_0,\gamma}_{s_0+\sigma}, \tag{15.107}$$

for some $\sigma = \sigma(k_0,\tau,\nu) > 0$.

In order to verify the inversion assumption (6.30)-(6.34) that is required to construct an approximate inverse (and thus to define the next approximate solution of the Nash-Moser nonlinear iteration), we decompose the operator $\mathcal{L}_n$ in (15.45) as

$$\mathcal{L}_n = \mathfrak{L}_n^{<} + \mathcal{R}_n + \mathcal{R}_n^{\perp} \tag{15.108}$$

where

$$\mathfrak{L}_n^{<} := \Pi_{K_n}\big(\omega\cdot\partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_n\big)\Pi_{K_n} + \Pi^{\perp}_{K_n}, \quad \mathcal{R}_n^{\perp} := \Pi^{\perp}_{K_n}\big(\omega\cdot\partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_n\big)\Pi^{\perp}_{K_n} - \Pi^{\perp}_{K_n}, \tag{15.109}$$

the diagonal operator $\mathcal{D}_n$ is defined in (15.22) (with $\nu = n$), and

$$K_n := K_0^{\chi^n}, \qquad K_0 > 0$$

is the scale of the nonlinear Nash-Moser iterative scheme.

Lemma 15.11. (First order Melnikov non-resonance conditions) *For all $\lambda = (\omega, \mathtt{h})$ in*

$$\Lambda^{\gamma,I}_{n+1} := \Lambda^{\gamma,I}_{n+1}(i) := \big\{\lambda \in \mathbb{R}^\nu \times [\mathtt{h}_1,\mathtt{h}_2] : |\omega\cdot\ell + \mu^n_j| \ge 2\gamma j^{\frac12}\langle\ell\rangle^{-\tau}, \quad \forall |\ell| \le K_n,\ j \in \mathbb{N}^+ \setminus \mathbb{S}^+\big\}, \tag{15.110}$$

the operator $\mathfrak{L}_n^{<}$ in (15.109) is invertible and there is an extension of the inverse operator (that we denote in the same way) to the whole $\mathbb{R}^\nu \times [\mathtt{h}_1,\mathtt{h}_2]$ satisfying the estimate

$$\|(\mathfrak{L}_n^{<})^{-1} g\|_s^{k_0,\gamma} \lesssim_{k_0} \gamma^{-1}\|g\|^{k_0,\gamma}_{s+\mu}, \tag{15.111}$$

where $\mu = k_0 + \tau(k_0+1)$ is the constant in (2.18) with $k_0 = k+1$.

Proof. By (15.68), similarly as in (15.69) one has $\gamma^{|\alpha|}|\partial_\lambda^\alpha(\omega\cdot\ell+\mu_j^n)|\lesssim\gamma\langle\ell\rangle|j|^{\frac12}$ for all $1\le|\alpha|\le k_0$. Hence Lemma A.4 can be applied to $f(\lambda)=\omega\cdot\ell+\mu_j^n(\lambda)$ with $M=C\gamma\langle\ell\rangle|j|^{\frac12}$ and $\rho=2\gamma j^{\frac12}\langle\ell\rangle^{-\tau}$. Thus, following the proof of Lemma 2.5 with $\omega\cdot\ell+\mu_j^n(\lambda)$ instead of $\omega\cdot\ell$, we obtain (15.111). □

Standard smoothing properties imply that the operator $\mathcal{R}_n^\perp$ defined in (15.109) satisfies, for all $b>0$,

$$\|\mathcal{R}_n^\perp h\|_{s_0}^{k_0,\gamma}\lesssim K_n^{-b}\|h\|_{s_0+b+1}^{k_0,\gamma}\,,\quad \|\mathcal{R}_n^\perp h\|_{s}^{k_0,\gamma}\lesssim \|h\|_{s+1}^{k_0,\gamma}\,. \tag{15.112}$$

By (15.105), (15.108), Theorem 15.5, Proposition 14.3, and estimates (15.111), (15.112), (15.107), we deduce the following theorem.

Theorem 15.12. (Almost-invertibility of $\mathcal{L}_\omega$) *Assume* (6.9). *Let* $\mathtt{a},\mathtt{b}$ *as in* (15.10) *and* M *as in* (15.16). *Let* $S>s_0$, *and assume the smallness condition* (15.41). *Then for all*

$$(\omega,\mathtt{h})\in\mathbf{\Lambda}_{n+1}^\gamma:=\mathbf{\Lambda}_{n+1}^\gamma(i):=\mathtt{\Lambda}_{n+1}^\gamma\cap\mathtt{\Lambda}_{n+1}^{\gamma,I} \tag{15.113}$$

(see (15.44), (15.110)*) the operator* $\mathcal{L}_\omega$ *defined in* (6.27) *(see also* (7.5)*) can be decomposed as*

$$\begin{aligned}&\mathcal{L}_\omega=\mathcal{L}_\omega^<+\mathcal{R}_\omega+\mathcal{R}_\omega^\perp\,,\\ &\mathcal{L}_\omega^<:=\mathcal{V}_n\mathfrak{L}_n^<\mathcal{V}_n^{-1}\,,\quad \mathcal{R}_\omega:=\mathcal{V}_n\mathcal{R}_n\mathcal{V}_n^{-1}\,,\quad \mathcal{R}_\omega^\perp:=\mathcal{V}_n\mathcal{R}_n^\perp\mathcal{V}_n^{-1}\,,\end{aligned} \tag{15.114}$$

where $\mathcal{L}_\omega^<$ *is invertible and there is an extension of the inverse operator (that we denote in the same way) to the whole* $\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$ *satisfying, for some* $\sigma:=\sigma(k_0,\tau,\nu)>0$ *and for all* $s_0\le s\le S$, *the estimates*

$$\|(\mathcal{L}_\omega^<)^{-1}g\|_s^{k_0,\gamma}\lesssim_S\gamma^{-1}\big(\|g\|_{s+\sigma}^{k_0,\gamma}+\|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|g\|_{s_0+\sigma}^{k_0,\gamma}\big) \tag{15.115}$$

(with $\mu(\mathtt{b})$ *defined in* (15.17)*) and*

$$\|\mathcal{R}_\omega h\|_s^{k_0,\gamma}\lesssim_S\varepsilon\gamma^{-2(M+1)}N_{n-1}^{-\mathtt{a}}\big(\|h\|_{s+\sigma}^{k_0,\gamma}+\|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,, \tag{15.116}$$

$$\|\mathcal{R}_\omega^\perp h\|_{s_0}^{k_0,\gamma}\lesssim_S K_n^{-b}\big(\|h\|_{s_0+b+\sigma}^{k_0,\gamma}+\|\mathfrak{I}_0\|_{s_0+\mu(\mathtt{b})+\sigma+b}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,,\qquad\forall b>0\,, \tag{15.117}$$

$$\|\mathcal{R}_\omega^\perp h\|_{s}^{k_0,\gamma}\lesssim_S\|h\|_{s+\sigma}^{k_0,\gamma}+\|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\,. \tag{15.118}$$

Notice that (15.116)-(15.118) *hold on the whole* $\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$.

This theorem provides the decomposition (6.30) with estimates (6.31)-(6.34). As a consequence, it allows to deduce Theorem 6.8.

16 The Nash-Moser iteration

In this section we prove Theorem 5.1. It will be a consequence of Theorem 16.2 below where we construct iteratively a sequence of better and better approximate solutions of the equation $\mathcal{F}(i,\alpha)=0$, with $\mathcal{F}(i,\alpha)$ defined in (5.13). We consider the finite-dimensional subspaces

$$E_n:=\Big\{\mathfrak{I}(\varphi)=(\Theta,I,z)(\varphi),\ \ \Theta=\Pi_n\Theta,\ I=\Pi_nI,\ z=\Pi_nz\Big\}$$

where Π_n is the projector

$$\Pi_n:=\Pi_{K_n}:z(\varphi,x)=\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{S}_0^c}z_{\ell,j}e^{\mathrm{i}(\ell\cdot\varphi+jx)}\ \mapsto\Pi_nz(\varphi,x):=\sum_{|(\ell,j)|\le K_n}z_{\ell,j}e^{\mathrm{i}(\ell\cdot\varphi+jx)} \tag{16.1}$$

with $K_n=K_0^{\chi^n}$ (see (6.29)) and we denote with the same symbol $\Pi_np(\varphi):=\sum_{|\ell|\le K_n}p_\ell e^{\mathrm{i}\ell\cdot\varphi}$. We define $\Pi_n^\perp:=\mathrm{Id}-\Pi_n$. The projectors Π_n, $\Pi_n^\perp$ satisfy the smoothing properties (2.6), (2.7) for the weighted Whitney-Sobolev norm $\|\cdot\|_s^{k_0,\gamma}$ defined in (2.3).

In view of the Nash-Moser Theorem 16.2 we introduce the following constants:

$$\mathtt{a}_1 := \max\{6\sigma_1 + 13, \chi p(\tau+1)(4\mathtt{d}+1) + \chi(\mu(\mathtt{b}) + 2\sigma_1) + 1\}, \qquad \mathtt{a}_2 := \chi^{-1}\mathtt{a}_1 - \mu(\mathtt{b}) - 2\sigma_1, \tag{16.2}$$

$$\mu_1 := 3(\mu(\mathtt{b}) + 2\sigma_1) + 1, \qquad \mathtt{b}_1 := \mathtt{a}_1 + \mu(\mathtt{b}) + 3\sigma_1 + 3 + \chi^{-1}\mu_1, \qquad \chi = 3/2, \tag{16.3}$$

$$\sigma_1 := \max\{\bar\sigma\,, s_0 + 2k_0 + 5\}, \tag{16.4}$$

where $\bar\sigma := \bar\sigma(\tau,\nu,k_0) > 0$ is defined in Theorem 6.8, $s_0 + 2k_0 + 5$ is the largest loss of regularity in the estimates of the Hamiltonian vector field X_P in Lemma 6.1, $\mu(\mathtt{b})$ is defined in (15.17), $\mathtt{b}$ is the constant $\mathtt{b} := [\mathtt{a}] + 2 \in \mathbb{N}$ where $\mathtt{a}$ is defined in (15.10), and the exponent p in (6.28) satisfies

$$p\mathtt{a} > (\chi - 1)\mathtt{a}_1 + \chi\sigma_1 = \frac{1}{2}\mathtt{a}_1 + \frac{3}{2}\sigma_1\,. \tag{16.5}$$

By (15.10), $\mathtt{a} \geq \chi(\tau+1)(4\mathtt{d}+1) + 1$. Hence, by the definition of $\mathtt{a}_1$ in (16.2), there exists $p := p(\tau,\nu,k_0)$ such that (16.5) holds. For example we fix

$$p := \frac{3(\mu(\mathtt{b}) + 3\sigma_1 + 1)}{\mathtt{a}}\,. \tag{16.6}$$

Remark 16.1. *The constant* $\mathtt{a}_1$ *is the exponent in* (16.11). *The constant* $\mathtt{a}_2$ *is the exponent in* (16.9). *The constant* μ_1 *is the exponent in* $(\mathcal{P}3)_n$. *The choice of the constants* $\mu_1, \mathtt{b}_1, \mathtt{a}_1$ *allows the convergence of the iterative scheme* (16.22)-(16.23), *see Lemma 16.4. The conditions required along the iteration are* $\mathtt{a}_1 > (2\sigma_1+4)\chi/(2-\chi) = 6\sigma_1 + 12$, $\mathtt{b}_1 > \mathtt{a}_1 + \mu(\mathtt{b}) + 3\sigma_1 + 2 + \chi^{-1}\mu_1$, *as well as* $p\mathtt{a} > (\chi-1)\mathtt{a}_1 + \chi\sigma_1$ *and* $\mu_1 > 3(\mu(\mathtt{b}) + 2\sigma_1)$.

In addition we require $\mathtt{a}_1 \geq \chi p(\tau+1)(4\mathtt{d}+1) + \chi(\mu(\mathtt{b}) + 2\sigma_1)$ *so that* $\mathtt{a}_2 > p(\tau+1)(4\mathtt{d}+1)$. *This condition is used in the proof of Lemma 16.5.*

In this section, given $W = (\mathfrak{I}, \beta)$ where $\mathfrak{I} = \mathfrak{I}(\lambda)$ is the periodic component of a torus as in (5.15), and $\beta = \beta(\lambda) \in \mathbb{R}^\nu$ we denote $\|W\|_s^{k_0,\gamma} := \max\{\|\mathfrak{I}\|_s^{k_0,\gamma}, |\beta|^{k_0,\gamma}\}$, where $\|\mathfrak{I}\|_s^{k_0,\gamma}$ is defined in (5.16).

Theorem 16.2. (Nash-Moser) *There exist* δ_0, $C_* > 0$, *such that, if*

$$K_0^{\tau_3}\varepsilon\gamma^{-2M-3} < \delta_0, \quad \tau_3 := \max\{p\tau_2, 2\sigma_1 + \mathtt{a}_1 + 4\}, \quad K_0 := \gamma^{-1}, \quad \gamma := \varepsilon^a, \quad 0 < a < \frac{1}{\tau_3 + 2M + 3}\,, \tag{16.7}$$

where the constant M *is defined in* (15.16) *and* $\tau_2 := \tau_2(\tau,\nu)$ *is defined in Theorem 15.4, then, for all* $n \geq 0$*:*

$(\mathcal{P}1)_n$ *there exists a* k_0 *times differentiable function* $\tilde W_n : \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] \to E_{n-1} \times \mathbb{R}^\nu$, $\lambda = (\omega, \mathtt{h}) \mapsto \tilde W_n(\lambda) := (\tilde{\mathfrak{I}}_n, \tilde\alpha_n - \omega)$, *for* $n \geq 1$, *and* $\tilde W_0 := 0$, *satisfying*

$$\|\tilde W_n\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \leq C_*\varepsilon\gamma^{-1}\,. \tag{16.8}$$

Let $\tilde U_n := U_0 + \tilde W_n$ *where* $U_0 := (\varphi, 0, 0, \omega)$. *The difference* $\tilde H_n := \tilde U_n - \tilde U_{n-1}$, $n \geq 1$, *satisfies*

$$\|\tilde H_1\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \leq C_*\varepsilon\gamma^{-1}\,, \quad \|\tilde H_n\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \leq C_*\varepsilon\gamma^{-1}K_{n-1}^{-\mathtt{a}_2}\,, \quad \forall n \geq 2. \tag{16.9}$$

$(\mathcal{P}2)_n$ *Setting* $\tilde\imath_n := (\varphi, 0, 0) + \tilde{\mathfrak{I}}_n$, *we define*

$$\mathcal{G}_0 := \Omega \times [\mathtt{h}_1, \mathtt{h}_2]\,, \quad \mathcal{G}_{n+1} := \mathcal{G}_n \cap \mathbf{\Lambda}_{n+1}^\gamma(\tilde\imath_n)\,, \quad n \geq 0\,, \tag{16.10}$$

where $\mathbf{\Lambda}_{n+1}^\gamma(\tilde\imath_n)$ *is defined in* (15.113). *Then, for all* $\lambda \in \mathcal{G}_n$, *setting* $K_{-1} := 1$, *we have*

$$\|\mathcal{F}(\tilde U_n)\|_{s_0}^{k_0,\gamma} \leq C_*\varepsilon K_{n-1}^{-\mathtt{a}_1}\,. \tag{16.11}$$

$(\mathcal{P}3)_n$ (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* $\lambda \in \mathcal{G}_n$.

Proof. To simplify notation, in this proof we denote $\| \|^{k_0,\gamma}$ by $\| \|$.

STEP 1: *Proof of* $(\mathcal{P}1, \mathcal{P}2, \mathcal{P}3)_0$. One has $\|\mathcal{F}(U_0)\|_s \lesssim \varepsilon$ by (5.13), (6.2), then take C_* large enough.

STEP 2: *Assume that* $(\mathcal{P}1, \mathcal{P}2, \mathcal{P}3)_n$ *hold for some* $n \geq 0$, *and prove* $(\mathcal{P}1, \mathcal{P}2, \mathcal{P}3)_{n+1}$. We are going to define the next approximation $\tilde{U}_{n+1}$ by a modified Nash-Moser scheme. To this aim, we prove the almost-approximate invertibility of the linearized operator

$$L_n := L_n(\lambda) := d_{i,\alpha}\mathcal{F}(\tilde{\imath}_n(\lambda))$$

by applying Theorem 6.8 to $L_n(\lambda)$. To prove that the inversion assumptions (6.30)-(6.34) hold, we apply Theorem 15.12 with $i = \tilde{\imath}_n$. By (16.7) (and recalling the relation $N_0 = K_0^p$ in (6.28)), the smallness condition (15.41) holds for ε small enough. Therefore Theorem 15.12 applies, and we deduce that (6.30)-(6.34) hold for all $\lambda \in \mathbf{\Lambda}^{\gamma}_{n+1}(\tilde{\imath}_n)$, see (15.113).

Now we apply Theorem 6.8 to the linearized operator $L_n(\lambda)$ with $\Lambda_o = \mathbf{\Lambda}^{\gamma}_{n+1}(\tilde{\imath}_n)$ and

$$S = s_0 + \mathtt{b}_1 \quad \text{where } \mathtt{b}_1 \text{ is defined in (16.3).} \tag{16.12}$$

It implies the existence of an almost-approximate inverse $\mathbf{T}_n := \mathbf{T}_n(\lambda, \tilde{\imath}_n(\lambda))$ satisfying

$$\|\mathbf{T}_n g\|_s \lesssim_{s_0+\mathtt{b}_1} \gamma^{-1}\big(\|g\|_{s+\sigma_1} + \|\tilde{\mathfrak{I}}_n\|_{s+\mu(\mathtt{b})+\sigma_1}\|g\|_{s_0+\sigma_1}\big) \quad \forall s_0 < s \leq s_0 + \mathtt{b}_1\,, \tag{16.13}$$

$$\|\mathbf{T}_n g\|_{s_0} \lesssim_{s_0+\mathtt{b}_1} \gamma^{-1}\|g\|_{s_0+\sigma_1} \tag{16.14}$$

because $\sigma_1 \geq \bar{\sigma}$ by (16.4), where $\bar{\sigma}$ is the loss in (6.47). For all $\lambda \in \mathcal{G}_{n+1} = \mathcal{G}_n \cap \mathbf{\Lambda}^{\gamma}_{n+1}(\tilde{\imath}_n)$ (see (16.10)), we define

$$U_{n+1} := \tilde{U}_n + H_{n+1}\,, \quad H_{n+1} := (\widehat{\mathfrak{I}}_{n+1}, \widehat{\alpha}_{n+1}) := -\mathbf{\Pi}_n\mathbf{T}_n\Pi_n\mathcal{F}(\tilde{U}_n) \in E_n \times \mathbb{R}^{\nu} \tag{16.15}$$

where $\mathbf{\Pi}_n$ is defined by (see (16.1))

$$\mathbf{\Pi}_n(\mathfrak{I}, \alpha) := (\Pi_n\mathfrak{I}, \alpha)\,, \quad \mathbf{\Pi}_n^{\perp}(\mathfrak{I}, \alpha) := (\Pi_n^{\perp}\mathfrak{I}, 0)\,, \quad \forall(\mathfrak{I}, \alpha)\,. \tag{16.16}$$

We show that the iterative scheme in (16.15) is rapidly converging. We write

$$\mathcal{F}(U_{n+1}) = \mathcal{F}(\tilde{U}_n) + L_nH_{n+1} + Q_n \tag{16.17}$$

where $L_n := d_{i,\alpha}\mathcal{F}(\tilde{U}_n)$ and Q_n is defined by difference. Then, by the definition of H_{n+1} in (16.15), we have (recall also (16.16))

$$\begin{aligned}
\mathcal{F}(U_{n+1}) &= \mathcal{F}(\tilde{U}_n) - L_n\mathbf{\Pi}_n\mathbf{T}_n\Pi_n\mathcal{F}(\tilde{U}_n) + Q_n \\
&= \mathcal{F}(\tilde{U}_n) - L_n\mathbf{T}_n\Pi_n\mathcal{F}(\tilde{U}_n) + L_n\mathbf{\Pi}_n^{\perp}\mathbf{T}_n\Pi_n\mathcal{F}(\tilde{U}_n) + Q_n \\
&= \mathcal{F}(\tilde{U}_n) - \Pi_nL_n\mathbf{T}_n\Pi_n\mathcal{F}(\tilde{U}_n) + (L_n\mathbf{\Pi}_n^{\perp} - \Pi_n^{\perp}L_n)\mathbf{T}_n\Pi_n\mathcal{F}(\tilde{U}_n) + Q_n \\
&= \Pi_n^{\perp}\mathcal{F}(\tilde{U}_n) + R_n + Q_n + P_n
\end{aligned} \tag{16.18}$$

where

$$R_n := (L_n\mathbf{\Pi}_n^{\perp} - \Pi_n^{\perp}L_n)\mathbf{T}_n\Pi_n\mathcal{F}(\tilde{U}_n)\,, \qquad P_n := -\Pi_n(L_n\mathbf{T}_n - \mathrm{Id})\Pi_n\mathcal{F}(\tilde{U}_n)\,. \tag{16.19}$$

We first note that, for all $\lambda \in \Omega \times [\mathtt{h}_1, \mathtt{h}_2]$, $s \geq s_0$, by triangular inequality and by (5.13), (6.3), (16.4), (16.8) we have

$$\|\mathcal{F}(\tilde{U}_n)\|_s \lesssim_s \|\mathcal{F}(U_0)\|_s + \|\mathcal{F}(\tilde{U}_n) - \mathcal{F}(U_0)\|_s \lesssim_s \varepsilon + \|\tilde{W}_n\|_{s+\sigma_1} \tag{16.20}$$

and, by (16.8), (16.7),

$$\gamma^{-1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0} \leq 1\,. \tag{16.21}$$

We now prove the following inductive estimates of Nash-Moser type.

Lemma 16.3. *For all* $\lambda \in \mathcal{G}_{n+1}$ *we have, setting* $\mu_2 := \mu(\mathtt{b}) + 3\sigma_1$,

$$\|\mathcal{F}(U_{n+1})\|_{s_0} \lesssim_{s_0+\mathtt{b}_1} \frac{K_n^{\mu_2-\mathtt{b}_1}}{\gamma}(\varepsilon + \|\tilde{W}_n\|_{s_0+\mathtt{b}_1}) + K_n^{2\sigma_1+2}\|\mathcal{F}(\tilde{U}_n)\|^2_{s_0} + \frac{\varepsilon}{\gamma^{2M+3}}K_{n-1}^{-p\mathtt{a}}K_n^{\sigma_1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0} \tag{16.22}$$

$$\|W_1\|_{s_0+\mathtt{b}_1} \lesssim_{s_0+\mathtt{b}_1} \varepsilon\gamma^{-1}\,, \quad \|W_{n+1}\|_{s_0+\mathtt{b}_1} \lesssim_{s_0+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+2\sigma_1}\gamma^{-1}(\varepsilon + \|\tilde{W}_n\|_{s_0+\mathtt{b}_1})\,,\ n \geq 1\,. \tag{16.23}$$

Proof. We first estimate H_{n+1} defined in (16.15).

Estimates of H_{n+1}. By (16.15) and (2.6), (2.7), (16.13), (16.14), (16.8), we get

$$\begin{aligned}\|H_{n+1}\|_{s_0+\mathtt{b}_1} &\lesssim_{s_0+\mathtt{b}_1} \gamma^{-1}\big(K_n^{\sigma_1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0+\mathtt{b}_1} + K_n^{\mu(\mathtt{b})+2\sigma_1}\|\tilde{\mathfrak{I}}_n\|_{s_0+\mathtt{b}_1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0}\big)\\ &\overset{(16.20),(16.21)}{\lesssim_{s_0+\mathtt{b}_1}} K_n^{\mu(\mathtt{b})+2\sigma_1}\gamma^{-1}\big(\varepsilon + \|\tilde{W}_n\|_{s_0+\mathtt{b}_1}\big)\,, \qquad (16.24)\\ \|H_{n+1}\|_{s_0} &\lesssim_{s_0+\mathtt{b}_1} \gamma^{-1}K_n^{\sigma_1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0}\,. \qquad (16.25)\end{aligned}$$

Now we estimate the terms Q_n in (16.17) and P_n, R_n in (16.19) in $\| \ \|_{s_0}$ norm.

Estimate of Q_n. By (16.8), (16.4), (2.6), (16.25), (16.11), and since $3\sigma_1 - \mathtt{a}_1 \le 0$ and $a < 1/(1+3\sigma_1)$ (see (16.2), (16.7)), one has $\|\tilde{W}_n + tH_{n+1}\|_{2s_0+2k_0+5} \lesssim 1$ for all $t \in [0,1]$. Hence, by Taylor's formula, using (16.17), (5.13), (6.4), (16.25), (2.6), and $\varepsilon\gamma^{-2} \le 1$, we get

$$\|Q_n\|_{s_0} \lesssim_{s_0+\mathtt{b}_1} \varepsilon\|H_{n+1}\|^2_{s_0+1} \lesssim_{s_0+\mathtt{b}_1} K_n^{2\sigma_1+2}\|\mathcal{F}(\tilde{U}_n)\|^2_{s_0}\,. \qquad (16.26)$$

Estimate of P_n. By (6.48), $L_n\mathbf{T}_n - \mathrm{Id} = \mathcal{P}(\tilde{\imath}_n) + \mathcal{P}_\omega(\tilde{\imath}_n) + \mathcal{P}^\perp_\omega(\tilde{\imath}_n)$. Accordingly, we decompose P_n in (16.19) as $P_n = -P_n^{(1)} - P_{n,\omega} - P^\perp_{n,\omega}$, where

$$P_n^{(1)} := \Pi_n\mathcal{P}(\tilde{\imath}_n)\Pi_n\mathcal{F}(\tilde{U}_n), \qquad P_{n,\omega} := \Pi_n\mathcal{P}_\omega(\tilde{\imath}_n)\Pi_n\mathcal{F}(\tilde{U}_n), \qquad P^\perp_{n,\omega} := \Pi_n\mathcal{P}^\perp_\omega(\tilde{\imath}_n)\Pi_n\mathcal{F}(\tilde{U}_n).$$

By (2.6)-(2.7),

$$\|\mathcal{F}(\tilde{U}_n)\|_{s_0+\sigma_1} \le \|\Pi_n\mathcal{F}(\tilde{U}_n)\|_{s_0+\sigma_1} + \|\Pi_n^\perp\mathcal{F}(\tilde{U}_n)\|_{s_0+\sigma_1} \le K_n^{\sigma_1}(\|\mathcal{F}(\tilde{U}_n)\|_{s_0} + K_n^{-\mathtt{b}_1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0+\mathtt{b}_1}). \quad (16.27)$$

By (6.49), (16.8), (16.27), and then (16.20), (2.6), we obtain

$$\begin{aligned}\|P_n^{(1)}\|_{s_0} &\lesssim_{s_0+\mathtt{b}_1} \gamma^{-1}K_n^{2\sigma_1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0}(\|\mathcal{F}(\tilde{U}_n)\|_{s_0} + K_n^{-\mathtt{b}_1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0+\mathtt{b}_1})\\ &\lesssim_{s_0+\mathtt{b}_1} \gamma^{-1}K_n^{2\sigma_1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0}(\|\mathcal{F}(\tilde{U}_n)\|_{s_0} + K_n^{\sigma_1-\mathtt{b}_1}(\varepsilon + \|\tilde{W}_n\|_{s_0+\mathtt{b}_1})). \qquad (16.28)\end{aligned}$$

By (6.50), (16.8), (2.6), we have

$$\|P_{n,\omega}\|_{s_0} \lesssim_{s_0+\mathtt{b}_1} \varepsilon\gamma^{-2M-3}N_{n-1}^{-\mathtt{a}}K_n^{\sigma_1}\|\mathcal{F}(\tilde{U}_n)\|_{s_0}\,, \qquad (16.29)$$

where $\mathtt{a}$ is defined in (15.10). By (6.51), (2.6), (16.3), (16.11), and then (16.20), (2.6), we get

$$\begin{aligned}\|P^\perp_{n,\omega}\|_{s_0} &\lesssim_{s_0+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+2\sigma_1-\mathtt{b}_1}\gamma^{-1}(\|\mathcal{F}(\tilde{U}_n)\|_{s_0+\mathtt{b}_1} + \varepsilon\|\tilde{W}_n\|_{s_0+\mathtt{b}_1})\\ &\lesssim_{s_0+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+3\sigma_1-\mathtt{b}_1}\gamma^{-1}(\varepsilon + \|\tilde{W}_n\|_{s_0+\mathtt{b}_1}). \qquad (16.30)\end{aligned}$$

Estimate of R_n. For $H := (\widehat{\mathfrak{I}}, \widehat{\alpha})$ we have $(L_n\mathbf{\Pi}_n^\perp - \Pi_n^\perp L_n)H = \varepsilon[d_iX_P(\tilde{\imath}_n), \Pi_n^\perp]\widehat{\mathfrak{I}}$ where X_P is the Hamiltonian vector field in (5.13). By (6.3), (2.7), (16.4), (16.8),

$$\|(L_n\mathbf{\Pi}_n^\perp - \Pi_n^\perp L_n)H\|_{s_0} \lesssim_{s_0+\mathtt{b}_1} \varepsilon K_n^{-\mathtt{b}_1+\sigma_1}(\|\widehat{\mathfrak{I}}\|_{s_0+\mathtt{b}_1} + \|\tilde{\mathfrak{I}}_n\|_{s_0+\mathtt{b}_1}\|\widehat{\mathfrak{I}}\|_{s_0+1})\,. \qquad (16.31)$$

Hence, by (16.19), (16.31), (16.13), (16.8), (2.6), and then (16.20), (2.6), (16.21), we get

$$\begin{aligned}\|R_n\|_{s_0} &\lesssim_{s_0+\mathtt{b}_1} \varepsilon\gamma^{-1}K_n^{\mu(\mathtt{b})+2\sigma_1-\mathtt{b}_1}(\|\Pi_n\mathcal{F}(\tilde{U}_n)\|_{s_0+\mathtt{b}_1} + \|\tilde{W}_n\|_{s_0+\mathtt{b}_1}\|\Pi_n\mathcal{F}(\tilde{U}_n)\|_{s_0+\sigma_1})\\ &\lesssim_{s_0+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+3\sigma_1-\mathtt{b}_1}(\varepsilon + \|\tilde{W}_n\|_{s_0+\mathtt{b}_1}). \qquad (16.32)\end{aligned}$$

We can finally estimate $\mathcal{F}(U_{n+1})$. By (16.18), (2.7), (16.20), (16.32), (16.26), (16.28)-(16.30), (16.8), we get (16.22). By (16.15) and (16.13) we have bound (16.23) for $W_1 := H_1$, namely

$$\|W_1\|_{s_0+\mathtt{b}_1} = \|H_1\|_{s_0+\mathtt{b}_1} \lesssim_{s_0+\mathtt{b}_1} \gamma^{-1}\|\mathcal{F}(U_0)\|_{s_0+\mathtt{b}_1+\sigma_1} \lesssim_{s_0+\mathtt{b}_1} \varepsilon\gamma^{-1}\,.$$

Estimate (16.23) for $W_{n+1} := \tilde{W}_n + H_{n+1}$, $n \ge 1$, follows by (16.24). ∎

Now that Lemma 16.3 has been proved, we continue the proof of Theorem 16.2. As a corollary of Lemma 16.3 we get the following lemma, where for clarity we use the extended notation $\| \ \|^{k_0,\gamma}$ (instead of $\|\|\|$ used above).

Lemma 16.4. *For all $\lambda \in \mathcal{G}_{n+1}$, we have*

$$\|\mathcal{F}(U_{n+1})\|_{s_0}^{k_0,\gamma} \le C_* \varepsilon K_n^{-\mathtt{a}_1}\,, \qquad \|W_{n+1}\|_{s_0+\mathtt{b}_1}^{k_0,\gamma} \le C_* \varepsilon \gamma^{-1} K_n^{\mu_1}\,, \tag{16.33}$$

$$\|H_1\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \le C\varepsilon\gamma^{-1}\,, \quad \|H_{n+1}\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \lesssim_{s_0} \varepsilon\gamma^{-1} K_n^{\mu(\mathtt{b})+2\sigma_1} K_{n-1}^{-\mathtt{a}_1}\,, \quad n \ge 1\,. \tag{16.34}$$

Proof. First note that, by (16.10), if $\lambda \in \mathcal{G}_{n+1}$, then $\lambda \in \mathcal{G}_n$ and so (16.11) and the inequality in $(\mathcal{P}3)_n$ hold. Then the first inequality in (16.33) follows by (16.22), $(\mathcal{P}2)_n$, $(\mathcal{P}3)_n$, $\gamma^{-1} = K_0 \le K_n$, $\varepsilon\gamma^{-2M-3} \le c$ small, and by (16.2), (16.3), (16.5)-(16.6) (see also Remark 16.1). For $n = 0$ we use also (16.7).

The second inequality in (16.33) for $n = 0$ follows directly from the bound for W_1 in (16.23); for $n = 1, 2$ one proves, by (16.23), that

$$\|W_2\|_{s_0+\mathtt{b}_1}^{k_0,\gamma} \lesssim_{s_0+\mathtt{b}_1} \varepsilon\gamma^{-2} K_1^{\mu(\mathtt{b})+2\sigma_1}, \quad \|W_3\|_{s_0+\mathtt{b}_1}^{k_0,\gamma} \lesssim_{s_0+\mathtt{b}_1} \varepsilon\gamma^{-3} (K_2 K_1)^{\mu(\mathtt{b})+2\sigma_1},$$

whence the second inequality in (16.33) for $n = 1, 2$ follows by the choice of μ_1 in (16.3) and $K_0 = \gamma^{-1}$ large enough (i.e., ε small enough); the second inequality in (16.33) for $n \ge 3$ is proved inductively by using (16.23), $(\mathcal{P}3)_n$, the choice of μ_1 in (16.3) and K_0 large enough.

Since $H_1 = W_1$, the first inequality in (16.34) follows by the first inequality in (16.23). For $n \ge 1$, estimate (16.34) follows by (2.6), (16.25) and (16.11). ☐

By Theorem A.2, we define a k_0 times differentiable extension $\tilde{H}_{n+1}$ of $(H_{n+1})_{|\mathcal{G}_{n+1}}$ to the whole $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, which satisfies the same bound for H_{n+1} in (16.34) and therefore, by the definition of $\mathtt{a}_2$ in (16.2), the estimate (16.9) at $n + 1$ holds.

Finally we define the functions

$$\tilde{W}_{n+1} := \tilde{W}_n + \tilde{H}_{n+1}\,, \quad \tilde{U}_{n+1} := \tilde{U}_n + \tilde{H}_{n+1} = U_0 + \tilde{W}_n + \tilde{H}_{n+1} = U_0 + \tilde{W}_{n+1}\,,$$

which are defined for all $\lambda \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ and satisfy

$$\tilde{W}_{n+1} = W_{n+1}\,, \quad \tilde{U}_{n+1} = U_{n+1}\,, \quad \forall \lambda \in \mathcal{G}_{n+1}\,.$$

Therefore $(\mathcal{P}2)_{n+1}$, $(\mathcal{P}3)_{n+1}$ are proved by Lemma 16.4. Moreover by (16.9), which has been proved up to the step $n + 1$, we have

$$\|\tilde{W}_{n+1}\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \le \sum_{k=1}^{n+1} \|\tilde{H}_k\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \le C_* \varepsilon\gamma^{-1}$$

and thus (16.8) holds also at the step $n + 1$. This completes the proof of Theorem 16.2. ☐

16.1 Proof of Theorem 5.1

Let $\gamma = \varepsilon^a$ with $a \in (0, a_0)$ and $a_0 := 1/(2M + 3 + \tau_3)$ where τ_3 is defined in (16.7). Then the smallness condition given by the first inequality in (16.7) holds for $0 < \varepsilon < \varepsilon_0$ small enough and Theorem 16.2 applies. By (16.9) the sequence of functions

$$\tilde{W}_n = \tilde{U}_n - (\varphi, 0, 0, \omega) := (\tilde{\mathfrak{I}}_n, \tilde{\alpha}_n - \omega) = (\tilde{\imath}_n - (\varphi, 0, 0), \tilde{\alpha}_n - \omega)$$

is a Cauchy sequence in $\| \ \|_{s_0}^{k_0,\gamma}$ and then it converges to a function

$$W_\infty := \lim_{n\to+\infty} \tilde{W}_n\,, \qquad \text{with} \quad W_\infty : \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] \to H_\varphi^{s_0} \times H_\varphi^{s_0} \times H_{\varphi,x}^{s_0} \times \mathbb{R}^\nu\,.$$

We define

$$U_\infty := (i_\infty, \alpha_\infty) = (\varphi, 0, 0, \omega) + W_\infty\,.$$

By (16.8) and (16.9) we also deduce that

$$\|U_\infty - U_0\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \le C_* \varepsilon\gamma^{-1}\,, \quad \|U_\infty - \tilde U_n\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \le C\varepsilon\gamma^{-1} K_n^{-\mathtt{a}_2}\,, \quad n \ge 1\,. \tag{16.35}$$

Moreover by Theorem 16.2-$(\mathcal{P}2)_n$, we deduce that $\mathcal{F}(\lambda, U_\infty(\lambda)) = 0$ for all λ belonging to

$$\bigcap_{n\ge 0} \mathcal{G}_n = \mathcal{G}_0 \cap \bigcap_{n\ge1} \mathbf{\Lambda}^\gamma_n(\tilde\imath_{n-1}) \overset{(15.113)}{=} \mathcal{G}_0 \cap \Big[\bigcap_{n\ge1} \mathtt{\Lambda}^\gamma_n(\tilde\imath_{n-1})\Big] \cap \Big[\bigcap_{n\ge1} \mathtt{\Lambda}^{\gamma,I}_n(\tilde\imath_{n-1})\Big]\,, \tag{16.36}$$

where $\mathcal{G}_0 = \mathtt{\Omega} \times [\mathtt{h}_1, \mathtt{h}_2]$ is defined in (16.10). By the first inequality in (16.35) we deduce estimates (5.19) and (5.20).

To conclude the proof of Theorem 5.1, now we prove that the Cantor set $\mathcal{C}^\gamma_\infty$ in (5.23) is contained in $\bigcap_{n\ge0} \mathcal{G}_n$. We first consider the set

$$\mathcal{G}_\infty := \mathcal{G}_0 \cap \Big[\bigcap_{n\ge1} \mathtt{\Lambda}^{2\gamma}_n(i_\infty)\Big] \cap \Big[\bigcap_{n\ge1} \mathtt{\Lambda}^{2\gamma,I}_n(i_\infty)\Big]\,. \tag{16.37}$$

Lemma 16.5. *$\mathcal{G}_\infty \subseteq \bigcap_{n\ge0} \mathcal{G}_n$, where $\mathcal{G}_n$ is defined in* (16.10).

Proof. We are going to apply the inclusion property (15.39). By (16.35), (6.28), we have, for all $n \ge 2$,

$$C(S) N_{n-1}^{(\tau+1)(4\mathtt{d}+1)} \gamma^{-4\mathtt{d}} \|i_\infty - \tilde\imath_{n-1}\|_{s_0+\mu(\mathtt{b})+\sigma_1} \le C(S) K_{n-1}^{p(\tau+1)(4\mathtt{d}+1)} C\varepsilon\gamma^{-1-4\mathtt{d}} K_{n-1}^{-\mathtt{a}_2} \le \gamma$$

taking ε small enough, by (16.7) and using $\mathtt{a}_2 \ge p(\tau+1)(4\mathtt{d}+1)$ (see (16.2)). For $n = 1$ we get as well $C(S) N_0^{(\tau+1)(4\mathtt{d}+1)} \gamma^{-4\mathtt{d}} \|i_\infty - \tilde\imath_0\|_{s_0+\mu(\mathtt{b})+\sigma_1} \le \gamma$ using the first inequality in (16.35) and recalling that $K_0 = \gamma^{-1}$, $\gamma = \varepsilon^a$ and $a[2+4\mathtt{d}+p(\tau+1)(4\mathtt{d}+1)] < 1$. Recall also that S has been fixed in (16.12). Therefore (15.39) in Theorem 15.4-$(\mathbf{S3})_\nu$ gives

$$\mathtt{\Lambda}^{2\gamma}_n(i_\infty) \subseteq \mathtt{\Lambda}^\gamma_n(\tilde\imath_{n-1})\,, \quad \forall n \ge 1\,.$$

By similar arguments we deduce that $\mathtt{\Lambda}^{2\gamma,I}_n(i_\infty) \subseteq \mathtt{\Lambda}^{\gamma,I}_n(\tilde\imath_{n-1})$, and the lemma is proved. ☐

Then we define the "final eigenvalues"

$$\mu^\infty_j := \mu^0_j(i_\infty) + r^\infty_j\,, \quad j \in \mathbb{N}^+ \setminus \mathbb{S}^+\,, \tag{16.38}$$

where $\mu^0_j(i_\infty)$ are defined in (15.8) (with $\mathtt{m}_{\frac12}, r_j$ depending on i_∞) and

$$r^\infty_j := \lim_{n\to+\infty} r^n_j(i_\infty)\,, \quad j \in \mathbb{N}^+ \setminus \mathbb{S}^+\,, \tag{16.39}$$

with r^n_j given in Theorem 15.4-$(\mathbf{S1})_\nu$. Note that the sequence $(r^n_j(i_\infty))_{n\in\mathbb{N}}$ is a Cauchy sequence in $|\ |^{k_0,\gamma}$ by (15.25). As a consequence its limit function $r^\infty_j(\omega, \mathtt{h})$ is well defined, it is k_0 times differentiable and satisfies

$$|r^\infty_j - r^n_j(i_\infty)|^{k_0,\gamma} \le C\varepsilon\gamma^{-2(M+1)} |j|^{-2\mathtt{m}} N_{n-1}^{-\mathtt{a}}\,, \ n \ge 0\,. \tag{16.40}$$

In particular, since $r^0_j(i_\infty) = 0$, we get $|r^\infty_j|^{k_0,\gamma} \le C\varepsilon\gamma^{-2(M+1)}|j|^{-2\mathtt{m}}$ (here $C := C(S, k_0)$, with S fixed in (16.12)). Now consider the *final Cantor set* $\mathcal{C}^\gamma_\infty$ in (5.23).

Lemma 16.6. *$\mathcal{C}^\gamma_\infty \subseteq \mathcal{G}_\infty$, where $\mathcal{G}_\infty$ is defined in* (16.37).

Proof. By (16.37), we have to prove that $\mathcal{C}^\gamma_\infty \subseteq \mathtt{\Lambda}^{2\gamma}_n(i_\infty)$, $\forall n \in \mathbb{N}$. We argue by induction. For $n = 0$ the inclusion is trivial, since $\mathtt{\Lambda}^{2\gamma}_0(i_\infty) = \mathtt{\Omega} \times [\mathtt{h}_1, \mathtt{h}_2] = \mathcal{G}_0$. Now assume that $\mathcal{C}^\gamma_\infty \subseteq \mathtt{\Lambda}^{2\gamma}_n(i_\infty)$ for some $n \ge 0$. For all $\lambda \in \mathcal{C}^\gamma_\infty \subseteq \mathtt{\Lambda}^{2\gamma}_n(i_\infty)$, by (15.23), (16.38), (16.40), we get

$$|(\mu^n_j - \mu^n_{j'})(i_\infty) - (\mu^\infty_j - \mu^\infty_{j'})| \le C\varepsilon\gamma^{-2(M+1)} N_{n-1}^{-\mathtt{a}} \big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big)$$

Therefore, for any $|\ell|, |j-j'| \leq N_n$ with $(\ell, j, j') \neq (0, j, j)$ (recall (5.23)) we have

$$\begin{aligned} |\omega \cdot \ell + \mu_j^n(i_\infty) - \mu_{j'}^n(i_\infty)| &\geq |\omega \cdot \ell + \mu_j^\infty - \mu_{j'}^\infty| - C\varepsilon\gamma^{-2(M+1)} N_{n-1}^{-\mathtt{a}} \big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big) \\ &\geq 4\gamma \langle \ell \rangle^{-\tau} j^{-\mathtt{d}} j'^{-\mathtt{d}} - C\varepsilon\gamma^{-2(M+1)} N_{n-1}^{-\mathtt{a}} \big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big) \\ &\geq 2\gamma \langle \ell \rangle^{-\tau} j^{-\mathtt{d}} j'^{-\mathtt{d}} \end{aligned}$$

provided

$$C\varepsilon\gamma^{-2M-3} N_{n-1}^{-\mathtt{a}} N_n^{\tau} \big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big) j^{\mathtt{d}} j'^{\mathtt{d}} \leq 1 .$$

Since $\mathtt{m} > \mathtt{d}$ (see (15.10)), one has $(j+N_n)^{\mathtt{d}} j^{\mathtt{d}-2\mathtt{m}} \lesssim_{\mathtt{d}} N_n^{\mathtt{d}}$ for all $j \geq 1$. Hence, using $|j-j'| \leq N_n$,

$$\big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big) j^{\mathtt{d}} j'^{\mathtt{d}} = \frac{j'^{\mathtt{d}}}{j^{2\mathtt{m}-\mathtt{d}}} + \frac{j^{\mathtt{d}}}{j'^{2\mathtt{m}-\mathtt{d}}} \leq \frac{(j+N_n)^{\mathtt{d}}}{j^{2\mathtt{m}-\mathtt{d}}} + \frac{(j'+N_n)^{\mathtt{d}}}{j'^{2\mathtt{m}-\mathtt{d}}} \lesssim_{\mathtt{d}} N_n^{\mathtt{d}} . \tag{16.41}$$

Therefore, for some $C_1 > 0$, one has, for any $n \geq 0$,

$$C\varepsilon\gamma^{-2M-3} N_{n-1}^{-\mathtt{a}} N_n^{\tau} \big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big) j^{\mathtt{d}} j'^{\mathtt{d}} \leq C_1 \varepsilon\gamma^{-2M-3} N_{n-1}^{-\mathtt{a}} N_n^{\tau+\mathtt{d}} \leq 1$$

for ε small enough, by (15.10), (16.7) and because $\tau_3 > p(\tau + \mathtt{d})$ (that follows since $\tau_2 > \tau_1 + \mathtt{a}$ where τ_2 has been fixed in Theorem 15.4). In conclusion we have proved that $\mathcal{C}_\infty^\gamma \subseteq \Lambda_{n+1}^{2\gamma}(i_\infty)$ (for the second Melnikov conditions with the $+$ sign in (15.29) we apply the same argument). Similarly we prove that $\mathcal{C}_\infty^\gamma \subseteq \Lambda_n^{2\gamma, I}(i_\infty)$ for all $n \in \mathbb{N}$. □

Lemmata 16.5, 16.6 imply the following inclusion.

Corollary 16.7. *$\mathcal{C}_\infty^\gamma \subseteq \bigcap_{n \geq 0} \mathcal{G}_n$, where $\mathcal{G}_n$ is defined in* (16.10).

A Whitney differentiable functions

In this Appendix we recall the notion of Whitney differentiable functions and the Whitney extension theorem, following the version of Stein [59]. Then we prove the lemmata stated in Section 2.1. The following definition is the adaptation of the one in Section 2.3, Chapter VI of [59] to Banach-valued functions.

Definition A.1. (Whitney differentiable functions) *Let F be a closed subset of $\mathbb{R}^n$, $n \geq 1$. Let Y be a Banach space. Let $k \geq 0$ be an integer, and $k < \rho \leq k+1$. We say that a function $f : F \to Y$ belongs to* $\mathrm{Lip}(\rho, F, Y)$ *if there exist functions*

$$f^{(j)} : F \to Y, \quad j \in \mathbb{N}^n, \ 0 \leq |j| \leq k,$$

with $f^{(0)} = f$, and a constant $M > 0$ such that if $R_j(x,y)$ is defined by

$$f^{(j)}(x) = \sum_{\ell \in \mathbb{N}^n : |j+\ell| \leq k} \frac{1}{\ell!} f^{(j+\ell)}(y) \, (x-y)^\ell + R_j(x,y), \quad x, y \in F, \tag{A.1}$$

then

$$\|f^{(j)}(x)\|_Y \leq M, \quad \|R_j(x,y)\|_Y \leq M |x-y|^{\rho - |j|}, \quad \forall x, y \in F, \ |j| \leq k . \tag{A.2}$$

An element of $\mathrm{Lip}(\rho, F, Y)$ *is in fact the collection $\{f^{(j)} : |j| \leq k\}$. The norm of $f \in \mathrm{Lip}(\rho, F, Y)$ is defined as the smallest M for which the inequality* (A.2) *holds, namely*

$$\|f\|_{\mathrm{Lip}(\rho, F, Y)} := \inf\{M > 0 : \text{(A.2) holds}\} . \tag{A.3}$$

If $F = \mathbb{R}^n$ by $\mathrm{Lip}(\rho, \mathbb{R}^n, Y)$ *we shall mean the linear space of the functions $f = f^{(0)}$ for which there exist $f^{(j)} = \partial_x^j f$, $|j| \leq k$, satisfying* (A.2).

Notice that, if $F = \mathbb{R}^n$, the $f^{(j)}$, $|j| \geq 1$, are uniquely determined by $f^{(0)}$ (which is not the case for a general F with for example isolated points).

In the case $F = \mathbb{R}^n$, $\rho = k + 1$ and Y is a Hilbert space, the space $\mathrm{Lip}(k+1, \mathbb{R}^n, Y)$ is isomorphic to the Sobolev space $W^{k+1,\infty}(\mathbb{R}^n, Y)$, with equivalent norms

$$C_1 \|f\|_{W^{k+1,\infty}(\mathbb{R}^n,Y)} \leq \|f\|_{\mathrm{Lip}(k+1,\mathbb{R}^n,Y)} \leq C_2 \|f\|_{W^{k+1,\infty}(\mathbb{R}^n,Y)} \tag{A.4}$$

where C_1, C_2 depend only on k, n. For $Y = \mathbb{C}$ this isomorphism is classical, see e.g. [59], and it is based on the Rademacher theorem concerning the a.e. differentiability of Lipschitz functions, and the fundamental theorem of calculus for the Lebesgue integral. Such a property may fail for a Banach valued function, but it holds for a Hilbert space, see Chapter 5 of [12] (more in general it holds if Y is reflexive or it satisfies the Radon-Nykodim property).

The following key result provides an extension of a Whitney differentiable function f defined on a closed subset F of $\mathbb{R}^n$ to the whole domain $\mathbb{R}^n$, with equivalent norm.

Theorem A.2. (Whitney extension Theorem) *Let F be a closed subset of $\mathbb{R}^n$, $n \geq 1$, Y a Banach space, $k \geq 0$ an integer, and $k < \rho \leq k+1$. There exists a linear continuous extension operator $\mathcal{E}_k : \mathrm{Lip}(\rho, F, Y) \to \mathrm{Lip}(\rho, \mathbb{R}^n, Y)$ which gives an extension $\mathcal{E}_k f \in \mathrm{Lip}(\rho, \mathbb{R}^n, Y)$ to any $f \in \mathrm{Lip}(\rho, F, Y)$. The norm of $\mathcal{E}_k$ has a bound independent of F,*

$$\|\mathcal{E}_k f\|_{\mathrm{Lip}(\rho,\mathbb{R}^n,Y)} \leq C \|f\|_{\mathrm{Lip}(\rho,F,Y)}, \quad \forall f \in \mathrm{Lip}(\rho, F, Y), \tag{A.5}$$

where C depends only on n, k (and not on F, Y).

Proof. This is Theorem 4 in Section 2.3, Chapter VI of [59]. The proof in [59] is written for real-valued functions $f : F \to \mathbb{R}$, but it also holds for functions $f : F \to Y$ for any (real or complex) Banach space Y, *with no change*. The extension operator $\mathcal{E}_k$ is defined in formula (18) in Section 2.3, Chapter VI of [59], and it is linear by construction. □

Clearly, since $\mathcal{E}_k f$ is an extension of f, one has

$$\|f\|_{\mathrm{Lip}(\rho,F,Y)} \leq \|\mathcal{E}_k f\|_{\mathrm{Lip}(\rho,\mathbb{R}^n,Y)} \leq C \|f\|_{\mathrm{Lip}(\rho,F,Y)}. \tag{A.6}$$

In order to extend a function defined on a closed set $F \subset \mathbb{R}^n$ with values in scales of Banach spaces (like $H^s(\mathbb{T}^{\nu+1})$), we observe that the extension provided by Theorem A.2 does not depend on the index of the space (namely s).

Lemma A.3. *Let F be a closed subset of $\mathbb{R}^n$, $n \geq 1$, let $k \geq 0$ be an integer, and $k < \rho \leq k+1$. Let $Y \subseteq Z$ be two Banach spaces. Then $\mathrm{Lip}(\rho, F, Y) \subseteq \mathrm{Lip}(\rho, F, Z)$. The two extension operators $\mathcal{E}_k^{(Z)} : \mathrm{Lip}(\rho, F, Z) \to \mathrm{Lip}(\rho, \mathbb{R}^n, Z)$ and $\mathcal{E}_k^{(Y)} : \mathrm{Lip}(\rho, F, Y) \to \mathrm{Lip}(\rho, \mathbb{R}^n, Y)$ provided by Theorem A.2 satisfy*

$$\mathcal{E}_k^{(Z)} f = \mathcal{E}_k^{(Y)} f \quad \forall f \in \mathrm{Lip}(\rho, F, Y).$$

As a consequence, we simply denote $\mathcal{E}_k$ the extension operator.

Proof. The lemma follows directly by the construction of the extension operator $\mathcal{E}_k$ in formula (18) in Section 2.3, Chapter VI of [59]. The explicit construction relies on a nontrivial decomposition in cubes of the domain $\mathbb{R}^n$ only. □

Thanks to the equivalence (A.6), Lemma A.3, and (A.4) which holds for functions valued in H^s, classical interpolation and tame estimates for products, projections, and composition of Sobolev functions can be easily extended to Whitney differentiable functions.

The difference between the Whitney-Sobolev norm introduced in Definition 2.1 and the norm in Definition A.1 (for $\rho = k+1$, $n = \nu + 1$, and target space $Y = H^s(\mathbb{T}^{\nu+1}, \mathbb{C})$) is the weight $\gamma \in (0, 1]$. Observe that the introduction of this weight simply amounts to the following rescaling $\mathcal{R}_\gamma$: given $u = (u^{(j)})_{|j| \leq k}$, we define $\mathcal{R}_\gamma u = U = (U^{(j)})_{|j| \leq k}$ as

$$\lambda = \gamma \mu, \qquad \gamma^{|j|} u^{(j)}(\lambda) = \gamma^{|j|} u^{(j)}(\gamma \mu) =: U^{(j)}(\mu) = U^{(j)}(\gamma^{-1} \lambda), \qquad U := \mathcal{R}_\gamma u. \tag{A.7}$$

Thus $u \in \mathrm{Lip}(k+1, F, s, \gamma)$ if and only if $U \in \mathrm{Lip}(k+1, \gamma^{-1}F, s, 1)$, with

$$\|u\|_{s,F}^{k+1,\gamma} = \|U\|_{s,\gamma^{-1}F}^{k+1,1} . \tag{A.8}$$

Under the rescaling $\mathcal{R}_\gamma$, (A.4) gives the equivalence of the two norms

$$\|f\|_{W^{k+1,\infty,\gamma}(\mathbb{R}^{\nu+1},H^s)} := \sum_{|\alpha|\le k+1} \gamma^{|\alpha|} \|\partial_\lambda^\alpha f\|_{L^\infty(\mathbb{R}^{\nu+1},H^s)} \sim_{\nu,k} \|f\|_{s,\mathbb{R}^{\nu+1}}^{k+1,\gamma} . \tag{A.9}$$

Moreover, given $u \in \mathrm{Lip}(k+1, F, s, \gamma)$, its extension

$$\tilde{u} := \mathcal{R}_\gamma^{-1} \mathcal{E}_k \mathcal{R}_\gamma u \in \mathrm{Lip}(k+1, \mathbb{R}^{\nu+1}, s, \gamma) \quad \text{satisfies} \quad \|u\|_{s,F}^{k+1,\gamma} \sim_{\nu,k} \|\tilde{u}\|_{s,\mathbb{R}^{\nu+1}}^{k+1,\gamma} . \tag{A.10}$$

Proof of Lemma 2.2. Inequalities (2.6)-(2.7) follow by

$$(\Pi_N u)^{(j)}(\lambda) = \Pi_N[u^{(j)}(\lambda)], \quad R_j^{(\Pi_N u)}(\lambda, \lambda_0) = \Pi_N[R_j^{(u)}(\lambda, \lambda_0)],$$

for all $0 \le |j| \le k$, $\lambda, \lambda_0 \in F$, and the usual smoothing estimates $\|\Pi_N f\|_s \le N^\alpha \|f\|_{s-\alpha}$ and $\|\Pi_N^\perp f\|_s \le N^{-\alpha}\|f\|_{s+\alpha}$ for Sobolev functions. ☐

Proof of Lemma 2.3. Inequality (2.8) follows from the classical interpolation inequality $\|u\|_s \le \|u\|_{s_0}^\theta \|u\|_{s_1}^{1-\theta}$, $s = \theta s_0 + (1-\theta)s_1$ for Sobolev functions, and from the Definition 2.1 of Whitney-Sobolev norms, since

$$\gamma^{|j|}\|u^{(j)}(\lambda)\|_s \le (\gamma^{|j|}\|u^{(j)}(\lambda)\|_{s_0})^\theta (\gamma^{|j|}\|u^{(j)}(\lambda)\|_{s_1})^{1-\theta} \le (\|u\|_{s_0,F}^{k+1,\gamma})^\theta (\|u\|_{s_1,F}^{k+1,\gamma})^{1-\theta},$$

$$\gamma^{k+1}\|R_j(\lambda,\lambda_0)\|_s \le (\gamma^{k+1}\|R_j(\lambda,\lambda_0)\|_{s_0})^\theta (\gamma^{k+1}\|R_j(\lambda,\lambda_0)\|_{s_1})^{1-\theta} \le (\|u\|_{s_0,F}^{k+1,\gamma})^\theta (\|u\|_{s_1,F}^{k+1,\gamma})^{1-\theta} |\lambda-\lambda_0|^{k+1-|j|}.$$

Inequality (2.9) follows from (2.8) by using the asymmetric Young inequality (like in the proof of Lemma 2.2 in [21]). ☐

Proof of Lemma 2.4. By (A.9)-(A.10), the lemma follows from the corresponding inequalities for functions in $W^{k+1,\infty,\gamma}(\mathbb{R}^{\nu+1}, H^s)$, which are proved, for instance, in [21] (formula (2.72), Lemma 2.30). ☐

For any $\rho > 0$, we define the $\mathcal{C}^\infty$ function $h_\rho : \mathbb{R} \to \mathbb{R}$,

$$h_\rho(y) := \frac{\chi_\rho(y)}{y} = \frac{\chi(y\rho^{-1})}{y}, \quad \forall y \in \mathbb{R} \setminus \{0\}, \quad h_\rho(0) := 0 , \tag{A.11}$$

where χ is the cut-off function introduced in (2.16), and $\chi_\rho(y) := \chi(y/\rho)$. Notice that the function h_ρ is of class $\mathcal{C}^\infty$ because $h_\rho(y) = 0$ for $|y| \le \rho/3$. Moreover by the properties of χ in (2.16) we have

$$h_\rho(y) = \frac{1}{y}, \ \forall |y| \ge \frac{2\rho}{3}, \qquad |h_\rho(y)| \le \frac{3}{\rho}, \ \forall y \in \mathbb{R} . \tag{A.12}$$

To prove Lemma 2.5, we use the following preliminary lemma.

Lemma A.4. *Let $f : \mathbb{R}^{\nu+1} \to \mathbb{R}$ and $\rho > 0$. Then the function*

$$g(\lambda) := h_\rho(f(\lambda)), \quad \forall \lambda \in \mathbb{R}^{\nu+1} , \tag{A.13}$$

where h_ρ is defined in (A.11)*, coincides with $1/f(\lambda)$ on the set $F := \{\lambda \in \mathbb{R}^{\nu+1} : |f(\lambda)| \ge \rho\}$.*

If the function f is in $W^{k+1,\infty}(\mathbb{R}^{\nu+1}, \mathbb{R})$, with estimates

$$\gamma^{|\alpha|}|\partial_\lambda^\alpha f(\lambda)| \le M , \quad \forall \alpha \in \mathbb{N}^{\nu+1} , \quad 1 \le |\alpha| \le k+1 , \tag{A.14}$$

for some $M \ge \rho$, then the function g is in $W^{k+1,\infty}(\mathbb{R}^{\nu+1}, \mathbb{R})$ and

$$\gamma^{|\alpha|}|\partial_\lambda^\alpha g(\lambda)| \le C_k \frac{M^{k+1}}{\rho^{k+2}} , \quad \forall \alpha \in \mathbb{N}^{\nu+1}, \quad 0 \le |\alpha| \le k+1. \tag{A.15}$$

Proof. By (A.12), $g(\lambda) = 1/f(\lambda)$ for all $\lambda \in F$. The derivatives of $h_\rho(y)$ are

$$h_\rho^{(m)}(y) = \partial_y^m \big(\frac{\chi_\rho(y)}{y} \big) = \sum_{m_1+m_2=m} C_{m_1,m_2} \rho^{-m_1} (\partial_y^{m_1} \chi)(\rho^{-1} y) y^{-m_2-1} \,, \quad m \geq 0,$$

that we may bound — we have just to consider $|y| \geq \rho/3$ (otherwise $h_\rho(y) = 0$) — as

$$|h_\rho^{(m)}(y)| \lesssim_m \sum_{m_1+m_2=m} \rho^{-m_1} |y|^{-m_2-1} \lesssim_m \rho^{-m-1} \,, \quad \forall y \in \mathbb{R} \,. \tag{A.16}$$

Using the Faà di Bruno formula, for $|\alpha| \geq 1$ we compute the derivatives of the composite function

$$\partial_\lambda^\alpha g(\lambda) = \partial_\lambda^\alpha h_\rho(f(\lambda)) = \sum_{1 \leq q \leq |\alpha|} \sum_{\substack{\sigma_1 + \ldots + \sigma_q = \alpha, \\ \sigma_j \neq 0, j=1,\ldots,q}} C_{q,\sigma_1,\ldots,\sigma_q} h_\rho^{(q)}(f(\lambda)) \partial_\lambda^{\sigma_1} f(\lambda) \ldots \partial_\lambda^{\sigma_q} f(\lambda) \,,$$

and, using (A.16), we get, for all $|\alpha| \geq 1$,

$$\begin{aligned} |\partial_\lambda^\alpha g(\lambda)| &\leq C_\alpha \sum_{1 \leq q \leq |\alpha|} \sum_{\substack{\sigma_1 + \ldots + \sigma_q = \alpha, \\ \sigma_j \neq 0, j=1,\ldots,q}} \rho^{-q-1} |\partial_\lambda^{\sigma_1} f(\lambda)| \ldots |\partial_\lambda^{\sigma_q} f(\lambda)| \\ &\overset{(A.14)}{\leq} C_\alpha \sum_{1 \leq q \leq |\alpha|} \rho^{-q-1} \gamma^{-|\alpha|} M^q \leq C_\alpha \rho^{-|\alpha|-1} \gamma^{-|\alpha|} M^{|\alpha|} \,. \end{aligned}$$

Formula (A.15) for $\alpha = 0$ holds by (A.12). ☐

Proof of Lemma 2.5. The function $(\omega \cdot \partial_\varphi)_{ext}^{-1} u$ defined in (2.15) is

$$\big((\omega \cdot \partial_\varphi)_{ext}^{-1} u \big)(\lambda, \varphi, x) = -\mathrm{i} \sum_{(\ell,j) \in \mathbb{Z}^{\nu+1}} g_\ell(\lambda) u_{\ell,j}(\lambda) \, e^{\mathrm{i}(\ell \cdot \varphi + jx)} \,,$$

where $g_\ell(\lambda) = h_\rho(\omega \cdot \ell)$ in (A.13) with $\rho = \gamma \langle \ell \rangle^{-\tau}$ and $f(\lambda) = \omega \cdot \ell$. The function $f(\lambda)$ satisfies (A.14) with $M = \gamma |\ell|$. Hence $g_\ell(\lambda)$ satisfies (A.15), namely

$$\gamma^{|\alpha|} |\partial_\lambda^\alpha g_\ell(\lambda)| \leq C_k \gamma^{-1} \langle \ell \rangle^\mu \quad \forall \alpha \in \mathbb{N}^{\nu+1}, \ 0 \leq |\alpha| \leq k+1, \tag{A.17}$$

where $\mu = k + 1 + (k+2)\tau$ is defined in (2.18). One has

$$\partial_\lambda^\alpha (g_\ell(\lambda) u_{\ell,j}(\lambda)) = \sum_{\alpha_1 + \alpha_2 = \alpha} C_{\alpha_1,\alpha_2} (\partial_\lambda^{\alpha_1} g_\ell)(\lambda) (\partial_\lambda^{\alpha_2} u_{\ell,j})(\lambda),$$

whence, by (A.17), we deduce

$$\gamma^{|\alpha|} \| \partial_\lambda^\alpha ((\omega \cdot \partial_\varphi)_{ext}^{-1} u)(\lambda) \|_s \leq C_k \gamma^{-1} \| u \|_{s+\mu, \mathbb{R}^{\nu+1}}^{k+1,\gamma}$$

and therefore (2.17). The proof is concluded by observing that the restriction of $(\omega \cdot \partial_\varphi)_{ext}^{-1} u$ to F gives $(\omega \cdot \partial_\varphi)^{-1} u$ as defined in (2.14), and (2.18) follows by (A.10). ☐

Proof of Lemma 2.6. Given $u \in \mathrm{Lip}(k+1, F, s, \gamma)$, we consider its extension $\tilde u \in \mathrm{Lip}(k+1, \mathbb{R}^{\nu+1}, s, \gamma)$ provided by (A.10). Then we observe that the composition $\mathtt{f}(\tilde u)$ is an extension of $\mathtt{f}(u)$, and therefore one has the inequality $\| \mathtt{f}(u) \|_{s,F}^{k+1,\gamma} \leq \| \mathtt{f}(\tilde u) \|_{s,\mathbb{R}^{\nu+1}}^{k+1,\gamma} \sim \| \mathtt{f}(\tilde u) \|_{W^{k+1,\infty,\gamma}(\mathbb{R}^{\nu+1}, H^s)}$ by (A.9). Then (2.19) follows by the Moser composition estimates for $\| \ \|_{s,\mathbb{R}^{\nu+1}}^{k+1,\gamma}$ (see for instance Lemma 2.31 in [21]), together with the equivalence of the norms in (A.9)-(A.10). ☐

B A Nash-Moser-Hörmander implicit function theorem

In this section we state the Nash-Moser-Hörmander theorem of [10], which we apply in Section 8 as a black box to prove Theorem 8.3.

Let $(E_a)_{a\geq 0}$ be a decreasing family of Banach spaces with continuous injections $E_b \hookrightarrow E_a$,

$$\|u\|_{E_a} \leq \|u\|_{E_b} \quad \text{for } a \leq b. \tag{B.1}$$

Set $E_\infty = \cap_{a\geq 0} E_a$ with the weakest topology making the injections $E_\infty \hookrightarrow E_a$ continuous. Assume that there exist linear smoothing operators $S_j : E_0 \to E_\infty$ for $j = 0, 1, \ldots$, satisfying the following inequalities, with constants C bounded when a and b are bounded, and independent of j,

$$\|S_j u\|_{E_a} \leq C\|u\|_{E_a} \qquad \text{for all } a; \tag{B.2}$$

$$\|S_j u\|_{E_b} \leq C2^{j(b-a)}\|S_j u\|_{E_a} \qquad \text{if } a < b; \tag{B.3}$$

$$\|u - S_j u\|_{E_b} \leq C2^{-j(a-b)}\|u - S_j u\|_{E_a} \qquad \text{if } a > b; \tag{B.4}$$

$$\|(S_{j+1} - S_j)u\|_{E_b} \leq C2^{j(b-a)}\|(S_{j+1} - S_j)u\|_{E_a} \quad \text{for all } a, b. \tag{B.5}$$

Set

$$R_0 u := S_1 u, \qquad R_j u := (S_{j+1} - S_j)u, \quad j \geq 1. \tag{B.6}$$

Thus

$$\|R_j u\|_{E_b} \leq C2^{j(b-a)}\|R_j u\|_{E_a} \quad \text{for all } a, b. \tag{B.7}$$

Bound (B.7) for $j \geq 1$ is (B.5), while, for $j = 0$, it follows from (B.1) and (B.3). We also assume that

$$\|u\|^2_{E_a} \leq C \sum_{j=0}^{\infty} \|R_j u\|^2_{E_a} \quad \forall a \geq 0, \tag{B.8}$$

with C bounded for a bounded (a sort of "orthogonality property" of the smoothing operators).

Suppose that we have another family F_a of decreasing Banach spaces with smoothing operators having the same properties as above. We use the same notation also for the smoothing operators.

Theorem B.1 ([10]). **(Existence)** *Let $a_1, a_2, \alpha, \beta, a_0, \mu$ be real numbers with*

$$0 \leq a_0 \leq \mu \leq a_1, \qquad a_1 + \frac{\beta}{2} < \alpha < a_1 + \beta, \qquad 2\alpha < a_1 + a_2. \tag{B.9}$$

Let U be a convex neighborhood of 0 in E_μ. Let Φ be a map from U to F_0 such that $\Phi : U \cap E_{a+\mu} \to F_a$ is of class C^2 for all $a \in [0, a_2 - \mu]$, with

$$\begin{aligned}\|\Phi''(u)[v, w]\|_{F_a} \leq\ & M_1(a)\big(\|v\|_{E_{a+\mu}}\|w\|_{E_{a_0}} + \|v\|_{E_{a_0}}\|w\|_{E_{a+\mu}}\big) \\ & + \{M_2(a)\|u\|_{E_{a+\mu}} + M_3(a)\}\|v\|_{E_{a_0}}\|w\|_{E_{a_0}}\end{aligned} \tag{B.10}$$

for all $u \in U \cap E_{a+\mu}$, $v, w \in E_{a+\mu}$, where $M_i : [0, a_2 - \mu] \to \mathbb{R}$, $i = 1, 2, 3$, are positive, increasing functions. Assume that $\Phi'(v)$, for $v \in E_\infty \cap U$ belonging to some ball $\|v\|_{E_{a_1}} \leq \delta_1$, has a right inverse $\Psi(v)$ mapping F_∞ to E_{a_2}, and that

$$\|\Psi(v)g\|_{E_a} \leq L_1(a)\|g\|_{F_{a+\beta-\alpha}} + \{L_2(a)\|v\|_{E_{a+\beta}} + L_3(a)\}\|g\|_{F_0} \quad \forall a \in [a_1, a_2], \tag{B.11}$$

where $L_i : [a_1, a_2] \to \mathbb{R}$, $i = 1, 2, 3$, are positive, increasing functions.

Then for all $A > 0$ there exists $\delta > 0$ such that, for every $g \in F_\beta$ satisfying

$$\sum_{j=0}^{\infty} \|R_j g\|^2_{F_\beta} \leq A^2\|g\|^2_{F_\beta}, \quad \|g\|_{F_\beta} \leq \delta, \tag{B.12}$$

there exists $u \in E_\alpha$ solving $\Phi(u) = \Phi(0) + g$. The solution u satisfies

$$\|u\|_{E_\alpha} \le C L_{123}(a_2)(1+A)\|g\|_{F_\beta}, \tag{B.13}$$

where $L_{123} = L_1 + L_2 + L_3$ and C is a constant depending on a_1, a_2, α, β. The constant δ is

$$\delta = 1/B, \quad B = C' L_{123}(a_2) \max\big\{1/\delta_1, 1+A, (1+A)L_{123}(a_2)M_{123}(a_2-\mu)\big\} \tag{B.14}$$

where $M_{123} = M_1 + M_2 + M_3$ and C' is a constant depending on a_1, a_2, α, β.

(Higher regularity) *Moreover, let $c > 0$ and assume that* (B.10) *holds for all $a \in [0, a_2 + c - \mu]$, $\Psi(v)$ maps F_∞ to E_{a_2+c}, and* (B.11) *holds for all $a \in [a_1, a_2 + c]$. If g satisfies* (B.12) *and, in addition, $g \in F_{\beta+c}$ with*

$$\sum_{j=0}^{\infty} \|R_j g\|^2_{F_{\beta+c}} \le A_c^2 \|g\|^2_{F_{\beta+c}} \tag{B.15}$$

for some A_c, then the solution u belongs to $E_{\alpha+c}$, with

$$\|u\|_{E_{\alpha+c}} \le C_c\big\{\mathcal{G}_1(1+A)\|g\|_{F_\beta} + \mathcal{G}_2(1+A_c)\|g\|_{F_{\beta+c}}\big\} \tag{B.16}$$

where

$$\mathcal{G}_1 := \tilde L_3 + \tilde L_{12}(\tilde L_3 \tilde M_{12} + L_{123}(a_2)\tilde M_3)(1+z^N), \quad \mathcal{G}_2 := \tilde L_{12}(1+z^N), \tag{B.17}$$

$$z := L_{123}(a_1)M_{123}(0) + \tilde L_{12}\tilde M_{12}, \tag{B.18}$$

$\tilde L_{12} := \tilde L_1 + \tilde L_2$, $\tilde L_i := L_i(a_2+c)$, $i = 1,2,3$; $\tilde M_{12} := \tilde M_1 + \tilde M_2$, $\tilde M_i := M_i(a_2+c-\mu)$, $i = 1,2,3$; N is a positive integer depending on c, a_1, α, β; and C_c depends on $a_1, a_2, \alpha, \beta, c$.

This theorem is proved in [10] using an iterative scheme similar to [33]. The main advantage with respect to the Nash-Moser implicit function theorems as presented in [62, 17] is the optimal regularity of the solution u in terms of the datum g (see (B.13), (B.16)). Theorem B.1 has the advantage of making explicit all the constants (unlike [33]), which is necessary to deduce the quantitative Theorem 8.3.

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 Cauchy problem for gravity water waves.* Invent. Math., 198, 71–163, 2014.

[3] Alazard T., Burq N., Zuily C., *Cauchy theory for the gravity water waves system with non-localized initial data.* Ann. Inst. H. Poincaré Anal. Non Linéaire, 33, 337-395, 2016.

[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] Alazard T., Métivier G., *Paralinearization of the Dirichlet to Neumann operator, and regularity of the three dimensional water waves*, Comm. Partial Differential Equations 34 (2009), no. 10-12, 1632-1704.

[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, AN 33, 1589-1638, 2016.

[9] Baldi P., Berti M., Montalto R., *KAM for autonomous quasi-linear perturbations of mKdV.* Bollettino Unione Matematica Italiana, 9, 143-188, 2016.

[10] Baldi P., Haus E., *A Nash-Moser-Hörmander implicit function theorem with applications to control and Cauchy problems for PDEs.* Preprint 2016 (arXiv:1609.00213).

[11] Bambusi D., Berti M., Magistrelli E., *Degenerate KAM theory for partial differential equations*, Journal Diff. Equations, 250, 8, 3379-3397, 2011.

[12] Benyamini Y., Lindenstrauss J., *Geometric nonlinear functional analysis. Vol. 1.* American Mathematical Society Colloquium Publications, 48. American Mathematical Society, Providence, RI, 2000.

[13] 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.

[14] Berti M., Biasco L., Procesi M., *KAM for Reversible Derivative Wave Equations.* Arch. Ration. Mech. Anal., 212(3):905–955, 2014.

[15] Berti M., Bolle Ph., *Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential*, Eur. Jour. Math. 15 (2013), 229-286.
[16] Berti M., Bolle Ph., *A Nash-Moser approach to KAM theory*, Fields Institute Communications, special volume "Hamiltonian PDEs and Applications", 255-284, 2015.
[17] Berti M., Bolle Ph., Procesi M., *An abstract Nash-Moser theorem with parameters and applications to PDEs*, Ann. I. H. Poincaré, 27, 377-399, 2010.
[18] Berti M., Corsi L., Procesi M. *An abstract Nash-Moser Theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds*, Comm. Math. Phys. 334, no. 3, 1413-1454, 2015.
[19] Berti M., Delort J-M., *Almost global existence of solutions for capillarity-gravity water waves equations with periodic spatial boundary conditions*, preprint arXiv:1702.04674.
[20] Berti M., Montalto R., *Quasi-periodic water waves*, J. Fixed Point Th. Appl. 19, no. 1, 129-156, 2017.
[21] Berti M., Montalto R., *KAM for gravity capillary water waves*, Memoires of AMS, Memo 891, to appear, preprint arXiv:1602.02411.
[22] Bourgain J., *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton, 2005.
[23] Corsi L., Feola R., Procesi M., *Finite dimensional invariant KAM tori for tame vector fields*, preprint, arXiv:1611.01641.
[24] Craig W., Nicholls D., *Travelling two and three dimensional capillary gravity water waves.* SIAM J. Math. Anal., 32(2):323–359 (electronic), 2000.
[25] Craig W., Nicholls D., *Traveling gravity water waves in two and three dimensions.* Eur. J. Mech. B Fluids, 21(6):615–641, 2002.
[26] Craig W., Sulem C., *Numerical simulation of gravity waves.* J. Comput. Phys., 108(1):73–83, 1993.
[27] Craig W., Worfolk P., *An integrable normal form for water waves in infinite depth.* Phys. D, 84 (1995), no. 3-4, 513-531.
[28] Dyachenko A.I., Lvov Y.V., Zakharov V.E., *Five-wave interaction on the surface of deep fluid*, Physica D 87 (1995) 233-261.
[29] Féjoz J., *Démonstration du théorème d'Arnold sur la stabilité du système planétaire (d'après Herman)*, Ergodic Theory Dynam. Systems 24 (5) (2004) 1521-1582.
[30] Feola R., Procesi M., *Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations.* J. Diff. Eq., 259, no. 7, 3389-3447, 2015.
[31] Germain P., Masmoudi N., Shatah J., *Global solutions for the gravity water waves equation in dimension 3*, Ann. of Math. (2), 175, 691–754, 2012.
[32] Eliasson L. H., Kuksin S., *KAM for nonlinear Schrödinger equation*, Annals of Math, (2) 172, no. 1, 371-435, 2010.
[33] Hörmander L., *The boundary problems of physical geodesy.* Arch. Rat. Mech. Anal. 62, no. 1, 1-52, 1976.
[34] Hörmander L., *The analysis of linear partial differential operators III.* Springer-Verlag, Berlin, 1990.
[35] Ifrim M., Tataru D., *Two dimensional gravity water waves with constant vorticity: I. Cubic lifespan*, arXiv:1510.07732.
[36] Ionescu A., Pusateri F., *Global solutions for the gravity water waves system in 2d*, Invent. Math., 199, 3, 653-804, 2015.
[37] Iooss G., Plotnikov P., *Existence of multimodal standing gravity waves.* J. Math. Fluid Mech., 7(suppl. 3):S349–S364, 2005.
[38] Iooss G., Plotnikov P., *Multimodal standing gravity waves: a completely resonant system.* J. Math. Fluid Mech., 7(suppl. 1):S110–S126, 2005.
[39] 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.
[40] Iooss G., Plotnikov P., *Asymmetrical tridimensional traveling gravity waves.* Arch. Rat. Mech. Anal., 200(3):789–880, 2011.
[41] Iooss G., Plotnikov P., Toland J., *Standing waves on an infinitely deep perfect fluid under gravity.* Arch. Ration. Mech. Anal., 177(3):367–478, 2005.
[42] Kappeler T., Pöschel J., *KdV & KAM*, Springer-Verlag, 2003.
[43] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional Anal. i Prilozhen. 2, 22-37, 95, 1987.
[44] Kuksin S., *Analysis of Hamiltonian PDEs*, volume 19 of Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, 2000.
[45] Lannes D., *Well-posedness of the water-waves equations*, J. Amer. Math. Soc., 3, 605–654, 18, 2005.
[46] Lannes D., *The water waves problem: mathematical analysis and asymptotics.* Mathematical Surveys and Monographs, 188, 2013.
[47] Levi-Civita T., *Détermination rigoureuse des ondes permanentes d' ampleur finie.* Math. Ann., 93 , pp. 264-314, 1925.
[48] Liu J., Yuan X., *A KAM theorem for Hamiltonian partial differential equations with unbounded perturbations.* Comm. Math. Phys., 307(3), 629–673, 2011.
[49] Métivier G., *Para-differential Calculus and Applications to the Cauchy Problem for Nonlinear Systems.* Pubblicazioni Scuola Normale Pisa, 5, 2008.
[50] Montalto R., *Quasi-periodic solutions of forced Kirchhoff equation.* NoDEA, Nonlinear Differ. Equ. Appl., 24(1), 9, 2017.
[51] Moser J., *Convergent series expansions for quasi-periodic motions*, Math.Ann. 169, 136-176, 1967.
[52] Plotnikov P., Toland J., *Nash-Moser theory for standing water waves.* Arch. Ration. Mech. Anal., 159(1):1–83, 2001.
[53] Pöschel J., *A KAM-Theorem for some nonlinear PDEs*, Ann. Sc. Norm. Pisa, 23, (1996) 119-148.
[54] Procesi C., Procesi M., *A KAM algorithm for the completely resonant nonlinear Schrödinger equation*, Advances in Mathematics, volume 272, 399-470, 2015.
[55] Pjartli A.S., *Diophantine approximations of submanifolds of a Euclidean space*, Funktsional. Anal. i Prilozhen. 3 (4) (1969) 59-62.
[56] Robbin J., Salamon D., *The exponential Vandermonde matrix*, Linear Algebra Appl. 317 (2000) 225.

[57] Rüssmann H., *Invariant tori in non-degenerate nearly integrable Hamiltonian systems*, Regul. Chaotic Dyn. 6 (2) (2001) 119-204.

[58] Saranen J., Vainikko G., *Periodic Integral and Pseudodifferential Equations with Numerical Approximation*, Springer Monographs in Mathematics, 2002.

[59] Stein E.M., *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, no. 30. Princeton University Press, Princeton, N.J. 1970.

[60] Wu S., *Global well-posedness of the 3-D full water wave problem*, Invent. Math., 1, 184, 125–220, 2011.

[61] Zakharov V. E., *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, J. Appl. Mech. Tech. Phys., 9, 190-194, 1968.

[62] Zehnder E., *Generalized implicit function theorems with applications to some small divisor problems. I-II.* Comm. Pure Appl. Math. 28 (1975), 91-140; and 29 (1976), 49-111.

[63] Zhang J., Gao M., Yuan X. *KAM tori for reversible partial differential equations.* Nonlinearity, 24(4):1189-1228, 2011.

Pietro Baldi, Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università di Napoli Federico II, Via Cintia, Monte S. Angelo, 80126, Napoli, Italy. *E-mail:* `pietro.baldi@unina.it`

Massimiliano Berti, SISSA, Via Bonomea 265, 34136, Trieste, Italy. *E-mail:* `berti@sissa.it`

Emanuele Haus, Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università di Napoli Federico II, Via Cintia, Monte S. Angelo, 80126, Napoli, Italy. *E-mail:* `emanuele.haus@unina.it`

Riccardo Montalto, University of Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland. *E-mail:* `riccardo.montalto@math.uzh.ch`