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 fully nonlinear nature of the gravity water waves equations – the highest order x-derivative appears in the nonlinear term but not in the linearization at the origin – 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 a Nash-Moser iterative scheme, 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 which lose derivatives both in time and space. Despite the fact that the depth parameter moves the linear frequencies by just exponentially small quantities, we are able to verify such non-resonance conditions for most values of the depth, extending degenerate KAM theory.

Keywords: Water waves, KAM for PDEs, quasi-periodic solutions, standing waves.
MSC 2010: 76B15, 37K55 (37C55, 35S05).

Contents

1 Introduction

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

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

where $g > 0$ is the acceleration of gravity. The unknowns of the problem are the free surface $y = \eta(t,x)$ 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 the fact 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 $(\eta(t,x),\psi(t,x))$ where $\psi(t,x)=\Phi(t,x,\eta(t,x))$ is, at each instant t, the trace at the free boundary of the velocity potential. Given 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 that $G(\eta,h)-D\tanh(hD)$ is in $OPS^{-\infty}$ when $\eta(x)\in\mathcal{C}^\infty$. This operator has been introduced in Craig-Sulem [26] and its properties are nowdays well-understood thanks to the works of Lannes [46]-[47], Alazard-Métivier [5], Alazard-Burq-Zuily [2], Alazard-Delort [4]. In Appendix A we provide a self-contained analysis of the Dirichlet-Neumann operator adapted to our purposes.

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, and $H_0^1(\mathbb{T})$ is the subspace of $H^1(\mathbb{T})$ of zero average functions. For simplicity of notation we denote the equivalence class $[\psi]$ by 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 $H_0^1(\mathbb{T})$, 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 H in (1.7) 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*.

In this paper we prove the first existence result of small amplitude time *quasi-periodic* standing waves solutions of the pure gravity water waves equations (1.4), for most values of the depth h, see Theorem 1.1.

The existence of standing water waves is a small divisor problem, which is particularly difficult because (1.4) is a fully nonlinear system of PDEs, the nonlinearity contains derivatives of order higher than those present in the linearized system at the origin, and the linear frequencies grow as $\sim j^{1/2}$. The existence of small amplitude time-periodic gravity standing wave solutions for bi-dimensional fluids has been first proved by Plotinkov and Toland [53] in finite depth and by Iooss, Plotnikov and Toland in [42] in infinite depth, see also [38], [39]. More recently, the existence of time periodic gravity-capillary standing wave solutions in infinite depth 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, in infinite depth, 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 [48]; 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 [25] for the gravity-capillary case (which is not a small divisor problem) and by Iooss and Plotinikov [40]-[41] in the pure gravity case (which is a small divisor problem).

From a physical point of view, it is natural to consider the depth h of the ocean as a fixed physical quantity and to introduce the space wavelength $2\pi\varsigma$ as an *internal* parameter. Rescaling time, space and amplitude of the solution $(\eta(t, x), \psi(t, x))$ of (1.4) as

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

we get that $(\tilde{\eta}(\tau, \tilde{x}), \tilde{\psi}(\tau, \tilde{x}))$ satisfies

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

Choosing the scaling parameters ς, μ, α such that $\frac{\varsigma^2}{\alpha\mu} = 1$, $\frac{g\alpha}{\varsigma\mu} = 1$ we obtain system (1.4) where the gravity constant g has been replaced by 1 and the depth parameter h by

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

Changing the parameter $\mathtt{h}$ can be interpreted as changing the space period $2\pi\varsigma$ of the solutions and not the depth h of the water, giving results for a *fixed* equation (1.4).

In the sequel 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 $\psi(t) \in \dot{H}^1(\mathbb{T}_x)$, actually belonging to more regular Sobolev spaces.

1.1 Main result

We look for small amplitude solutions of (1.14). Hence a fundamental rôle is played by 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 $\eta = 0$. 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) and (1.9), 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 equations (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 $u(\widetilde{\omega}t) = (\eta, \psi)(\widetilde{\omega}t)$ of (1.14), with frequency $\widetilde{\omega} \in \mathbb{R}^\nu$ (to be determined), close to 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)

$$\eta(t, x) = \sum_{j \in \mathbb{S}^+} a_j \cos\big(\omega_j(\mathtt{h})t\big)\cos(jx), \quad \psi(t, x) = -\sum_{j \in \mathbb{S}^+} \frac{a_j}{\omega_j(\mathtt{h})} \sin\big(\omega_j(\mathtt{h})t\big)\cos(jx), \quad a_j > 0\,, \tag{1.22}$$

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 vector $\vec{a} := (a_j)_{j\in\mathbb{S}^+}$, with $a_j > 0$ for all $j \in \mathbb{S}^+$ and $|\vec{a}| \leq \varepsilon_0$, there exists a Cantor-like set $\mathcal{G} \subset [\mathtt{h}_1, \mathtt{h}_2]$ with asymptotically full measure as $\vec{a} \to 0$, i.e.*

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

such that, for any $\mathtt{h} \in \mathcal{G}$, the gravity water waves system (1.14) has a time quasi-periodic solution $u(\widetilde{\omega}t, x) = (\eta(\widetilde{\omega}t, x), \psi(\widetilde{\omega}t, x))$, with Sobolev regularity $(\eta, \psi) \in H^{\bar{s}}(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R}^2)$, with a Diophantine frequency vector $\widetilde{\omega} := \widetilde{\omega}(\mathtt{h}, \vec{a}) := (\widetilde{\omega}_j)_{j\in\mathbb{S}^+} \in \mathbb{R}^\nu$, of the form

$$\begin{aligned} \eta(\widetilde{\omega}t, x) &= \sum_{j\in\mathbb{S}^+} a_j \cos(\widetilde{\omega}_j t)\cos(jx) + r_1(\widetilde{\omega}t, x), \\ \psi(\widetilde{\omega}t, x) &= -\sum_{j\in\mathbb{S}^+} \frac{a_j}{\omega_j(\mathtt{h})} \sin(\widetilde{\omega}_j t)\cos(jx) + r_2(\widetilde{\omega}t, x) \end{aligned} \tag{1.23}$$

with $\widetilde{\omega}(\mathtt{h}, \vec{a}) \to \vec{\omega}(\mathtt{h}) := (\omega_j(\mathtt{h}))_{j\in\mathbb{S}^+}$ as $\vec{a} \to 0$, and the functions $r_1(\varphi, x), r_2(\varphi, x)$ are $o(|\vec{a}|)$-small in $H^{\bar{s}}(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$, i.e. $\|r_i\|_{\bar{s}}/|\vec{a}| \to 0$ as $|\vec{a}| \to 0$ for $i = 1, 2$. The solution $(\eta(\widetilde{\omega}t, x), \psi(\widetilde{\omega}t, x))$ is even in x, η is even in t and ψ is odd in t. In addition these quasi-periodic solutions are linearly stable, see Theorem 1.2.

Let us make some comments on the result.

No global wellposedness 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 [33] 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 [37].

In the case of periodic boundary conditions, Ifrim-Tataru [36] 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 Nash-Moser-KAM iterative procedure used to prove Theorem 1.1 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. 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 $|\vec{a}| \leq \varepsilon_0$. 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.

The dynamics of the pure gravity and gravity-capillary water waves equations is very different:

- (i) the pure gravity water waves vector field in (1.14) is a *singular perturbation* of the linearized vector field at the origin in (1.15), which, after symmetrization, is $|D_x|^{\frac12}\tanh^{\frac12}(\mathtt{h}|D_x|)$; in fact, the linearization of the nonlinearity gives rise to a transport vector field $V\partial_x$, see (1.43). On the other hand, the gravity capillary vector field is quasi-linear and contains derivatives of the same order as the linearized vector field at the origin, which is $\sim |D_x|^{\frac32}$. This difference, which is well known in the water waves literature, requires a very different analysis of the linearized operator (Sections 6-12) with respect to the gravity capillary case in [1], [21], see Remark 1.4.

- (ii) The linear frequencies ω_j in (1.19) of the pure gravity water waves grow like $\sim j^{\frac12}$ as $j \to +\infty$, while, in presence of surface tension κ, the linear frequencies are $\sqrt{(1+\kappa j^2) j \tanh(\mathtt{h}j)} \sim j^{\frac32}$. This makes a substantial difference for the development of KAM theory. In presence of a sublinear growth of the linear frequencies $\sim j^\alpha$, $\alpha < 1$, one may impose only very weak second order Melnikov non-resonance conditions, see e.g. (1.36), which lose also space (and not only time) derivatives along the KAM reducibility scheme. This is not the case of the abstract infinite-dimensional KAM theorems [44], [45], [54] where the linear frequencies grow as j^α, $\alpha \geq 1$, and the perturbation is bounded. In this paper we overcome the difficulties posed by the sublinear growth $\sim j^{\frac12}$ and by the unboundedness of the water waves vector field thanks to a regularization procedure performed on the linearized PDE at each approximate quasi-periodic solution obtained along a Nash-Moser iterative scheme, see the regularized system (1.41). This regularization strategy is in principle applicable to a broad class of PDEs where the second order Melnikov non-resonance conditions lose space derivatives.

(iii) The linear frequencies (1.19) vary with $\mathtt{h}$ only by exponentially small quantities: they admit the asymptotic expansion

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

uniformly in $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, where the constant C_k depends only on k and $\mathtt{h}_1$. Nevertheless we shall be able, extending the degenerate KAM theory approach in [11], [21], to use the finite depth parameter $\mathtt{h}$ to impose the required Melnikov non-resonance conditions, see (1.36) and Sections 3 and 4.2. On the other hand, for the gravity capillary water waves considered in [21], the surface tension parameter κ moves the linear frequencies $\sqrt{(1+\kappa j^2)j\tanh(\mathtt{h}j)}$ of polynomial quantities $O(j^{3/2})$.

Linear stability. The quasi-periodic solutions $u(\widetilde{\omega}t) = (\eta(\widetilde{\omega}t), \psi(\widetilde{\omega}t))$ found in Theorem 1.1 are linearly stable. Since 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 [53], [42], [38], [39], [1]), we state precisely the result.

The quasi-periodic solutions (1.23) are mainly supported in Fourier space on the tangential sites $\mathbb{S}^+$. As a consequence, 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. In 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 4.

The classical normal form formulation of KAM theory for lower dimensional tori, see for instance [44]-[45], [54], [43], [29], [55], [13]-[14], [63], [49], provides, when applicable, existence and linear stability of quasi-periodic solutions at the same time. On the other hand, existence (without linear stability) of periodic and quasi-periodic solutions of PDEs has been proved by using the Lyapunov-Schmidt decomposition combined with Nash-Moser implicit function theorems, see e.g. [22], [24], [53], [42], [38], [39], [25], [6], [1] and references therein. In this paper we follow the Nash Moser approach to KAM theory outlined in [16] and implemented in [8], [21], which combines ideas of both formulations, see Section 1.2 "Analysis of the linearized operators" and Section 5.

We prove that around each torus filled by the quasi-periodic solutions (1.23) of the Hamiltonian system (1.14) constructed in Theorem 1.1 there exist symplectic coordinates $(\phi, y, w) = (\phi, y, \eta, \psi) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_{\mathbb{S}^+}^{\perp}$ (see (5.16) and [16]) in which the water waves Hamiltonian reads

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

where $K_{\ge 3}$ collects the terms at least cubic in the variables (y, w) (see (5.18) and note that, at a solution, one has $\partial_\phi K_{00} = 0$, $K_{10} = \widetilde{\omega}$, $K_{01} = 0$ by Lemma 5.4). The (ϕ, y) coordinates are modifications of the action-angle variables and w is a translation of the cartesian variable z in the normal subspace, see (5.16). In these coordinates the quasi-periodic solution reads $t \mapsto (\widetilde{\omega}t, 0, 0)$ and the corresponding linearized water waves equations are

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

The self-adjoint operator $K_{02}(\widetilde{\omega}t)$ is defined in (5.18) and $JK_{02}(\widetilde{\omega}t)$ is the restriction to $H_{\mathbb{S}^+}^{\perp}$ of the linearized water waves vector field $J\partial_u \nabla_u H(u(\widetilde{\omega}t))$ (computed explicitly in (6.8)) up to a finite dimensional remainder, see Lemma 6.1.

We have the following result of linear stability for the quasi-periodic solutions found in Theorem 1.1.

Theorem 1.2. (Linear stability) *The quasi-periodic solutions* (1.23) *of the pure gravity water waves system are linearly stable, meaning that for all s belonging to a suitable interval* $[s_1, s_2]$*, for any initial datum* $y(0) \in \mathbb{R}^\nu$, $w(0) \in H_x^{s-\frac{1}{4}} \times H_x^{s+\frac{1}{4}}$*, the solutions* $y(t)$, $w(t)$ *of system* (1.27) *satisfy*

$$y(t) = y(0), \quad \|w(t)\|_{H_x^{s-\frac{1}{4}} \times H_x^{s+\frac{1}{4}}} \leq C\big(\|w(0)\|_{H_x^{s-\frac{1}{4}} \times H_x^{s+\frac{1}{4}}} + |y(0)|\big) \quad \forall t \in \mathbb{R}.$$

In fact, by (1.27), the actions $y(t) = y(0)$ do not evolve in time and the third equation reduces to the linear PDE

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

Sections 6-14 imply the existence of a transformation $(H_x^s \times H_x^s) \cap H_{\mathbb{S}^+}^\perp \to (H_x^{s-\frac{1}{4}} \times H_x^{s+\frac{1}{4}}) \cap H_{\mathbb{S}^+}^\perp$, bounded and invertible for all $s \in [s_1, s_2]$, such that, in the new variables $\mathtt{w}_\infty$, the homogeneous equation $\dot{w} = JK_{02}(\widetilde{\omega}t)[w]$ transforms into a system of infinitely many uncoupled scalar and time independent ODEs of the form

$$\partial_t \mathtt{w}_{\infty,j} = -\mathrm{i}\mu_j^\infty \mathtt{w}_{\infty,j}\,, \quad \forall j \in \mathbb{S}_0^c\,, \tag{1.29}$$

where i is the imaginary unit, $\mathbb{S}_0^c := \mathbb{Z} \setminus \mathbb{S}_0$, $\mathbb{S}_0 := \mathbb{S}^+ \cup (-\mathbb{S}^+) \cup \{0\} \subseteq \mathbb{Z}$, the eigenvalues μ_j^∞ are (see (4.26), (4.27))

$$\mu_j^\infty := \mathtt{m}_{\frac{1}{2}}^\infty |j|^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h}|j|) + \mathfrak{r}_j^\infty \in \mathbb{R}, \quad j \in \mathbb{S}_0^c\,, \quad \mathfrak{r}_j^\infty = \mathfrak{r}_{-j}^\infty\,, \tag{1.30}$$

and $\mathtt{m}_{\frac{1}{2}}^\infty = 1 + O(|\vec{a}|^c)$, $\sup_{j \in \mathbb{S}_0^c} |j|^{\frac{1}{2}} |\mathfrak{r}_j^\infty| = O(|\vec{a}|^c)$ for some $c > 0$, see (4.28). Since μ_j^∞ are even in j, equations (1.29) can be equivalently written in the basis $(\cos(jx))_{j \in \mathbb{N} \setminus \mathbb{S}^+}$ of functions even in x; in Section 14, for convenience, we represent even operators in the exponential basis $(e^{\mathrm{i}jx})_{j \in \mathbb{S}_0^c}$. The above result is the *reducibility* of the linearized quasi-periodically time dependent equation $\dot{w} = JK_{02}(\widetilde{\omega}t)[w]$. The *Floquet exponents* of the quasi-periodic solution are the purely imaginary numbers $\{0, \mathrm{i}\mu_j^\infty, j \in \mathbb{S}_0^c\}$ (the null Floquet exponent comes from the action component $\dot{y} = 0$). Since μ_j^∞ are real, the Sobolev norms of the solutions of (1.29) are constant.

The reducibility of the linear equation $\dot{w} = JK_{02}(\widetilde{\omega}t)[w]$ is obtained by two well-separated procedures:

1. First, we 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, see (1.41). We perform such a reduction in Sections 6-13.

2. Then, we implement in Section 14 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 both in time and in space. This loss is compensated along the KAM scheme by the smoothing nature of the variable coefficients remainders. Actually, in Section 14 we explicitly state only a result of almost-reducibility (in Theorems 14.3-14.4 we impose only finitely many Melnikov non-resonance conditions and there appears a remainder $\mathcal{R}_n$ of size $O(N_n^{-\mathtt{a}})$, where $\mathtt{a} > 0$ is the large parameter fixed in (14.7)), because this is sufficient for the construction of the quasi-periodic solutions. However the frequencies of the quasi-periodic solutions that we construct in Theorem 1.1 satisfy all the infinitely many Melnikov non-resonance conditions in (4.29) and Theorems 14.3-14.4 pass to the limit as $n \to \infty$, leading to (1.29).

We shall explain these steps in detail in Section 1.2. In the pioneering works of Plotnikov-Toland [53] and Iooss-Plotnikov-Toland [42] dealing with time-periodic solutions of the water waves equations, as well as in [1], the latter diagonalization is not required. 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 definitely not true for quasi-periodic solutions.

Literature about KAM for quasilinear PDEs. KAM theory for PDEs has been developed to a large extent for bounded perturbations and for linear frequencies growing in a superlinear way, as j^α, $\alpha \geq 1$. The case $\alpha = 1$, which corresponds to 1d wave and Klein-Gordon equations, is more delicate. In the sublinear case $\alpha < 1$, as far as we know, there are no general KAM results, since the second order Melnikov conditions lose space derivatives. Since the eigenvalues of $-\Delta$ on $\mathbb{T}^d$ grow, according to the Weyl law, like $\sim j^{2/d}$, $j \in \mathbb{N}$, one could regard the KAM results for multidimensional Schrödinger and wave equations in [22], [29], [15], [18], [55], under this perspective. Actually the proof of these results exploits specific properties of clustering of the eigenvalues of the Laplacian.

The existence of quasi-periodic solutions of PDEs with *unbounded* perturbations (i.e. the nonlinearity contains derivatives) has been first proved by Kuksin [45] and Kappeler-Pöschel [43] for KdV, then by Liu-Yuan [49], Zhang-Gao-Yuan [63] for derivative NLS, and by Berti-Biasco-Procesi [13]-[14] for derivative wave equation. 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 or fully nonlinear PDEs the first KAM results have been recently proved by Baldi-Berti-Montalto in [7], [8], [9] for perturbations of Airy, KdV and mKdV equations, introducing tools of pseudo-differential calculus for the spectral analysis of the linearized equations. In particular, [7] concerns quasi-periodically forced perturbations of the Airy equation

$$u_t + u_{xxx} + \varepsilon f(\omega t, x, u, u_x, u_{xx}, u_{xxx}) = 0 \tag{1.31}$$

where the forcing frequency ω is an external parameter. The key step is the reduction of the linearized operator at an approximate solution to constant coefficients up to a sufficiently *smoothing* remainder, followed by a KAM reducibility scheme leading to its complete diagonalization. Once such a reduction has been achieved, the second order Melnikov nonresonance conditions required for the diagonalization are easily imposed since the frequencies are $\sim j^3$ and using ω as external parameters. Because of the purely differential structure of (1.31), the required tools of pseudo-differential calculus are mainly multiplication operators and Fourier multipliers. These techniques have been extended by Feola-Procesi [31] for quasi-linear forced perturbations of Schrödinger equations and by Montalto [51] for the forced Kirchhoff equation.

The paper [8] deals with the more difficult case of completely resonant autonomous Hamiltonian perturbed KdV equations of the form

$$u_t + u_{xxx} - 6uu_x + f(x, u, u_x, u_{xx}, u_{xxx}) = 0\,. \tag{1.32}$$

Since the Airy equation $u_t + u_{xxx} = 0$ possesses only 2π-periodic solutions, the existence of quasi-periodic solutions of (1.32) is entirely due to the nonlinearity, which determines the modulation of the tangential frequencies of the solutions with respect to its amplitudes. This is achieved via "weak" Birkhoff normal form transformations that are close to the identity up to finite rank operators. The paper [8] implements the general symplectic procedure proposed in [16] for autonomous PDEs, which reduces the construction of an approximate inverse of the linearized operator to the construction of an approximate inverse of its restriction to the normal directions. This is obtained along the lines of [7], but with more careful size estimates because (1.32) is a completely resonant PDE. The symplectic procedure of [16] is also applied in [21] and in Section 5 of the present paper. We refer to [23] and [32] for a similar reduction which applies also to PDEs which are not Hamiltonian, but for example reversible.

By further extending these techniques, the existence of quasi-periodic solutions of gravity capillary water waves has been recently proved in [21]. In items (i)-(iii) after Theorem 1.1 we have described the major differences between the pure gravity and gravity-capillary water waves equations and we postpone to Remark 1.4 more comments about the differences regarding the reducibility of the linearized equations.

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.

1.2 Ideas of the proof

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

(i) The nonlinear water waves system (1.14) is a singular perturbation of (1.15).

(ii) The dispersion relation (1.19) is sublinear, i.e. $\omega_j \sim \sqrt{j}$ for $j \to \infty$.

(iii) The linear frequencies $\omega_j(\mathtt{h}) = j^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h} j)$ vary with $\mathtt{h}$ of just exponentially small quantities.

We present below the key ideas to solve these three major problems.

Nash-Moser Theorem 4.1 of hypothetical conjugation. In Section 4 we rescale $u \mapsto \varepsilon u$ and introduce the action angle variables $(\theta, I) \in \mathbb{T}^\nu \times \mathbb{R}^\nu$ on the tangential sites (see (1.25))

$$\eta_j := \sqrt{\frac{2}{\pi}}\, \omega_j^{\frac{1}{2}} \sqrt{\xi_j + I_j} \cos(\theta_j), \quad \psi_j := -\sqrt{\frac{2}{\pi}}\, \omega_j^{-\frac{1}{2}} \sqrt{\xi_j + I_j}\, \sin(\theta_j)\,, \quad j \in \mathbb{S}^+\,, \tag{1.33}$$

where $\xi_j > 0$, $j = 1, \ldots, \nu$, the variables I_j satisfy $|I_j| < \xi_j$, so that system (1.14) becomes the Hamiltonian system generated by

$$H_\varepsilon = \vec{\omega}(\mathtt{h}) \cdot I + \frac{1}{2}(z, \Omega z)_{L^2} + \varepsilon P\,, \quad \vec{\omega}(\mathtt{h}) := (j^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h} j))_{j \in \mathbb{S}^+}, \tag{1.34}$$

where P is given in (4.8). The unperturbed actions ξ_j in (1.33) and the unperturbed amplitudes a_j in (1.22) and Theorem 1.1 are related by the identity $a_j = \varepsilon \sqrt{(2/\pi)}\, \omega_j^{\frac{1}{2}} \sqrt{\xi_j}$ for all $j \in \mathbb{S}^+$.

The expected quasi-periodic solutions of the autonomous Hamiltonian system generated by H_ε will have shifted frequencies $\widetilde{\omega}_j$ – 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 frequencies $\omega \in \mathbb{R}^\nu$ as parameters, introducing "counter-terms" $\alpha \in \mathbb{R}^\nu$ (as in [21], in the spirit of Herman-Féjoz [30]) in the family of Hamiltonians (see (4.9))

$$H_\alpha := \alpha \cdot I + \frac{1}{2}(z, \Omega z)_{L^2} + \varepsilon P\,.$$

Then the first goal (Theorem 4.1) is to prove that, for ε small enough, there exist $\alpha_\infty(\omega, \mathtt{h}, \varepsilon)$, close to ω, and a ν-dimensional embedded torus $i_\infty(\varphi; \omega, \mathtt{h}, \varepsilon)$ of the form

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

close to $(\varphi, 0, 0)$, defined for all $(\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, such that, for all $(\omega, \mathtt{h})$ belonging to the set $\mathcal{C}^\gamma_\infty$ defined in (4.20), $(i_\infty, \alpha_\infty)(\omega, \mathtt{h}, \varepsilon)$ is a zero of the nonlinear operator (see (4.10))

$$\mathcal{F}(i, \alpha, \omega, \mathtt{h}, \varepsilon) := \begin{pmatrix} \omega \cdot \partial_\varphi \theta(\varphi) - \alpha - \varepsilon \partial_I P(i(\varphi)) \\ \omega \cdot \partial_\varphi I(\varphi) + \varepsilon \partial_\theta P(i(\varphi)) \\ \omega \cdot \partial_\varphi z(\varphi) - J(\Omega z(\varphi) + \varepsilon \nabla_z P(i(\varphi))) \end{pmatrix}. \tag{1.35}$$

The explicit set $\mathcal{C}^\gamma_\infty$ requires ω to satisfy, in addition to the Diophantine property

$$|\omega \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau} \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\,, \qquad \langle \ell \rangle := \max\{1, |\ell|\}, \quad |\ell| := \max_{i=1,\ldots,\nu} |\ell_i|\,,$$

the first and second Melnikov non-resonance conditions stated in (4.20), in particular

$$|\omega \cdot \ell + \mu^\infty_j(\omega, \mathtt{h}) - \mu^\infty_{j'}(\omega, \mathtt{h})| \geq 4\gamma j^{-\mathtt{d}} j'^{-\mathtt{d}} \langle \ell \rangle^{-\tau}, \ \forall \ell \in \mathbb{Z}^\nu, \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+, \ (\ell, j, j') \neq (0, j, j)\,, \tag{1.36}$$

where $\mu^\infty_j(\omega, \mathtt{h})$ are the "final eigenvalues" in (4.18), defined for all $(\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ (we use the abstract Whitney extension theorem in Appendix B). The torus i_∞, the conter-term α_∞ and the final eigenvalues $\mu^\infty_j(\omega, \mathtt{h})$ are $\mathcal{C}^{k_0}$ differentiable with respect to the parameters $(\omega, \mathtt{h})$. The value of k_0 is fixed in Section 3, depending only on the unperturbed linear frequencies, so that transversality conditions like (1.39) hold, see Proposition 3.4. The value of the counterterm $\alpha := \alpha_\infty(\omega, \mathtt{h}, \varepsilon)$ is adjusted along the Nash-Moser iteration in order to control the average of the first component of the Hamilton equation (4.10), especially for solving the linearized equation (5.35), in particular (5.39).

Theorem 4.1 follows by the Nash-Moser Theorem 15.1 which relies on the analysis of the linearized operators $d_{i,\alpha}\mathcal{F}$ at an approximate solution, performed in Sections 5-14. The formulation of Theorem 4.1

is convenient as it completely decouples the Nash-Moser iteration required to prove Theorem 1.1 and the discussion about the measure of the set of parameters $\mathcal{C}^{\gamma}_{\infty}$ where all the Melnikov non-resonance conditions are verified. In Section 4.2 we are able to prove positive measure estimates, if the exponent $\mathtt{d}$ in (1.36) is large enough and $\gamma = o(1)$ as $\varepsilon \to 0$. Since such a value of $\mathtt{d}$ determines the number of regularization steps to be performed on the linearized operator, we prefer to first discuss how we fix it, applying degenerate KAM theory.

Proof of Theorem 1.1: degenerate KAM theory and measure estimates. In order to prove the existence of quasi-periodic solutions of the system with Hamiltonian H_ε in (1.34), thus (1.14), and not only of the system with modified Hamiltonian H_α with $\alpha := \alpha_\infty(\omega, \mathtt{h}, \varepsilon)$, we have to prove that the curve of the unperturbed linear tangential frequencies

$$[\mathtt{h}_1, \mathtt{h}_2] \ni \mathtt{h} \mapsto \vec{\omega}(\mathtt{h}) := \big(\sqrt{j \tanh(\mathtt{h} j)}\big)_{j \in \mathbb{S}^+} \in \mathbb{R}^\nu \tag{1.37}$$

intersects the image $\alpha_\infty(\mathcal{C}^{\gamma}_{\infty})$ of the set $\mathcal{C}^{\gamma}_{\infty}$ under the map α_∞, for "most" values of $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$. Setting

$$\omega_\varepsilon(\mathtt{h}) := \alpha_\infty^{-1}(\vec{\omega}(\mathtt{h}), \mathtt{h}) \,, \tag{1.38}$$

where $\alpha_\infty^{-1}(\cdot, \mathtt{h})$ is the inverse of the function $\alpha_\infty(\cdot, \mathtt{h})$ at a fixed $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, if the vector $(\omega_\varepsilon(\mathtt{h}), \mathtt{h})$ belongs to $\mathcal{C}^{\gamma}_{\infty}$, then Theorem 4.1 implies the existence of a quasi-periodic solution of the system with Hamiltonian H_ε with Diophantine frequency $\omega_\varepsilon(\mathtt{h})$.

In Theorem 4.2 we prove that for all the values of $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$ except a set of small measure $O(\gamma^{1/k_0^*})$ (where k_0^* is the index of non-degeneracy appearing below in (1.39)), the vector $(\omega_\varepsilon(\mathtt{h}), \mathtt{h})$ belongs to $\mathcal{C}^{\gamma}_{\infty}$. Since the parameter interval $[\mathtt{h}_1, \mathtt{h}_2]$ is fixed, independently of the $O(\varepsilon)$-neighborhood of the origin where we look for the solutions, the small divisor constant γ in the definition of $\mathcal{C}^{\gamma}_{\infty}$ (see e.g. (1.36)) can be taken as $\gamma = \varepsilon^a$ with $a > 0$ as small as needed, see (4.22), so that all the quantities $\varepsilon \gamma^{-\kappa}$ that we encounter along the proof are $\ll 1$.

The first task is to prove a transversality property for the unperturbed tangential frequencies $\vec{\omega}(\mathtt{h})$ in (1.37) and the normal ones $\vec{\Omega}(\mathtt{h}) := (\Omega_j(\mathtt{h}))_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+} := (\omega_j(\mathtt{h}))_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+}$. Exploiting the fact that the maps $\mathtt{h} \mapsto \omega_j(\mathtt{h}^4)$ are *analytic*, simple – namely injective in j – in the subspace of functions even in x, and they grow asymptotically like $\sqrt{j}$ for $j \to \infty$, we first prove that the linear frequencies $\omega_j(\mathtt{h})$ are *non-degenerate* in the sense of Bambusi-Berti-Magistrelli [11] (i.e. they are not contained in any hyperplane). This is verified in Lemma 3.2 using a generalized Vandermonde determinant (see Lemma 3.3). Then in Proposition 3.4 we translate this qualitative non-degeneracy condition into quantitative transversality information: there exist $k_0^* > 0, \rho_0 > 0$ such that, for all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$,

$$\max_{0 \le k \le k_0^*} \big| \partial_{\mathtt{h}}^k \big(\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h}) \big) \big| \ge \rho_0 \langle \ell \rangle \,, \quad \forall \ell \ne 0 \,, \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \,, \tag{1.39}$$

and similarly for the 0th, 1st and 2nd order Melnikov non-resonance condition with the + sign. We call (following [58]) k_0^* the index of non-degeneracy and ρ_0 the amount of non-degeneracy. 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]).

The transversality condition (1.39) is stable under perturbations that are small in $\mathcal{C}^{k_0}$-norm, where $k_0 := k_0^* + 2$, see Lemma 4.4. Since $\omega_\varepsilon(\mathtt{h})$ in (1.38) and the perturbed Floquet exponents $\mu_j^\infty(\mathtt{h}) = \mu_j^\infty(\omega_\varepsilon(\mathtt{h}), \mathtt{h})$ in (4.26) are small perturbations of the unperturbed linear frequencies $\sqrt{j \tanh(\mathtt{h} j)}$ in $\mathcal{C}^{k_0}$-norm, the transversality property (1.39) still holds for the perturbed frequencies. As a consequence, by applying the classical Rüssmann lemma (Theorem 17.1 in [58]) we prove that, for most $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$, the 0th, 1st and 2nd Melnikov conditions on the perturbed frequencies hold if $\mathtt{d} > \frac{3}{4} k_0^*$, see Lemma 4.5 and (4.46).

The larger is $\mathtt{d}$, the weaker are the Melnikov conditions (1.36), and the stronger will be the loss of space derivatives due to the small divisors in the reducibility scheme of Section 14. 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 6-13, 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 (14.8), which is large with respect to $\mathtt{d}$ and τ by (14.7). We will explain in detail this procedure below.

Analysis of the linearized operators. In order to prove the existence of a solution of $\mathcal{F}(i,\alpha)=0$ in (1.35), proving the Nash-Moser Theorem 4.1, the key step is to show that the linearized operator $d_{i,\alpha}\mathcal{F}$ obtained at any approximate solution along the iterative scheme admits an *almost approximate inverse* satisfying tame estimates in Sobolev spaces with loss of derivatives, see Theorem 5.6. Following the terminology of Zehnder [62], an *approximate inverse* is an operator which is an exact inverse at a solution (note that the operator $\mathcal{P}$ in (5.48) is zero when $\mathcal{F}(i,\alpha)=0$). The adjective *almost* refers to the fact that at the n-th step of the Nash-Moser iteration we shall require only finitely many non-resonance conditions of Diophantine type, therefore there remain operators (like (5.49)) that are Fourier supported on high frequencies of magnitude larger than $O(N_n)$ and thus they can be estimated as $O(N_n^{-a})$ for some $a>0$ (in suitable norms). The tame estimates (5.48)-(5.51) are sufficient for the convergence of a differentiable Nash-Moser scheme: the remainder (5.48) produces a quadratic error since it is of order $O(\mathcal{F}(i_n,\alpha_n))$; the remainder (5.49) arising from the almost-reducibility is small enough by taking $\mathtt{a}>0$ sufficiently large, as in (14.7); the remainder (5.50) arises by ultraviolet cut-off truncations and its contribution is small by usual differentiable Nash-Moser mechanisms, see for instance [17]. These abstract tame estimates imply the Nash-Moser Theorem 15.1.

In order to find an almost approximate inverse of $d_{i,\alpha}\mathcal{F}$ we first implement the strategy of Section 5 introduced in Berti-Bolle [16], which is based on the following simple observation: around an invariant torus there are symplectic coordinates (ϕ,y,w) in which the Hamiltonian assumes the normal form (1.26) and therefore the linearized equations at the quasi-periodic solution assume the *triangular* form as in (1.27). In these new coordinates it is immediate to solve the equations in the variables ϕ,y, and it remains to invert an operator acting on the w component, which is precisely $\mathcal{L}_\omega$ defined in (5.26). By Lemma 6.1 the operator $\mathcal{L}_\omega$ is a finite rank perturbation (see (6.5)) of the restriction to the normal subspace $H_{\mathbb{S}^+}^{\perp}$ in (1.25) of

$$\mathcal{L}=\omega\cdot\partial_\varphi+\begin{pmatrix}\partial_x V+G(\eta)B & -G(\eta)\\ (1+BV_x)+BG(\eta)B & V\partial_x-BG(\eta)\end{pmatrix} \tag{1.40}$$

where the functions B,V are given in (6.7), which is obtained linearizing the water waves equations (1.14) at a quasi-periodic approximate solution $(\eta,\psi)(\omega t,x)$ and changing ∂_t into the directional derivative $\omega\cdot\partial_\varphi$.

If $\mathcal{F}(i,\alpha)$ is not zero but it is small, we say that i is approximately invariant for X_{H_α}, and, following [16], in Section 5 we transform $d_{i,\alpha}\mathcal{F}$ into an *approximately* triangular operator, with an error of size $O(\mathcal{F}(i,\alpha))$. In this way, we have reduced the problem of almost approximately inverting $d_{i,\alpha}\mathcal{F}$ to the task of almost inverting the operator $\mathcal{L}_\omega$. The precise invertibility properties of $\mathcal{L}_\omega$ are stated in (5.29)-(5.33).

Remark 1.3. The main advantage of this approach is that the problem of inverting $d_{i,\alpha}\mathcal{F}$ on the whole space (i.e. both tangential and normal modes) is reduced to invert a PDE on the normal subspace $H_{\mathbb{S}^+}^{\perp}$ only. In this sense this is reminiscent of the Lyapunov-Schmidt decomposition, where the complete nonlinear problem is split into a bifurcation and a range equation on the orthogonal of the kernel. However, the Lyapunov-Schmidt approach is based on a splitting of the space $H^s(\mathbb{T}^{\nu+1})$ of functions $u(\varphi,x)$ of time and space, whereas the approach of [16] splits the phase space (of functions of x only) into $H_{\mathbb{S}^+}\oplus H_{\mathbb{S}^+}^{\perp}$ more similarly to a classical KAM theory formulation. □

The procedure of Section 5 is a preparation for the reducibility of the linearized water waves equations in the normal subspace developed in Sections 6-14, where we conjugate the operator $\mathcal{L}_\omega$ to a diagonal system of infinitely many decoupled, constant coefficients, scalar linear equations, see (1.42) below. First, in Sections 6-12, in order to use the tools of pseudo-differential calculus, it is convenient to ignore the projection on the normal subspace $H_{\mathbb{S}^+}^{\perp}$ and to perform a regularization procedure on the operator $\mathcal{L}$ acting on the whole space, see Remark 6.2. Then, in Section 13, we project back on $H_{\mathbb{S}^+}^{\perp}$. Our approach involves two well separated procedures that we describe in detail:

1. **Symmetrization and diagonalization of $\mathcal{L}$ up to smoothing operators**. The goal of Sections 6-12 is to conjugate $\mathcal{L}$ to an operator of the form

$$\omega\cdot\partial_\varphi+\mathrm{i}\mathtt{m}_{\frac12}|D|^{\frac12}\tanh^{\frac12}(\mathtt{h}|D|)+\mathrm{i}r(D)+\mathcal{T}_8(\varphi) \tag{1.41}$$

where $\mathtt{m}_{\frac12}\approx 1$ 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 $\mathcal{T}_8(\varphi)$, as well as $\partial_\varphi^\beta\mathcal{T}_8$ for all $|\beta|\le\beta_0$ 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 13 we restrict the operator in (1.41) to $H_{\mathbb{S}_+}^{\perp}$ and in Section 14 we implement an iterative diagonalization scheme to reduce quadratically the size of the perturbation, completing the conjugation of $\mathcal{L}_\omega$ to a diagonal, constant coefficient system of the form

$$\omega \cdot \partial_\varphi + \mathrm{i}\mathrm{Op}(\mu_j) \tag{1.42}$$

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.

We underline that all the transformations performed in Sections 6-14 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 inverting the final operator (1.42) directly provide good tame estimates for the inverse of the operator $\mathcal{L}_\omega$ in (6.5).

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 (1.42) 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.

Remark 1.4. (Comparison with the gravity-capillary linearized PDE) With respect to the gravity capillary water waves in infinite depth in [1], [21], the reduction in decreasing orders of the linearized operator is completely different. The linearized operator in the gravity-capillary case is like

$$\omega \cdot \partial_\varphi + \mathrm{i}|D_x|^{\frac{3}{2}} + V(\varphi, x)\partial_x \,,$$

the term $V\partial_x$ is a lower order perturbation of $|D_x|^{\frac{3}{2}}$, and it can be reduced to constant coefficients by conjugating the operator with a "semi-Fourier Integral Operator" A of type $(\frac{1}{2}, \frac{1}{2})$ (like in [1] and [21]): the commutator of $|D_x|^{\frac{3}{2}}$ and A produces a new operator of order 1, and one chooses appropriately the symbol of A for the reduction of $V\partial_x$. Instead, in the pure gravity case we have a linearized operator of the type

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

where the term $V\partial_x$ is a *singular* perturbation of $\mathrm{i}|D_x|^{\frac{1}{2}}$. The commutator between $|D_x|^{\frac{1}{2}}$ and any bounded pseudo-differential operator produces operators of order $\leq 1/2$, which do not interact with $V\partial_x$. Hence one uses the commutator with $\omega\cdot\partial_\varphi$ (which is the leading term of the unperturbed operator) to produce operators of order 1 that cancel out $V\partial_x$. This is why our first task is to straighten the first order vector field (1.44), which corresponds to a time quasi-periodic transport operator. Furthermore, the fact that the unperturbed linear operator is $\sim |D|^{\frac{1}{2}}$, unlike $\sim |D|^{\frac{3}{2}}$, also affects the conjugation analysis of the lower order operators, where the contribution of the commutator with $\omega \cdot \partial_\varphi$ is always of order higher than the commutator with $|D_x|^{\frac{1}{2}}$. As a consequence, in the procedure of reduction of the symbols to constant coefficients in Sections 11-12, we remove first their dependence on φ, and then their dependence on x. We also note that in [21], since the second order Melnikov conditions do not lose space derivatives, there is no need to perform such reduction steps at negative orders before starting with the KAM reducibility algorithm. $\square$

We now explain in details the steps of the conjugation of the quasi-periodic linear operator (1.40) described in the items 1 and 2 above. We underline that all the coefficients of the linearized operator $\mathcal{L}$ in (1.40) 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 (15.1)) 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 6.1 the linearized good unknown of Alinhac, as in [1] and [21]. This is indeed the same change of variable introduced by Lannes [46] (see also [47]) for proving energy estimates for the local existence theory. Subsequently, the nonlinear good unknown of Alinhac has been introduced by Alazard-Métivier [5], see also [2]-[4] to perform the paralinearization of the Dirichlet-Neumann operator. In these new variables, the linearized operator (1.40) becomes the more symmetric operator (see (6.15))

$$\mathcal{L}_0 = \omega \cdot \partial_\varphi + \begin{pmatrix} \partial_x V & -G(\eta) \\ a & V\partial_x \end{pmatrix} = \omega \cdot \partial_\varphi + \begin{pmatrix} V\partial_x & 0 \\ 0 & V\partial_x \end{pmatrix} + \begin{pmatrix} V_x & -G(\eta) \\ a & 0 \end{pmatrix}, \tag{1.43}$$

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 Appendix A we provide a self-contained proof of such a representation. We cannot directly use a result already existing in the literature (for the Cauchy problem) because we have to provide tame estimates for the action of $G(\eta)$ on Sobolev spaces of time-space variables (φ, x) and to control its smooth dependence with respect to the parameters $(\omega, \mathtt{h})$. We can neither directly apply the corresponding result of [21], which is given in the case $\mathtt{h} = +\infty$.

Notice that the first order transport operator $V\partial_x$ in (1.43) is a singular perturbation of $\mathcal{L}_0$ evaluated at $(\eta, \psi) = 0$, i.e. $\omega \cdot \partial_\varphi + \left(\begin{smallmatrix} 0 & -G(0) \\ 1 & 0 \end{smallmatrix}\right)$.

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)

$$\omega \cdot \partial_\varphi + V(\varphi, x)\partial_x \tag{1.44}$$

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 *symmetry* induced by the *reversible* structure of the water waves equations, 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 (7.4). Actually we solve equation (7.6) 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 7.1, [53], [42] and [52]. We solve this problem in Section 7 using a Nash-Moser implicit function theorem. Actually, after having inverted the linearized operator at an approximate solution (Lemma 7.2), we apply the Nash-Moser-Hörmander Theorem C.1, proved in Baldi-Haus [10]. We cannot directly use already existing results for equation (7.6) because we have to prove tame estimates and Lipschitz dependence of the solution with respect to the approximate torus, as well as its smooth dependence with respect to the parameters $(\omega, \mathtt{h})$, see Lemmata 7.4-7.5.

We remark that, when searching for time periodic solutions as in [42], [53], the corresponding transport equation is not a small-divisor problem and has been solved in [53] by a direct ODE analysis.

In Lemma 7.6 we apply this change of variable to the whole operator $\mathcal{L}_0$ in (1.43), obtaining the new conjugated system (see (7.31))

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

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

3. Change of the space variable. In Section 8 we introduce a change of variable induced by a diffeomorphism of $\mathbb{T}_x$ of the form (independent of φ)

$$y = x + \alpha(x) \qquad \Leftrightarrow \qquad x = y + \breve{\alpha}(y) . \tag{1.45}$$

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

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

see (8.5), 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 11 the function $\alpha(x)$ (see (11.23)) in order to eliminate the dependence on x from the time average $\langle a_7 \rangle_\varphi(x)$ in (11.17)-(11.18) of the coefficient of $|D_x|^{\frac{1}{2}}$. The advantage of introducing the diffeomorphism (1.45) at this step, rather than in Section 11 where it is used, is that it is easier 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 (on the other hand, performing this transformation in Section 11 would require delicate estimates of the symbols obtained after an Egorov-type analysis).

4. Symmetrization of the order $1/2$. In Section 9 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 (9.10)-(9.14), up to lower order operators. The function a_7 is close to 1 and $\breve{a}_4$ is small in ε, see (9.17). Notice that the off-diagonal operators in $\mathcal{L}_3$ are opposite to each other, unlike in $\mathcal{L}_2$. Then, in the complex unknown $h = \eta + \mathrm{i}\psi$, the first component of such an operator reads

$$(h, \bar{h}) \mapsto \omega \cdot \partial_\varphi h + \mathrm{i} a_7 |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} h + a_8 h + P_5 h + Q_5 \bar{h} \tag{1.46}$$

(which corresponds to (10.1) 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 $(h, \bar{h})$.

In Sections 10-12 we perform the reduction to constant coefficients of (1.46) up to smoothing operators, dealing separately with the diagonal and off-diagonal operators.

5. Symmetrization of the lower orders. In Section 10 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 10.3)

$$(h, \bar{h}) \mapsto \omega \cdot \partial_\varphi h + \mathrm{i} a_7(\varphi, x) |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} h + a_8 h + P_6 h + Q_6 \bar{h} , \tag{1.47}$$

where P_6 is in $OPS^{-\frac{1}{2}}$ and $Q_6 \in OPS^{-M}$ for a constant M large enough fixed in Section 14, in view of the reducibility scheme.

6. Time and space reduction at the order $1/2$. In Section 11 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 (1.47) 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 8), see formulas (11.19) and (11.23), the final outcome is a linear operator of the form, see (11.31),

$$(h, \bar{h}) \mapsto \omega \cdot \partial_\varphi h + \mathrm{i} \mathtt{m}_{\frac{1}{2}} |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} h + (a_8 + a_9 \mathcal{H}) h + P_7 h + \mathcal{T}_7(h, \bar{h}) , \tag{1.48}$$

where $\mathcal{H}$ is the Hilbert transform. This linear operator has the constant coefficient $\mathtt{m}_{\frac{1}{2}} \approx 1$ at the order $1/2$, 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 (11.39).

7. Reduction of the lower orders. In Section 12 we further diagonalize the linear operator in (1.48), reducing it to constant coefficients up to regularizing smoothing operators of very negative order $|D|^{-M}$. This step, based on standard pseudo-differential calculus, is not needed in [21], because the second order Melnikov conditions in [21] do not lose space derivatives. We apply an iterative sequence of pseudo-differential transformations that eliminate first the φ- and then the x-dependence of the diagonal symbols. The final system has the form

$$(h, \bar{h}) \mapsto \omega \cdot \partial_\varphi h + \mathrm{i} \mathtt{m}_{\frac{1}{2}} |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} h + \mathrm{i} r(D) h + \mathcal{T}_8(\varphi)(h, \bar{h}) \tag{1.49}$$

where the constant Fourier multiplier $r(\xi)$ is real, even $r(\xi) = r(-\xi)$, it satisfies (see (12.78))

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

and the variable coefficient operator $\mathcal{T}_8(\varphi)$ is regularizing and satisfies tame estimates, see more precisely (12.85). We also remark that the operator (1.49) is reversible and even, since all the previous transformations that we performed are reversibility preserving and even.

At this point the procedure of diagonalization of $\mathcal{L}$ up to smoothing operators is complete. Thus, in Section 13, restricting the operator (1.49) to $H_{\mathbb{S}^+}^{\perp}$, we obtain the reduction of $\mathcal{L}_\omega$ up to smoothing remainders. We are now ready to begin the KAM reduction procedure.

8. KAM reducibility. In order to decrease quadratically the size of the resulting perturbation $\mathcal{R}_0$ (see (14.4)) we apply the KAM diagonalization iterative scheme of Section 14, which converges because the operators

$$\langle D \rangle^{\mathtt{m}+\mathtt{b}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}, \quad \partial_{\varphi_i}^{s_0+\mathtt{b}} \langle D \rangle^{\mathtt{m}+\mathtt{b}} \mathcal{R}_0 \langle D \rangle^{\mathtt{m}+\mathtt{b}+1} , \ i = 1, \ldots, \nu , \tag{1.50}$$

satisfy tame estimates for some $\mathtt{b} := \mathtt{b}(\tau, k_0) \in \mathbb{N}$ and $\mathtt{m} := \mathtt{m}(k_0)$ that are large enough (independently of s), see Lemma 14.2. Such conditions hold under the assumption that M (the order of regularization of the remainder) is chosen large enough as in (14.8) (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. Actually, for the construction of the quasi-periodic solutions, it is sufficient to prove the almost-reducibility of the linearized operator, in the sense that the remainder $\mathcal{R}_n$ in Theorem 14.4 is not zero but it is of order $O(\varepsilon \gamma^{-2(M+1)} N_{n-1}^{-\mathtt{a}})$, which can be obtained imposing only the finitely many Diophantine conditions (14.41), (14.26).

The big difference of the KAM reducibility scheme of Section 14 with respect to the one developed in [21] is that the second order Melnikov non-resonance conditions that we impose are very weak, see (14.26), 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 (see the Fourier truncations $|\ell|, |j - j'| \leq N_{\mathtt{n}-1}$ in (14.26)).

After the above almost-diagonalization of the linearized operator we almost-invert it, by imposing the first order Melnikov non-resonance conditions in (14.92), see Lemma 14.9. 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 almost inverse of $\mathcal{L}_\omega$ which satisfies tame estimates, see Theorem 14.10.

At this point the proof of the Nash-Moser Theorem 4.1, given in Section 15, follows in a usual way, in the same setting of [21].

Notation. 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).

The notation $a \lesssim_{s,\alpha,M} b$ means that $a \leq C(s,\alpha,M)b$ for some constant $C(s,\alpha,M) > 0$ depending on the Sobolev index s and the constants α, M. 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 3) are considered as fixed constants. Similarly, the set $\mathbb{S}^+$ of tangential sites is 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.

We use the multi-index notation: if $k = (k_1, \ldots, k_{\nu+1}) \in \mathbb{N}^{\nu+1}$ we denote $|k| := k_1 + \ldots + k_{\nu+1}$ and $k! := k_1! \cdots k_{\nu+1}!$ and if $\lambda = (\lambda_1, \ldots, \lambda_{\nu+1}) \in \mathbb{R}^{\nu+1}$, we denote the derivative $\partial_\lambda^k := \partial_{\lambda_1}^{k_1} \ldots \partial_{\lambda_{\nu+1}}^{k_{\nu+1}}$ and $\lambda^k := \lambda_1^{k_1} \cdots \lambda_{\nu+1}^{k_{\nu+1}}$. Recalling 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 $\| \cdot \|_{s,F}^{k+1,\gamma}$.

Definition 2.1. (Whitney-Sobolev functions) *Let F be a closed subset of $\mathbb{R}^{\nu+1}$. Let $k \geq 0$ be an integer, $\gamma \in (0,1]$, and $s \geq s_0 > (\nu+1)/2$. 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 $u^{(j)} : F \to H^s_{(\varphi,x)}$, $j \in \mathbb{N}^\nu$, $0 \leq |j| \leq k$ 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*

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

then

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

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

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

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 $u^{(j)} = \partial_\lambda^j u$, $|j| \leq k$, 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 $u^{(j)}$, $|j| \geq 1$, are uniquely determined as the partial derivatives $u^{(j)} = \partial_\lambda^j u$, $|j| \leq k$, 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

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

Theorem B.2 and (B.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 (B.9))

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

By Lemma B.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

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

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*

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

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

Proof. See Appendix B. □

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*

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

(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*

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

Proof. See Appendix B. □

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*

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

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 $s \ge s_0$,

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

Let $\|\beta\|_{2s_0+k+2}^{k+1,\gamma} \le \delta(s_0,k)$ *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 B. □

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

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

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

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

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 $(\omega \cdot \partial_\varphi)^{-1}_{ext} u = (\omega \cdot \partial_\varphi)^{-1} u$ 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*

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

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

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

Proof. See Appendix B. □

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* $\mathrm{Lip}(k+1, F, s, \gamma)$ *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 $\|u\|^{k+1,\gamma}_{s_0,F} \leq C_0$, then, for all $s \geq s_0 > (d+1)/2$,*

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

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 $\|u\|^{k+1,\gamma}_{s_0,F} \leq C_0$.

Proof. See Appendix B. □

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

$$(Au)(\varphi, x) := (A(\varphi)u(\varphi, \cdot))(x). \tag{2.20}$$

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

$$\mathcal{R} \begin{pmatrix} \eta \\ \psi \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \eta \\ \psi \end{pmatrix} \tag{2.21}$$

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 as in (2.20) 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

$$u(\varphi, x) = \sum_{j\in\mathbb{Z}} u_j(\varphi) e^{\mathrm{i}jx} = \sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} u_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + jx)}, \tag{2.22}$$

is

$$Au(\varphi, x) = \sum_{j,j'\in\mathbb{Z}} A_j^{j'}(\varphi) u_{j'}(\varphi) e^{\mathrm{i}jx} = \sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \sum_{\ell'\in\mathbb{Z}^\nu, j'\in\mathbb{Z}} A_j^{j'}(\ell-\ell') u_{\ell',j'} e^{\mathrm{i}(\ell\cdot\varphi + jx)}. \tag{2.23}$$

We shall identify an operator A with the matrix $\big(A_j^{j'}(\ell-\ell')\big)_{j,j'\in\mathbb{Z},\ell,\ell'\in\mathbb{Z}^\nu}$, which is Töplitz with respect to the index ℓ. In this paper we always consider Töplitz operators as in (2.20), (2.23).

The matrix entries $A_j^{j'}(\ell-\ell')$ of a bounded operator $A : H^s \to H^s$ (as in (2.23)) satisfy

$$\sum_{\ell,j} |A_j^{j'}(\ell-\ell')|^2 \langle \ell, j\rangle^{2s} \le \|A\|^2_{\mathcal{L}(H^s)} \langle \ell', j'\rangle^{2s}, \quad \forall (\ell', j') \in \mathbb{Z}^{\nu+1}, \tag{2.24}$$

where $\|A\|_{\mathcal{L}(H^s)} := \sup\{\|Ah\|_s : \|h\|_s = 1\}$ is the operator norm (consider $h = e^{\mathrm{i}(\ell',j')\cdot(\varphi,x)}$).

Definition 2.7. *Given a linear operator A as in* (2.23) *we define the operator*

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

$$(\Pi_N A)_j^{j'}(\ell-\ell') := \begin{cases} A_j^{j'}(\ell-\ell') & \text{if} \quad \langle \ell-\ell', j-j'\rangle \le N \\ 0 & \text{otherwise}. \end{cases} \tag{2.25}$$

   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* $\langle \ell-\ell', j-j'\rangle^b A_j^{j'}(\ell-\ell')$.
4. $\partial_{\varphi_m} A(\varphi) = [\partial_{\varphi_m}, A] = \partial_{\varphi_m} \circ A - A \circ \partial_{\varphi_m}$ **(differentiated operator)** *whose matrix elements are* $\mathrm{i}(\ell_m - \ell'_m) A_j^{j'}(\ell-\ell')$.

Similarly the commutator $[\partial_x, A]$ is represented by the matrix with entries $\mathrm{i}(j-j')A_j^{j'}(\ell-\ell')$.

Given linear operators A, B as in (2.23) we have that (see Lemma 2.4 in [21])

$$\| \, |A+B| u\|_s \le \| \, |A| \, [\![u]\!] \, \|_s + \| \, |B| \, [\![u]\!] \, \|_s, \quad \| \, |AB| u\|_s \le \| \, |A||B| \, [\![u]\!] \, \|_s, \tag{2.26}$$

where, for a given a function $u(\varphi, x)$ expanded in Fourier series as in (2.22), we define the majorant function

$$[\![u]\!](\varphi, x) := \sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} |u_{\ell,j}| e^{\mathrm{i}(\ell\cdot\varphi + jx)}. \tag{2.27}$$

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

$$\|u\|_s = \|[\![u]\!]\|_s. \tag{2.28}$$

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 [35].

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*

$$\left|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)\right| \leq C_{\alpha,\beta}\langle\xi\rangle^{m-\beta}, \quad \forall \alpha,\beta\in\mathbb{N}. \tag{2.29}$$

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 (2.29)*, 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

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, \quad A_i \in OPS^m, \quad i=1,\dots,4 \tag{2.30}$$

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

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

Along the paper we consider φ-dependent 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 3. Note that $\partial_\lambda^k A=\mathrm{Op}(\partial_\lambda^k a)$, $\forall k\in\mathbb{N}^{\nu+1}$.

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

$$|D|^m := \mathrm{Op}\big(\chi(\xi)|\xi|^m\big), \tag{2.31}$$

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

$$\mathcal{H}(e^{\mathrm{i}jx}) := -\mathrm{i}\,\mathrm{sign}(j)e^{\mathrm{i}jx}, \ \forall j\neq 0, \quad \mathcal{H}(1):=0, \tag{2.32}$$

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

$$\pi_0 u := \frac{1}{2\pi}\int_{\mathbb{T}} u(x)\,dx, \tag{2.33}$$

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

$$\langle D\rangle^m := \pi_0 + |D|^m := \mathrm{Op}\left((1-\chi(\xi))+\chi(\xi)|\xi|^m\right), \quad \xi\in\mathbb{R}. \tag{2.34}$$

We now recall the pseudo-differential norm introduced in Definition 2.11 in [21] (inspired by Métivier [50], 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*

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

where

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

For a matrix of pseudo differential operators $A \in OPS^m$ as in (2.30)*, we define its pseudo differential norm*

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

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

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

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

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

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

$$|\mathrm{Op}\,(a)|_{0,s,\alpha}^{k_0,\gamma} = \sum_{|k|\le k_0} \gamma^{|k|} \sup_{\lambda\in\Lambda_0} \|\partial_\lambda^k a(\lambda)\|_s = \|a\|_{W^{k_0,\infty,\gamma}(\Lambda_0,H^s)} \sim_{k_0} \|a\|_s^{k_0,\gamma} , \quad \forall \alpha \in \mathbb{N} , \tag{2.39}$$

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

$$|\mathrm{Op}(g)|_{m,s,\alpha}^{k_0,\gamma} = |\mathrm{Op}(g)|_{m,0,\alpha}^{k_0,\gamma} \le C(m,\alpha,g,k_0) , \quad \forall s \ge 0 . \tag{2.40}$$

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

$$|\mathrm{Op}(\langle a\rangle_\varphi)|_{m,s,\alpha}^{k_0,\gamma} \lesssim |\mathrm{Op}(a)|_{m,s,\alpha}^{k_0,\gamma} , \quad |\mathrm{Op}(\langle a\rangle_{\varphi,x})|_{m,s,\alpha}^{k_0,\gamma} \lesssim |\mathrm{Op}(a)|_{m,0,\alpha}^{k_0,\gamma} , \quad \forall s \ge 0 . \tag{2.41}$$

The norm $|\ |_{0,s,0}$ controls the action of a pseudo-differential operator on the Sobolev spaces H^s, see Lemma 2.28. The norm $|\ |_{m,s,\alpha}^{k_0,\gamma}$ is closed under composition and satisfies tame estimates.

Composition. If $A = a(x,D) \in OPS^m$, $B = b(x,D) \in OPS^{m'}$ then the composition operator $AB := A \circ B = \sigma_{AB}(x,D)$ is a pseudo-differential operator in $OPS^{m+m'}$ whose symbol σ_{AB} has the following asymptotic expansion: for all $N \ge 1$,

$$\sigma_{AB}(x,\xi) = \sum_{\beta=0}^{N-1} \frac{1}{\mathrm{i}^\beta \beta!} \partial_\xi^\beta a(x,\xi)\,\partial_x^\beta b(x,\xi) + r_N(x,\xi) \qquad \text{where} \qquad r_N := r_{N,AB} \in S^{m+m'-N} , \tag{2.42}$$

and the remainder r_N has the explicit formula

$$r_N(x,\xi) := r_{N,AB}(x,\xi) := \frac{1}{\mathrm{i}^N (N-1)!}\int_0^1 (1-\tau)^{N-1} \sum_{j\in\mathbb{Z}} (\partial_\xi^N a)(x,\xi+\tau j)\widehat{(\partial_x^N b)}(j,\xi) e^{\mathrm{i}jx}\,d\tau . \tag{2.43}$$

We remind the following composition estimate proved in Lemma 2.13 in [21].

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 $\alpha \in \mathbb{N}$, $s \ge s_0$,*

$$|AB|_{m+m',s,\alpha}^{k_0,\gamma} \lesssim_{m,\alpha,k_0} C(s)|A|_{m,s,\alpha}^{k_0,\gamma}|B|_{m',s_0+\alpha+|m|,\alpha}^{k_0,\gamma} + C(s_0)|A|_{m,s_0,\alpha}^{k_0,\gamma}|B|_{m',s+\alpha+|m|,\alpha}^{k_0,\gamma} . \tag{2.44}$$

Moreover, for any integer $N \geq 1$, the remainder $R_N := \mathrm{Op}(r_N)$ in (2.42) *satisfies*

$$\begin{aligned} |R_N|^{k_0,\gamma}_{m+m'-N,s,\alpha} \lesssim_{m,N,\alpha,k_0} & C(s)|A|^{k_0,\gamma}_{m,s,N+\alpha}|B|^{k_0,\gamma}_{m',s_0+2N+|m|+\alpha,\alpha} \\ & + C(s_0)|A|^{k_0,\gamma}_{m,s_0,N+\alpha}|B|^{k_0,\gamma}_{m',s+2N+|m|+\alpha,\alpha}. \end{aligned} \tag{2.45}$$

Both (2.44)-(2.45) *hold with the constant $C(s_0)$ interchanged with $C(s)$.*

Analogous estimates hold if A and B are matrix operators of the form (2.30).

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

$$|A \circ g(D)|^{k_0,\gamma}_{m+m',s,\alpha} \lesssim_{k_0,\alpha} |A|^{k_0,\gamma}_{m,s,\alpha}|\mathrm{Op}(g)|^{k_0,\gamma}_{m',0,\alpha} \lesssim_{k_0,\alpha,m'} |A|^{k_0,\gamma}_{m,s,\alpha}. \tag{2.46}$$

By (2.42) 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

$$[A, B] = \mathrm{Op}(a \star b). \tag{2.47}$$

By (2.42) the symbol $a \star b \in S^{m+m'-1}$ admits the expansion

$$a \star b = -\mathrm{i}\{a, b\} + \mathtt{r}_2(a, b) \qquad \text{where} \qquad \{a, b\} := \partial_\xi a\, \partial_x b - \partial_x a\, \partial_\xi b \in S^{m+m'-1} \tag{2.48}$$

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

$$\mathtt{r}_2(a, b) := r_{2,AB} - r_{2,BA} \in S^{m+m'-2}. \tag{2.49}$$

By Lemma 2.10 we deduce the following corollary.

Lemma 2.11. (Commutator) *If $A = a(\lambda, \varphi, x, D) \in OPS^m$ and $B = b(\lambda, \varphi, x, D) \in OPS^{m'}$, $m, m' \in \mathbb{R}$, then the commutator $[A, B] := AB - BA \in OPS^{m+m'-1}$ satisfies*

$$\begin{aligned} |[A, B]|^{k_0,\gamma}_{m+m'-1,s,\alpha} \lesssim_{m,m',\alpha,k_0} & C(s)|A|^{k_0,\gamma}_{m,s+2+|m'|+\alpha,\alpha+1}|B|^{k_0,\gamma}_{m',s_0+2+|m|+\alpha,\alpha+1} \\ & + C(s_0)|A|^{k_0,\gamma}_{m,s_0+2+|m'|+\alpha,\alpha+1}|B|^{k_0,\gamma}_{m',s+2+|m|+\alpha,\alpha+1}. \end{aligned} \tag{2.50}$$

Proof. Use the expansion in (2.42) with $N = 1$ for both AB and BA, then use (2.45) and (2.37). $\square$

Given two linear operators A and B, we define inductively the operators $\mathrm{Ad}^n_A(B)$, $n \in \mathbb{N}$ in the following way: $\mathrm{Ad}_A(B) := [A, B]$ and $\mathrm{Ad}^{n+1}_A(B) := [A, \mathrm{Ad}^n_A(B)]$, $n \in \mathbb{N}$. Iterating the estimate (2.50), one deduces

$$\begin{aligned} |\mathrm{Ad}^n_A(B)|^{k_0,\gamma}_{nm+m'-n,s,\alpha} \lesssim_{m,m',s,\alpha,k_0} & (|A|^{k_0,\gamma}_{m,s_0+\mathtt{c}_n(m,m',\alpha),\alpha+n})^n|B|^{k_0,\gamma}_{m',s+\mathtt{c}_n(m,m',\alpha),\alpha+n} \\ & + (|A|^{k_0,\gamma}_{m,s_0+\mathtt{c}_n(m,m',\alpha),\alpha+n})^{n-1}|A|^{k_0,\gamma}_{m,s+\mathtt{c}_n(m,m',\alpha),\alpha+n}|B|^{k_0,\gamma}_{m',s_0+\mathtt{c}_n(m,m',\alpha),\alpha+n} \end{aligned} \tag{2.51}$$

for suitable constants $\mathtt{c}_n(m, m', \alpha) > 0$.

We remind the following estimate for the adjoint operator proved in Lemma 2.16 in [21].

Lemma 2.12. (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 L^2-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 (2.30).

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

Lemma 2.13. (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*

$$C(s_0, \alpha, k_0) |A|_{0, s_0+\alpha, \alpha}^{k_0, \gamma} \leq 1/2 \,, \tag{2.52}$$

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

$$|\Phi^{-1} - \mathrm{Id}|_{0,s,\alpha}^{k_0,\gamma} \leq C(s, \alpha, k_0) |A|_{0, s+\alpha, \alpha}^{k_0, \gamma} \,. \tag{2.53}$$

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 (2.30).

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

2.4 Integral operators and Hilbert transform

In this section we consider integral operators with a $\mathcal{C}^\infty$ kernel, which are the operators in $OPS^{-\infty}$. As in the previous section, they are k_0 times differentiable with respect to $\lambda := (\omega, \mathtt{h})$ in an open set $\Lambda_0 \subseteq \mathbb{R}^{\nu+1}$.

Lemma 2.14. *Let $K := K(\lambda, \cdot) \in \mathcal{C}^\infty(\mathbb{T}^\nu \times \mathbb{T} \times \mathbb{T})$. Then the integral operator*

$$(\mathcal{R}u)(\varphi, x) := \int_{\mathbb{T}} K(\lambda, \varphi, x, y) u(\varphi, y)\, dy \tag{2.54}$$

is in $OPS^{-\infty}$ and, for all $m, s, \alpha \in \mathbb{N}$, $|\mathcal{R}|_{-m,s,\alpha}^{k_0,\gamma} \leq C(m, s, \alpha, k_0) \|K\|_{\mathcal{C}^{s+m+\alpha}}^{k_0,\gamma}$.

Proof. See Lemma 2.32 in [21]. □

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

$$Pu(\varphi, x) := u(\varphi, x + p(\varphi, x)) \,. \tag{2.55}$$

Lemma 2.15. *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 $\|p\|_{2s_0+k_0+1}^{k_0,\gamma} \leq \delta$, then the integral operator $\mathcal{R}$ in (2.54) 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*

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

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

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

Proof. See Lemma 2.34 in [21]. □

We now recall some properties of the Hilbert transform $\mathcal{H}$ defined as a Fourier multiplier in (2.32). The commutator between $\mathcal{H}$ and the multiplication operator by a smooth function a is a regularizing operator in $OPS^{-\infty}$, as stated in Lemma 2.35 in [21] (see also Lemma B.5 in [6], Appendices H and I in [42]).

Lemma 2.16. *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 $m, s, \alpha \in \mathbb{N}$,*

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

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 [42] for similar statements).

Lemma 2.17. *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* (2.55)*. 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 $K(\lambda, \varphi, x, z) := -\frac{1}{\pi}\partial_z \log(1 + g(\lambda, \varphi, x, z))$ with

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

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

Lemma 2.18. *Let $g(\lambda, \varphi, \xi)$ be a family of Fourier multipliers with $\partial^k_\lambda g(\lambda, \varphi, \cdot) \in S^{-\infty}$, for all $k \in \mathbb{N}^{\nu+1}$, $|k| \leq k_0$. Then the operator $\mathrm{Op}(g)$ admits the integral representation*

$$\big[\mathrm{Op}(g)u\big](\varphi, x) = \int_{\mathbb{T}} K_g(\lambda, \varphi, x, y)u(\varphi, y)\, dy\,, \qquad K_g(\lambda, \varphi, x, y) := \frac{1}{2\pi}\sum_{j\in\mathbb{Z}} g(\lambda, \varphi, j)e^{\mathrm{i}j(x-y)}\,, \tag{2.56}$$

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

$$\|K_g\|^{k_0,\gamma}_{\mathcal{C}^s} \lesssim |\mathrm{Op}(g)|^{k_0,\gamma}_{-1,s+s_0,0} + |\mathrm{Op}(g)|^{k_0,\gamma}_{-s-s_0-1,0,0}\,. \tag{2.57}$$

Proof. The lemma follows by differentiating the explicit expression of the integral Kernel K_g in (2.56). □

2.5 Reversible, Even, Real operators

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

Definition 2.19. (Even operator) *A linear operator $A := A(\varphi)$ as in* (2.23) *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

$$A \text{ is even} \quad \Longleftrightarrow \quad A^{j'}_j(\varphi) + A^{-j'}_j(\varphi) = A^{j'}_{-j}(\varphi) + A^{-j'}_{-j}(\varphi)\,, \quad \forall j, j' \in \mathbb{Z},\ \varphi \in \mathbb{T}^\nu. \tag{2.58}$$

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

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

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 2.21. *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*
$$\tilde{a}(x,\xi) := \frac{1}{2}\big(a(x,\xi) + a(-x,-\xi)\big) \tag{2.59}$$
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 2.22. 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 (2.59) 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
$$\langle \tilde{a} \rangle_x(\xi) := \frac{1}{2\pi}\int_{\mathbb{T}} \tilde{a}(x,\xi)\,dx \qquad \text{satisfies} \qquad \langle \tilde{a} \rangle_x(-\xi) = \langle \tilde{a} \rangle_x(\xi)\,,$$
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 (2.21), 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
$$\mathcal{C} := \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \end{pmatrix}, \qquad \mathcal{C}^{-1} = \begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix}. \tag{2.60}$$
We get that the *real* operator $\mathcal{R}$ acting in the complex coordinates $(u,\bar{u}) = \mathcal{C}^{-1}(\eta,\psi)$ takes the form
$$\begin{aligned} &\mathbf{R} = \mathcal{C}^{-1}\mathcal{R}\mathcal{C} := \begin{pmatrix} \mathcal{R}_1 & \mathcal{R}_2 \\ \overline{\mathcal{R}_2} & \overline{\mathcal{R}_1} \end{pmatrix}, \\ &\mathcal{R}_1 := \frac{1}{2}\big\{(A+D) - \mathrm{i}(B-C)\big\}, \quad \mathcal{R}_2 := \frac{1}{2}\big\{(A-D) + \mathrm{i}(B+C)\big\} \end{aligned} \tag{2.61}$$
where the *conjugate* operator $\overline{A}$ is defined by
$$\overline{A}(u) := \overline{A(\bar{u})}\,. \tag{2.62}$$
We say that a matrix operator acting on the complex variables $(u,\bar{u})$ is REAL if it has the structure in (2.61) 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 2.23. *Let* $\mathbf{R}$ *be a real operator as in* (2.61). *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*

$$(\mathcal{R}_i)_j^{j'}(-\varphi)=-\overline{(\mathcal{R}_i)_{-j}^{-j'}(\varphi)}\quad\forall\varphi\in\mathbb{T}^\nu\,,\quad i.e.\ (\mathcal{R}_i)_j^{j'}(\ell)=-\overline{(\mathcal{R}_i)_{-j}^{-j'}(\ell)}\quad\forall\ell\in\mathbb{Z}^\nu\,. \tag{2.63}$$

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*

$$(\mathcal{R}_i)_j^{j'}(-\varphi)=\overline{(\mathcal{R}_i)_{-j}^{-j'}(\varphi)}\quad\forall\varphi\in\mathbb{T}^\nu\,,\quad i.e.\ (\mathcal{R}_i)_j^{j'}(\ell)=\overline{(\mathcal{R}_i)_{-j}^{-j'}(\ell)}\quad\forall\ell\in\mathbb{Z}^\nu\,. \tag{2.64}$$

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 family of linear operators as in (2.23), k_0 times differentiable with respect to λ in an open set $\Lambda_0\subset\mathbb{R}^{\nu+1}$.

Definition 2.24. **($\mathcal{D}^{k_0}$-σ-tame)** *Let* $\sigma\ge 0$. *A linear operator* $A:=A(\lambda)$ *as in* (2.20) *is* $\mathcal{D}^{k_0}$-σ-*tame if there exists a non-decreasing function* $[s_0,S]\to[0,+\infty)$, $s\mapsto\mathfrak{M}_A(s)$, *possibly with* $S=+\infty$, *such that for all* $s_0\le s\le S$, *for all* $u\in H^{s+\sigma}$

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

We say that $\mathfrak{M}_A(s)$ *is a* TAME CONSTANT *of the operator* A. *The constant* $\mathfrak{M}_A(s):=\mathfrak{M}_A(k_0,\sigma,s)$ *may also depend 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 (2.61)*)*

$$A=\begin{pmatrix}A_1 & A_2\\ \overline{A_2} & \overline{A_1}\end{pmatrix}, \tag{2.66}$$

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

Note that the tame constants $\mathfrak{M}_A(s)$ are not uniquely determined. Moreover, if $S<+\infty$, every linear operator A that is uniformly bounded in λ (together with its derivatives $\partial_\lambda^k A$) as an operator from $H^{s+\sigma}$ to H^s is $\mathcal{D}^{k_0}$-σ-tame. The relevance of Definition 2.24 is that, for the remainder operators which we shall obtain along the reducibility of the linearized operator in Sections 6-14, we are able to prove bounds of the tame constants $\mathfrak{M}_A(s)$ better than the trivial operator norm.

Remark 2.25. In Sections 6-14 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 $\mathfrak{M}_A(s)\le C(S)(1+\|\mathfrak{I}_0\|_{s+\mu}^{k_0,\gamma})$, for all $s_0\le s\le S$. □

An immediate consequence of (2.65) (with $k=0$, $s=s_0$) is that $\|A\|_{\mathcal{L}(H^{s_0+\sigma},H^{s_0})}\le 2\mathfrak{M}_A(s_0)$.

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 (2.23) we have, for all $|k|\le k_0$, $j'\in\mathbb{Z}$, $\ell'\in\mathbb{Z}^\nu$,

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

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

Lemma 2.26. (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 a tame constant satisfying*

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

The same estimate holds if A, B are matrix operators as in (2.66).

Proof. See Lemma 2.20 in [21]. □

We now discuss the action of a $\mathcal{D}^{k_0}$-σ-tame operator $A(\lambda)$ on a family of Sobolev functions $u(\lambda) \in H^s$.

Lemma 2.27. (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, we have*

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

The same estimate holds if A is a matrix operator as in (2.66).

Proof. See Lemma 2.22 in [21]. □

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

Lemma 2.28. *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 a tame constant satisfying*

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

As a consequence

$$\|Ah\|_s^{k_0,\gamma} \leq C(s_0, k_0)|A|_{0,s_0,0}^{k_0,\gamma}\|h\|_s^{k_0,\gamma} + C(s, k_0)|A|_{0,s,0}^{k_0,\gamma}\|h\|_{s_0}^{k_0,\gamma} . \tag{2.69}$$

The same statement holds if A is a matrix operator of the form (2.66).

Proof. See Lemma 2.21 in [21] for the proof of (2.68), then apply Lemma 2.27 to deduce (2.69). □

In view of the KAM reducibility scheme of Section 14, 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 2.29. ($\mathcal{D}^{k_0}$-modulo-tame) *A linear operator $A := A(\lambda)$ as in* (2.20) *is $\mathcal{D}^{k_0}$-modulo-tame if there exists a non-decreasing function $[s_0, S] \to [0, +\infty)$, $s \mapsto \mathfrak{M}_A^\sharp(s)$, such that 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 $s_0 \leq s \leq S$, $u \in H^s$,*

$$\sup_{|k|\leq k_0} \sup_{\lambda \in \Lambda_0} \gamma^{|k|} \| \, |\partial_\lambda^k A| u\|_s \leq \mathfrak{M}_A^\sharp(s_0)\|u\|_s + \mathfrak{M}_A^\sharp(s)\|u\|_{s_0} . \tag{2.70}$$

The constant $\mathfrak{M}_A^\sharp(s)$ is called a MODULO-TAME CONSTANT *of the operator A.*

For a matrix operator as in (2.66) *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 2.30. *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* (2.66).

Proof. See Lemma 2.24 in [21]. □

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

Lemma 2.31. **(Sum and composition)** *Let A, B be $\mathcal{D}^{k_0}$-modulo-tame operators with modulo-tame constants respectively $\mathfrak{M}^\sharp_A(s)$ and $\mathfrak{M}^\sharp_B(s)$. Then $A+B$ is $\mathcal{D}^{k_0}$-modulo-tame with a modulo-tame constant satisfying*

$$\mathfrak{M}^\sharp_{A+B}(s) \leq \mathfrak{M}^\sharp_A(s) + \mathfrak{M}^\sharp_B(s) \,. \tag{2.71}$$

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

$$\mathfrak{M}^\sharp_{AB}(s) \leq C(k_0)\big(\mathfrak{M}^\sharp_A(s)\mathfrak{M}^\sharp_B(s_0) + \mathfrak{M}^\sharp_A(s_0)\mathfrak{M}^\sharp_B(s)\big) \,. \tag{2.72}$$

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 a modulo-tame constant respectively $\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A}(s)$ and $\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} B}(s)$. Then $\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}(AB)$ is $\mathcal{D}^{k_0}$-modulo-tame with a modulo-tame constant satisfying

$$\begin{aligned}\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}(AB)}(s) \leq C(\mathtt{b})C(k_0)\Big(&\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A}(s)\mathfrak{M}^\sharp_B(s_0) + \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A}(s_0)\mathfrak{M}^\sharp_B(s) \\ &+ \mathfrak{M}^\sharp_A(s)\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} B}(s_0) + \mathfrak{M}^\sharp_A(s_0)\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} B}(s)\Big)\end{aligned} \tag{2.73}$$

for some constants $C(k_0), C(\mathtt{b}) \geq 1$. The same statement holds if A and B are matrix operators as in (2.66).

Proof. The estimates (2.71), (2.72) are proved in Lemma 2.25 of [21]. The bound (2.73) is proved as the estimate (2.76) of Lemma 2.25 in [21], replacing $\langle \partial_\varphi\rangle^{\mathtt{b}}$ (cf. Definition 2.3 in [21]) with $\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}$. $\square$

Iterating (2.72)-(2.73), one estimates $\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A^n}(s)$, and arguing as in Lemma 2.26 of [21] we deduce the following lemma.

Lemma 2.32. **(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*

$$4C(\mathtt{b})C(k_0)\mathfrak{M}^\sharp_A(s_0) \leq 1/2 \,. \tag{2.74}$$

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 they admit modulo-tame constants satisfying

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

The same statement holds if A is a matrix operator of the form (2.66).

Corollary 2.33. *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*

$$4C(\mathtt{b})C(k_0)\mathfrak{M}^\sharp_{\langle D\rangle^m A \langle D\rangle^{-m}}(s_0) \leq 1/2 \,. \tag{2.75}$$

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 and they admit modulo-tame constants satisfying

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

The same statement holds if A is a matrix operator of the form (2.66).

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 2.32, since the smallness condition (2.75) is (2.74) with $A = A_m$, and $\Phi_m^{-1} = \mathrm{Id} + \langle D\rangle^m \check{A}\langle D\rangle^{-m}$. $\square$

Lemma 2.34. **(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 a modulo-tame constant satisfying*

$$\mathfrak{M}^\sharp_{\Pi_N^\perp A}(s) \leq N^{-\mathtt{b}}\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} A}(s) \,, \quad \mathfrak{M}^\sharp_{\Pi_N^\perp A}(s) \leq \mathfrak{M}^\sharp_A(s) \,. \tag{2.76}$$

The same estimate holds when A is a matrix operator of the form (2.66).

Proof. As in Lemma 2.27 in [21], replacing $\langle \partial_\varphi \rangle^{\mathtt{b}}$ (cf. Definition 2.3 in [21]) with $\langle \partial_{\varphi,x} \rangle^{\mathtt{b}}$. $\square$

In order to verify that an operator is modulo-tame, we shall use the following Lemma. Notice that the right hand side of (2.77) below contains tame constants (not modulo-tame) of operators which control more space and time derivatives than $\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} A \langle D \rangle^{\mathtt{m}}$.

Lemma 2.35. *Let* $\mathtt{b}, \mathtt{m} \geq 0$. *Then*

$$\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} A \langle D \rangle^{\mathtt{m}}}(s) \lesssim_{s_0,\mathtt{b}} \mathfrak{M}_{\langle D \rangle^{\mathtt{m}+\mathtt{b}} A \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}}(s) + \max_{i=1,\ldots,\nu} \left\{ \mathfrak{M}_{\partial_{\varphi_i}^{s_0+\mathtt{b}} \langle D \rangle^{\mathtt{m}+\mathtt{b}} A \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}}(s) \right\}. \tag{2.77}$$

Proof. We denote by $\mathbb{M}(s, \mathtt{b})$ the right hand side in (2.77). For any $\alpha, \beta \in \mathbb{N}$, the matrix elements of the operator $\partial_{\varphi_i}^\alpha \langle D \rangle^\beta A \langle D \rangle^{\beta+1}$ are $\mathrm{i}^\alpha (\ell_i - \ell_i')^\alpha \langle j \rangle^\beta A_j^{j'}(\ell - \ell') \langle j' \rangle^{\beta+1}$. Then, by (2.67) with $\sigma = 0$, applied to the operators $\langle D \rangle^{\mathtt{m}+\mathtt{b}} A \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}$ and $\partial_{\varphi_i}^{s_0+\mathtt{b}} \langle D \rangle^{\mathtt{m}+\mathtt{b}} A \langle D \rangle^{\mathtt{m}+\mathtt{b}+1}$, we get, using the inequality $\langle \ell - \ell' \rangle^{2(s_0+\mathtt{b})} \lesssim_{\mathtt{b}} 1 + \max_{i=1,\ldots,\nu} |\ell_i - \ell_i'|^{2(s_0+\mathtt{b})}$, the bound

$$\begin{aligned} \gamma^{2|k|} \sum\nolimits_{\ell,j} \langle \ell, j \rangle^{2s} \langle \ell - \ell' \rangle^{2(s_0+\mathtt{b})} \langle j \rangle^{2(\mathtt{m}+\mathtt{b})} |\partial_\lambda^k A_j^{j'}(\ell - \ell')|^2 \langle j' \rangle^{2(\mathtt{m}+\mathtt{b}+1)} \\ \lesssim_{\mathtt{b}} \mathbb{M}^2(s_0, \mathtt{b}) \langle \ell', j' \rangle^{2s} + \mathbb{M}^2(s, \mathtt{b}) \langle \ell', j' \rangle^{2s_0}. \end{aligned} \tag{2.78}$$

For all $|k| \leq k_0$, by Cauchy-Schwarz inequality and using that

$$\langle \ell - \ell', j - j' \rangle^{\mathtt{b}} \lesssim_{\mathtt{b}} \langle \ell - \ell' \rangle^{\mathtt{b}} \langle j - j' \rangle^{\mathtt{b}} \lesssim_{\mathtt{b}} \langle \ell - \ell' \rangle^{\mathtt{b}} (\langle j \rangle^{\mathtt{b}} + \langle j' \rangle^{\mathtt{b}}) \lesssim_{\mathtt{b}} \langle \ell - \ell' \rangle^{\mathtt{b}} \langle j \rangle^{\mathtt{b}} \langle j' \rangle^{\mathtt{b}} \tag{2.79}$$

we get

$$\begin{aligned} \| |\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} \partial_\lambda^k A \langle D \rangle^{\mathtt{m}}| h \|_s^2 &\lesssim_{\mathtt{b}} \sum\nolimits_{\ell,j} \langle \ell, j \rangle^{2s} \Big(\sum\nolimits_{\ell',j'} |\langle \ell - \ell' \rangle^{\mathtt{b}} \langle j \rangle^{\mathtt{m}+\mathtt{b}} \partial_\lambda^k A_j^{j'}(\ell - \ell') \langle j' \rangle^{\mathtt{m}+\mathtt{b}}| |h_{\ell',j'}| \Big)^2 \\ &\lesssim_{\mathtt{b}} \sum\nolimits_{\ell,j} \langle \ell, j \rangle^{2s} \Big(\sum\nolimits_{\ell',j'} \langle \ell - \ell' \rangle^{s_0+\mathtt{b}} \langle j \rangle^{\mathtt{m}+\mathtt{b}} |\partial_\lambda^k A_j^{j'}(\ell - \ell')| \langle j' \rangle^{\mathtt{m}+\mathtt{b}+1} |h_{\ell',j'}| \frac{1}{\langle \ell - \ell' \rangle^{s_0} \langle j' \rangle} \Big)^2 \\ &\lesssim_{s_0,\mathtt{b}} \sum\nolimits_{\ell,j} \langle \ell, j \rangle^{2s} \sum\nolimits_{\ell',j'} \langle \ell - \ell' \rangle^{2(s_0+\mathtt{b})} \langle j \rangle^{2(\mathtt{m}+\mathtt{b})} |\partial_\lambda^k A_j^{j'}(\ell - \ell')|^2 \langle j' \rangle^{2(\mathtt{m}+\mathtt{b}+1)} |h_{\ell',j'}|^2 \\ &\lesssim_{s_0,\mathtt{b}} \sum\nolimits_{\ell',j'} |h_{\ell',j'}|^2 \sum\nolimits_{\ell,j} \langle \ell, j \rangle^{2s} \langle \ell - \ell' \rangle^{2(s_0+\mathtt{b})} \langle j \rangle^{2(\mathtt{m}+\mathtt{b})} |\partial_\lambda^k A_j^{j'}(\ell - \ell')|^2 \langle j' \rangle^{2(\mathtt{m}+\mathtt{b}+1)} \\ &\overset{(2.78)}{\lesssim_{s_0,\mathtt{b}}} \gamma^{-2|k|} \sum\nolimits_{\ell',j'} |h_{\ell',j'}|^2 \big(\mathbb{M}^2(s_0, \mathtt{b}) \langle \ell', j' \rangle^{2s} + \mathbb{M}^2(s, \mathtt{b}) \langle \ell', j' \rangle^{2s_0} \big) \\ &\lesssim_{s_0,\mathtt{b}} \gamma^{-2|k|} \big(\mathbb{M}^2(s_0, \mathtt{b}) \|h\|_s^2 + \mathbb{M}^2(s, \mathtt{b}) \|h\|_{s_0}^2 \big) \end{aligned} \tag{2.80}$$

using (2.28), whence the claimed statement follows. $\square$

Lemma 2.36. *Let* π_0 *be the projector defined in* (2.33) *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* (2.23) *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 a tame constant satisfying, for all* $s \geq s_0$,

$$\begin{aligned} \mathfrak{M}_{\langle D \rangle^{m_1} (\partial_\varphi^\beta (A \pi_0 B - \pi_0)) \langle D \rangle^{m_2}}(s) \lesssim_{m,s,\beta_0,k_0} \mathfrak{M}_{A-\mathrm{Id}}(s + \beta_0 + m_1) \big(1 + \mathfrak{M}_{B^*-\mathrm{Id}}(s_0 + m_2) \big) \\ + \mathfrak{M}_{B^*-\mathrm{Id}}(s + \beta_0 + m_2) \big(1 + \mathfrak{M}_{A-\mathrm{Id}}(s_0 + m_1) \big). \end{aligned} \tag{2.81}$$

The same estimate holds if A, B *are matrix operators of the form* (2.66) *and* π_0 *is replaced by the matrix operator* Π_0 *defined in* (10.2).

Proof. A direct calculation shows that $\langle D \rangle^{m_1} \big(A \pi_0 B - \pi_0 \big) \langle D \rangle^{m_2} [h] = g_1 (h, g_2)_{L^2_x} + (h, g_3)_{L^2_x}$ where g_1, g_2, g_3 are the functions defined by

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

The estimate (2.81) then follows by computing for any $\beta \in \mathbb{N}^\nu$, $k \in \mathbb{N}^{\nu+1}$ with $|\beta| \leq \beta_0$, $|k| \leq k_0$, the operator $\partial_\lambda^k \partial_\varphi^\beta \big(\langle D \rangle^{m_1} (A \pi_0 B - \pi_0) \langle D \rangle^{m_2} \big)$. $\square$

2.7 Tame estimates for the flow of pseudo-PDEs

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

$$\begin{cases} \partial_\tau u = \mathrm{i} a(\varphi, x)|D|^{\frac{1}{2}} u \\ u(0, x) = u_0(\varphi, x) , \end{cases} \qquad \varphi \in \mathbb{T}^\nu , \quad x \in \mathbb{T} , \tag{2.82}$$

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 $\Phi^\tau := \Phi(\tau) := \Phi(\lambda, \varphi, \tau)$ satisfies the equation

$$\begin{cases} \partial_\tau \Phi(\tau) = \mathrm{i} a(\varphi, x)|D|^{\frac{1}{2}} \Phi(\tau) \\ \Phi(0) = \mathrm{Id} . \end{cases} \tag{2.83}$$

Since the function $a(\varphi, x)$ is real valued, usual energy estimates imply that the flow $\Phi(\tau)$ is a bounded operator mapping H^s_x to H^s_x. In the Appendix of [21] it is proved that the flow $\Phi(\tau)$ satisfies also tame estimates in $H^s_{\varphi,x}$, see Proposition 2.37 below. Moreover, since (2.82) is an autonomous equation, its flow $\Phi(\varphi, \tau)$ satisfies the group property

$$\Phi(\varphi, \tau_1 + \tau_2) = \Phi(\varphi, \tau_1) \circ \Phi(\varphi, \tau_2) , \qquad \Phi(\varphi, \tau)^{-1} = \Phi(\varphi, -\tau) , \tag{2.84}$$

and, since $a(\lambda, \cdot)$ is k_0 times differentiable with respect to the parameter λ, then $\Phi(\lambda, \varphi, \tau)$ is k_0 times differentiable with respect to λ as well. Also notice that $\Phi^{-1}(\tau) = \Phi(-\tau) = \overline{\Phi}(\tau)$, 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

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

is even and reversibility preserving.

The operator $\partial_\lambda^k \partial_\varphi^\beta \Phi$ loses $|D_x|^{\frac{|\beta|+|k|}{2}}$ derivatives, which, in (2.86) 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 2.37. *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*

$$\|a\|_{2s_0+|m_1|+2} \leq \delta(s, m_1) , \quad \|a\|^{k_0,\gamma}_{s_0+\sigma(\beta_0,k_0,m_1,m_2)} \leq 1 , \tag{2.85}$$

then the flow $\Phi(\tau) := \Phi(\lambda, \varphi, \tau)$ of (2.82) satisfies

$$\sup_{\tau\in[0,1]} \|\langle D\rangle^{-m_1} \partial_\lambda^k \partial_\varphi^\beta \Phi(\tau) \langle D\rangle^{-m_2} h\|_s \lesssim_{s,\beta_0,k_0,m_1,m_2} \gamma^{-|k|} \Big(\|h\|_s + \|a\|^{k_0,\gamma}_{s+\sigma(|\beta|,|k|,m_1,m_2)} \|h\|_{s_0} \Big) \tag{2.86}$$

$$\sup_{\tau\in[0,1]} \|\partial_\lambda^k (\Phi(\tau) - \mathrm{Id}) h\|_s \lesssim_s \gamma^{-|k|} \big(\|a\|^{k_0,\gamma}_{s_0} \|h\|_{s+\frac{|k|+1}{2}} + \|a\|^{k_0,\gamma}_{s+s_0+k_0+\frac{3}{2}} \|h\|_{s_0+\frac{|k|+1}{2}} \big) . \tag{2.87}$$

Proof. The proof is similar to Propositions A.7, A.10 and A.11 in [21] with, in addition, the presence of $\langle D\rangle^{-m_1}$ and $\langle D\rangle^{-m_2}$ in (2.86). □

We consider also the dependence of the flow Φ with respect to the torus $i := i(\varphi)$ and the estimates for the adjoint operator Φ^*.

Lemma 2.38. *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 $\|a\|_{s_1+\sigma(\beta_0)} \leq \delta(s)$ with $\delta(s) > 0$ small enough, then the following estimate holds:*

$$\sup_{\tau\in[0,1]} \|\langle D\rangle^{-m_1} \partial_\varphi^\beta \Delta_{12} \Phi(\tau) \langle D\rangle^{-m_2} h\|_{s_1} \lesssim_{s_1} \|\Delta_{12} a\|_{s_1+\sigma(|\beta|)} \|h\|_{s_1} , \tag{2.88}$$

where $\Delta_{12}\Phi := \Phi(i_2) - \Phi(i_1)$ *and* $\Delta_{12}a := a(i_2) - a(i_1)$. *Moreover, for any* $k \in \mathbb{N}^{\nu+1}$, $|k| \leq k_0$, *for all* $s \geq s_0$,

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

Finally, 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. The proof is similar to Propositions A.13, A.14, A.17 and A.18 of [21]. □

3 Degenerate KAM theory

In this section we extend the degenerate KAM theory approach of [11] and [21].

Definition 3.1. *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 3.1 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

$$\vec{\omega}(\mathtt{h}) := (\omega_j(\mathtt{h}))_{j \in \mathbb{S}^+}\,, \quad \vec{\Omega}(\mathtt{h}) := (\Omega_j(\mathtt{h}))_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+} := (\omega_j(\mathtt{h}))_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+}\,, \tag{3.1}$$

where $\omega_j(\mathtt{h}) = \sqrt{j \tanh(\mathtt{h} j)}$ are defined in (1.19).

Lemma 3.2. (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 $\omega_{j_1}(\mathtt{h}), \ldots, \omega_{j_N}(\mathtt{h})$ with $1 \leq j_1 < j_2 < \ldots < j_N$ the function $[\mathtt{h}_1, \mathtt{h}_2] \ni \mathtt{h} \mapsto (\omega_{j_1}(\mathtt{h}), \ldots, \omega_{j_N}(\mathtt{h})) \in \mathbb{R}^N$ is non-degenerate according to Definition 3.1, namely that, for all $c \in \mathbb{R}^N \setminus \{0\}$, the function $\mathtt{h} \mapsto c_1\omega_{j_1}(\mathtt{h}) + \ldots + c_N\omega_{j_N}(\mathtt{h})$ is not identically zero on the interval $[\mathtt{h}_1, \mathtt{h}_2]$. We shall prove, equivalently, that the function

$$\mathtt{h} \mapsto c_1\omega_{j_1}(\mathtt{h}^4) + \ldots + c_N\omega_{j_N}(\mathtt{h}^4)$$

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

$$\tanh(z) = \sum_{n=0}^{\infty} T_n z^{2n+1} = z - \frac{z^3}{3} + \frac{2}{15}z^5 + \ldots\,,$$

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

$$\tanh(z^4) = \sum_{n=0}^{\infty} T_n z^{4(2n+1)} = z^4\Big(1 + \sum_{n \geq 1} T_n z^{8n}\Big) = z^4 - \frac{z^{12}}{3} + \frac{2}{15}z^{20} + \ldots$$

is convergent in $|z| < (\pi/2)^{1/4}$. Moreover $|\sum_{n\geq 1} T_n z^{8n}| < 1$ in a ball $|z| < \delta$, for some positive δ sufficiently small. As a consequence, also the real function

$$g_1(\mathtt{h}) := \omega_1(\mathtt{h}^4) = \sqrt{\tanh(\mathtt{h}^4)} = \mathtt{h}^2\Big(1 + \sum_{n\geq 1} T_n \mathtt{h}^{8n}\Big)^{1/2} = \sum_{n=0}^{+\infty} b_n \frac{\mathtt{h}^{8n+2}}{(8n+2)!} = \mathtt{h}^2 - \frac{\mathtt{h}^{10}}{6} + \ldots \tag{3.2}$$

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 $j \geq 1$,

$$g_j(\mathtt{h}) := \omega_j(\mathtt{h}^4) = \sqrt{j}\sqrt{\tanh(\mathtt{h}^4 j)} = \sqrt{j}\, g_1(j^{1/4}\mathtt{h}) = \sum_{n=0}^{+\infty} b_n j^{2n+1} \frac{\mathtt{h}^{8n+2}}{(8n+2)!}\,, \tag{3.3}$$

which is analytic on the whole $\mathbb{R}$, similarly as g_1.

Now fix N integers $1 \leq j_1 < j_2 < \ldots < j_N$. We prove that for all $c \in \mathbb{R}^N \setminus \{0\}$, the analytic function $c_1 g_{j_1}(\mathtt{h}) + \ldots + c_N g_{j_N}(\mathtt{h})$ is not identically zero. Suppose, by contradiction, that there exists $c \in \mathbb{R}^N \setminus \{0\}$ such that

$$c_1 g_{j_1}(\mathtt{h}) + \ldots + c_N g_{j_N}(\mathtt{h}) = 0 \quad \forall \mathtt{h} \in \mathbb{R}. \tag{3.4}$$

The real analytic function $g_1(\mathtt{h})$ defined in (3.2) 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 $b_{n_1}, \ldots, b_{n_N}$ with $n_1 < n_2 < \ldots < n_N$. We differentiate with respect to $\mathtt{h}$ the identity in (3.4) and we find

$$\begin{cases} c_1\big(D_{\mathtt{h}}^{(8n_1+2)} g_{j_1}\big)(\mathtt{h}) + \ldots + c_N\big(D_{\mathtt{h}}^{(8n_1+2)} g_{j_N}\big)(\mathtt{h}) = 0 \\ c_1\big(D_{\mathtt{h}}^{(8n_2+2)} g_{j_1}\big)(\mathtt{h}) + \ldots + c_N\big(D_{\mathtt{h}}^{(8n_2+2)} g_{j_N}\big)(\mathtt{h}) = 0 \\ \ldots\ldots\ldots \\ c_1\big(D_{\mathtt{h}}^{(8n_N+2)} g_{j_1}\big)(\mathtt{h}) + \ldots + c_N\big(D_{\mathtt{h}}^{(8n_N+2)} g_{j_N}\big)(\mathtt{h}) = 0\,. \end{cases}$$

As a consequence the $N \times N$-matrix

$$\mathcal{A}(\mathtt{h}) := \begin{pmatrix} \big(D_{\mathtt{h}}^{(8n_1+2)} g_{j_1}\big)(\mathtt{h}) & \ldots & \big(D_{h}^{(8n_1+2)} g_{j_N}\big)(\mathtt{h}) \\ \big(D_{\mathtt{h}}^{(8n_2+2)} g_{j_1}\big)(\mathtt{h}) & \ldots & \big(D_{h}^{(8n_2+2)} g_{j_N}\big)(\mathtt{h}) \\ \vdots & \ddots & \vdots \\ \big(D_{\mathtt{h}}^{(8n_N+2)} g_{j_1}\big)(\mathtt{h}) & \ldots & \big(D_{\mathtt{h}}^{(8n_N+2)} g_{j_N}\big)(\mathtt{h}) \end{pmatrix} \tag{3.5}$$

is singular for all $\mathtt{h} \in \mathbb{R}$, and so the analytic function

$$\det \mathcal{A}(\mathtt{h}) = 0 \quad \forall \mathtt{h} \in \mathbb{R} \tag{3.6}$$

is identically zero. In particular at $\mathtt{h} = 0$ we have $\det \mathcal{A}(0) = 0$. On the other hand, by (3.3) and the multi-linearity of the determinant we compute

$$\det \mathcal{A}(0) := \det \begin{pmatrix} b_{n_1} j_1^{2n_1+1} & \ldots & b_{n_1} j_N^{2n_1+1} \\ b_{n_2} j_1^{2n_2+1} & \ldots & b_{n_2} j_N^{2n_2+1} \\ \vdots & \ddots & \vdots \\ b_{n_N} j_1^{2n_N+1} & \ldots & b_{n_N} j_N^{2n_N+1} \end{pmatrix} = b_{n_1} \ldots b_{n_N} \det \begin{pmatrix} j_1^{2n_1+1} & \ldots & j_N^{2n_1+1} \\ j_1^{2n_2+1} & \ldots & j_N^{2n_2+1} \\ \vdots & \ddots & \vdots \\ j_1^{2n_N+1} & \ldots & j_N^{2n_N+1} \end{pmatrix}.$$

This is a generalized Vandermonde determinant. We use the following result.

Lemma 3.3. *Let $x_1, \ldots, x_N, \alpha_1, \ldots, \alpha_N$ be real numbers, with $0 < x_1 < \ldots < x_N$ and $\alpha_1 < \ldots < \alpha_N$. Then*

$$\det \begin{pmatrix} x_1^{\alpha_1} & \ldots & x_N^{\alpha_1} \\ \vdots & \ddots & \vdots \\ x_1^{\alpha_N} & \ldots & x_N^{\alpha_N} \end{pmatrix} > 0\,.$$

Proof. The lemma is proved in [57]. $\square$

Since $1 \leq j_1 < j_2 < \ldots < j_N$ and the exponents $\alpha_j := 2n_j + 1$ are increasing $\alpha_1 < \ldots < \alpha_N$, Lemma 3.3 implies that $\det \mathcal{A}(0) \neq 0$ (recall that $b_{n_1}, \ldots, b_{n_N} \neq 0$). This is a contradiction with (3.6).

In order to conclude the proof of Lemma 3.2 we have to prove that, for any N, for any $1 \leq j_1 < j_2 < \ldots < j_N$, the function $[\mathtt{h}_1, \mathtt{h}_2] \ni \mathtt{h} \mapsto (1, \omega_{j_1}(\mathtt{h}), \ldots, \omega_{j_N}(\mathtt{h})) \in \mathbb{R}^{N+1}$ is non-degenerate according to Definition 3.1, namely that, for all $c = (c_0, c_1, \ldots, c_N) \in \mathbb{R}^{N+1} \setminus \{0\}$, the function $\mathtt{h} \mapsto c_0 + c_1\omega_{j_1}(\mathtt{h}) + \ldots + c_N\omega_{j_N}(\mathtt{h})$ is not identically zero on the interval $[\mathtt{h}_1, \mathtt{h}_2]$. We shall prove, equivalently, that the real analytic function $\mathtt{h} \mapsto c_0 + c_1\omega_{j_1}(\mathtt{h}^4) + \ldots + c_N\omega_{j_N}(\mathtt{h}^4)$ is not identically zero on $\mathbb{R}$.

Suppose, by contradiction, that there exists $c = (c_0, c_1, \ldots, c_N) \in \mathbb{R}^{N+1} \setminus \{0\}$ such that

$$c_0 + c_1 g_{j_1}(\mathtt{h}) + \ldots + c_N g_{j_N}(\mathtt{h}) = 0 \quad \forall \mathtt{h} \in \mathbb{R}. \tag{3.7}$$

As above, we differentiate with respect to $\mathtt{h}$ the identity (3.7), and we find that the $(N+1) \times (N+1)$-matrix

$$\mathcal{B}(\mathtt{h}) := \begin{pmatrix} 1 & g_{j_1}(\mathtt{h}) & \ldots & g_{j_N}(\mathtt{h}) \\ 0 & (D_{\mathtt{h}}^{(8n_1+2)} g_{j_1})(\mathtt{h}) & \ldots & (D_{\mathtt{h}}^{(8n_1+2)} g_{j_N})(\mathtt{h}) \\ 0 & \vdots & \ddots & \vdots \\ 0 & (D_{\mathtt{h}}^{(8n_N+2)} g_{j_1})(\mathtt{h}) & \ldots & (D_{\mathtt{h}}^{(8n_N+2)} g_{j_N})(\mathtt{h}) \end{pmatrix} \tag{3.8}$$

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 (3.8) 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 (3.5). 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 (3.9)-(3.12) from the qualitative non-degeneracy condition of Lemma 3.2, the analyticity of the linear frequencies ω_j in (1.19), and their asymptotics (1.24).

Proposition 3.4. (Transversality) *There exist $k_0^* \in \mathbb{N}$, $\rho_0 > 0$ such that, for any $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$,*

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell\}| \geq \rho_0 \langle \ell \rangle, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \tag{3.9}$$

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h})\}| \geq \rho_0 \langle \ell \rangle, \quad \forall \ell \in \mathbb{Z}^\nu, \ j \in \mathbb{N}^+ \setminus \mathbb{S}^+, \tag{3.10}$$

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})\}| \geq \rho_0 \langle \ell \rangle, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+, \tag{3.11}$$

$$\max_{k \leq k_0^*} |\partial_{\mathtt{h}}^k \{\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) + \Omega_{j'}(\mathtt{h})\}| \geq \rho_0 \langle \ell \rangle, \quad \forall \ell \in \mathbb{Z}^\nu, \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \tag{3.12}$$

where $\vec{\omega}(\mathtt{h})$ and $\Omega_j(\mathtt{h})$ are defined in (3.1). *We recall the notation $\langle \ell \rangle := \max\{1, |\ell|\}$. We call (following [58]) ρ_0 the* "amount of non-degeneracy" *and k_0^* the* "index of non-degeneracy"*.*

Note that in (3.11) we exclude the index $\ell = 0$. In this case we directly have that, for all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$

$$|\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})| \geq c_1 |\sqrt{j} - \sqrt{j'}| = c_1 \frac{|j - j'|}{\sqrt{j} + \sqrt{j'}} \quad \forall j, j' \in \mathbb{N}^+, \qquad \text{where } c_1 := \sqrt{\tanh(\mathtt{h}_1)}\,. \tag{3.13}$$

Proof. All the inequalities (3.9)-(3.12) are proved by contradiction.

Proof of (3.9). 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

$$\forall k \in \mathbb{N}, \quad \forall m \geq k\,, \quad \Big| \partial_{\mathtt{h}}^k \vec{\omega}(\mathtt{h}_m) \cdot \frac{\ell_m}{\langle \ell_m \rangle} \Big| < \frac{1}{1+m}\,. \tag{3.14}$$

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 (3.14) 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 3.2.

Proof of (3.10). 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 (3.10) we can restrict to the indices $(\ell, j) \in \mathbb{Z}^\nu \times (\mathbb{N}^+ \setminus \mathbb{S}^+)$ satisfying

$$j^{\frac{1}{2}} < C_0|\ell| \,. \tag{3.15}$$

Arguing by contradiction (as for proving (3.9)), 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

$$\forall k \in \mathbb{N}, \quad \forall m \geq k \,, \quad \Big| \partial_{\mathtt{h}}^k \Big\{ \vec{\omega}(\mathtt{h}_m) \cdot \frac{\ell_m}{\langle\ell_m\rangle} + \frac{\Omega_{j_m}(\mathtt{h}_m)}{\langle\ell_m\rangle} \Big\} \Big| < \frac{1}{1+m} \,. \tag{3.16}$$

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

$$\mathtt{h}_{m_n} \to \bar{\mathtt{h}} \in [\mathtt{h}_1,\mathtt{h}_2] \,, \quad \frac{\ell_{m_n}}{\langle\ell_{m_n}\rangle} \to \bar c \in \mathbb{R}^\nu \,. \tag{3.17}$$

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 $|j_m| \leq C|\ell_m|^2$ for all m (see (3.15)), we have $j_{m_n} \to \bar\jmath$. Passing to the limit for $m_n \to +\infty$ in (3.16) we deduce, by (3.17), 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 3.2.

Case 2: (ℓ_{m_n}) is unbounded. Up to a subsequence, $|\ell_{m_n}| \to +\infty$. In this case the constant $\bar c$ in (3.17) is nonzero. Moreover, by (3.15), we also have that, up to a subsequence,

$$j_{m_n}^{\frac{1}{2}} \langle\ell_{m_n}\rangle^{-1} \to \bar d \in \mathbb{R}. \tag{3.18}$$

By (1.24), (3.17), (3.18), we get

$$\frac{\Omega_{j_{m_n}}(\mathtt{h}_{m_n})}{\langle\ell_{m_n}\rangle} = \frac{j_{m_n}^{\frac{1}{2}}}{\langle\ell_{m_n}\rangle} + \frac{r(j_{m_n}, \mathtt{h}_{m_n})}{\langle\ell_{m_n}\rangle} \to \bar d \,, \quad \partial_{\mathtt{h}}^k \frac{\Omega_{j_{m_n}}(\mathtt{h}_{m_n})}{\langle\ell_{m_n}\rangle} = \partial_{\mathtt{h}}^k \frac{r(j_{m_n}, \mathtt{h}_{m_n})}{\langle\ell_{m_n}\rangle} \to 0 \quad \forall k \geq 1 \tag{3.19}$$

as $m_n \to +\infty$. Passing to the limit in (3.16), by (3.19), (3.17) 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 3.2.

Proof of (3.11). For all $\mathtt{h} \in [\mathtt{h}_1,\mathtt{h}_2]$, by (3.13) and (1.19), we have

$$|\vec{\omega}(\mathtt{h}) \cdot \ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})| \geq |\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})| - |\vec{\omega}(\mathtt{h})||\ell| \geq c_1 |j^{\frac{1}{2}} - j'^{\frac{1}{2}}| - C|\ell| \geq \langle\ell\rangle$$

provided $|j^{\frac{1}{2}} - j'^{\frac{1}{2}}| \geq C_1\langle\ell\rangle$, for some $C_1 > 0$. Therefore in (3.11) we can restrict to the indices such that

$$|j^{\frac{1}{2}} - j'^{\frac{1}{2}}| < C_1\langle\ell\rangle \,. \tag{3.20}$$

Moreover in (3.11) we can also assume that $j \neq j'$, otherwise (3.11) reduces to (3.9), which is already proved. If, by contradiction, (3.11) 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

$$\forall k \in \mathbb{N} \,, \quad \forall m \geq k \,, \quad \Big| \partial_{\mathtt{h}}^k \Big\{ \vec{\omega}(\mathtt{h}_m) \cdot \frac{\ell_m}{\langle\ell_m\rangle} + \frac{\Omega_{j_m}(\mathtt{h}_m)}{\langle\ell_m\rangle} - \frac{\Omega_{j'_m}(\mathtt{h}_m)}{\langle\ell_m\rangle} \Big\} \Big| < \frac{1}{1+m} \,. \tag{3.21}$$

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

$$\mathtt{h}_{m_n} \to \bar{\mathtt{h}} \in [\mathtt{h}_1, \mathtt{h}_2]\,, \quad \ell_{m_n}/\langle \ell_{m_n}\rangle \to \bar{c} \in \mathbb{R}^\nu \setminus \{0\}\,. \tag{3.22}$$

We distinguish again two cases.

Case 1 : (ℓ_{m_n}) is unbounded. Using (3.20) we deduce that, up to a subsequence,

$$|j_m^{\frac12} - j_m'^{\frac12}|\langle \ell_m\rangle^{-1} \to \bar{d} \in \mathbb{R}\,. \tag{3.23}$$

Hence passing to the limit in (3.21) for $m_n \to +\infty$, we deduce by (3.22), (3.23), (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 3.2.

Case 2 : (ℓ_{m_n}) is bounded. By (3.20), we have that $|\sqrt{j_m} - \sqrt{j_m'}| \leq C$ 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 (3.21) 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 3.2.

(*ii*) $j_m, j_m' \to +\infty$. By (3.23) and (1.24), we deduce, passing to the limit in (3.21), that

$$\partial_{\mathtt{h}}^k \big\{ \vec{\omega}(\mathtt{h}) \cdot \bar{c} + \bar{d} \big\} = 0 \quad \forall k \in \mathbb{N}\,.$$

Hence the analytic function $\mathtt{h} \mapsto \vec{\omega}(\mathtt{h}) \cdot \bar{c} + \bar{d}$ is identically zero, which contradicts Lemma 3.2.

Proof of (3.12). The proof is similar to (3.10). 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 (3.10) we can restrict the analysis to the indices $(\ell, j, j') \in \mathbb{Z}^\nu \times (\mathbb{N}^+ \setminus \mathbb{S}^+)^2$ satisfying

$$\sqrt{j} + \sqrt{j'} < C_0|\ell|\,. \tag{3.24}$$

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

$$\forall k \in \mathbb{N}, \quad \forall m \geq k\,, \quad \Big| \partial_{\mathtt{h}}^k \Big\{ \vec{\omega}(\mathtt{h}_m) \cdot \frac{\ell_m}{\langle \ell_m\rangle} + \frac{\Omega_{j_m}(\mathtt{h}_m)}{\langle \ell_m\rangle} + \frac{\Omega_{j_m'}(\mathtt{h}_m)}{\langle \ell_m\rangle} \Big\} \Big| < \frac{1}{1+m}\,. \tag{3.25}$$

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

$$\mathtt{h}_{m_n} \to \bar{\mathtt{h}} \in [\mathtt{h}_1, \mathtt{h}_2]\,, \quad \frac{\ell_{m_n}}{\langle \ell_{m_n}\rangle} \to \bar{c} \in \mathbb{R}^\nu\,. \tag{3.26}$$

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 (3.24), 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 (3.25) we deduce, by (3.26), that

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

Therefore the analytic function $\mathtt{h} \mapsto \vec{\omega}(\mathtt{h}) \cdot \bar{c} + \langle \bar{\ell}\rangle^{-1}\Omega_{\bar{\jmath}}(\mathtt{h}) + \langle \bar{\ell}\rangle^{-1}\Omega_{\bar{\jmath}'}(\mathtt{h})$ is identically zero. This is in contradiction with Lemma 3.2.

Case 2: (ℓ_{m_n}) is unbounded. Up to a subsequence, $|\ell_{m_n}| \to +\infty$. In this case the constant $\bar{c}$ in (3.26) is nonzero. Moreover, by (3.24), we also have that, up to a subsequence,

$$(j_{m_n}^{\frac{1}{2}} + j_{m_n}'^{\frac{1}{2}})\langle \ell_{m_n}\rangle^{-1} \to \bar{d} \in \mathbb{R}\,. \tag{3.27}$$

By (1.24), (3.26), (3.27), passing to the limit as $m_n \to +\infty$ in (3.25) 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 3.2. □

4 Nash-Moser theorem and measure estimates

Rescaling $u \mapsto \varepsilon u$, we write (1.14) as 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

$$P_\varepsilon(u, \mathtt{h}) := P_\varepsilon(u) := \frac{1}{2\varepsilon}\int_{\mathbb{T}} \psi\,\big(G(\varepsilon\eta, \mathtt{h}) - G(0, \mathtt{h})\big)\psi\, dx\,. \tag{4.1}$$

We decompose the phase space

$$H^1_{0,\mathrm{even}} := \Big\{ u := (\eta, \psi) \in H^1_0(\mathbb{T}_x) \times \dot{H}^1(\mathbb{T}_x)\,, \quad u(x) = u(-x) \Big\} = H_{\mathbb{S}^+} \oplus H_{\mathbb{S}^+}^\perp \tag{4.2}$$

as the direct sum of the symplectic subspaces $H_{\mathbb{S}^+}$ and $H_{\mathbb{S}^+}^\perp$ defined in (1.25), we introduce action-angle variables on the tangential sites as in (1.33), and we leave unchanged the normal component z. The symplectic 2-form in (1.8) reads

$$\mathcal{W} := \Big(\sum\nolimits_{j\in\mathbb{S}^+} d\theta_j \wedge dI_j\Big) \oplus \mathcal{W}_{|H_{\mathbb{S}^+}^\perp} = d\Lambda, \tag{4.3}$$

where Λ is the Liouville 1-form

$$\Lambda_{(\theta,I,z)}[\widehat{\theta}, \widehat{I}, \widehat{z}] := -\sum_{j\in\mathbb{S}^+} I_j \widehat{\theta}_j - \frac{1}{2}\big(Jz\,,\,\widehat{z}\big)_{L^2}\,. \tag{4.4}$$

Hence the Hamiltonian system generated by $\mathcal{H}_\varepsilon$ transforms into the one generated by the Hamiltonian

$$H_\varepsilon := \mathcal{H}_\varepsilon \circ A = \varepsilon^{-2} H \circ \varepsilon A \tag{4.5}$$

where

$$A(\theta, I, z) := v(\theta, I) + z := \sum_{j\in\mathbb{S}^+} \sqrt{\frac{2}{\pi}} \begin{pmatrix} \omega_j^{1/2}\sqrt{\xi_j + I_j}\,\cos(\theta_j) \\ -\omega_j^{-1/2}\sqrt{\xi_j + I_j}\,\sin(\theta_j)\end{pmatrix} \cos(jx) + z\,. \tag{4.6}$$

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 $(\theta, I, z) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_{\mathbb{S}^+}^\perp$. The involution ρ in (1.11) becomes

$$\tilde{\rho} : (\theta, I, z) \mapsto (-\theta, I, \rho z)\,. \tag{4.7}$$

By (1.7) and (4.5) the Hamiltonian H_ε reads (up to a constant)

$$H_\varepsilon = \mathcal{N} + \varepsilon P\,, \quad \mathcal{N} := H_L \circ A = \vec{\omega}(\mathtt{h}) \cdot I + \frac{1}{2}(z, \Omega z)_{L^2}\,, \quad P := P_\varepsilon \circ A\,, \tag{4.8}$$

where $\vec{\omega}(\mathtt{h})$ is defined in (3.1) and Ω in (1.16). We look for an embedded invariant torus

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

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 (4.20)).

4.1 Nash-Moser theorem of hypothetical conjugation

For $\alpha \in \mathbb{R}^\nu$, we consider the modified Hamiltonian

$$H_\alpha := \mathcal{N}_\alpha + \varepsilon P\,, \quad \mathcal{N}_\alpha := \alpha \cdot I + \frac{1}{2}(z, \Omega z)_{L^2}\,. \tag{4.9}$$

We look for zeros of the nonlinear operator

$$\begin{aligned} \mathcal{F}(i,\alpha) &:= \mathcal{F}(i,\alpha,\omega,\mathtt{h},\varepsilon) := \omega \cdot \partial_\varphi i(\varphi) - X_{H_\alpha}(i(\varphi)) = \omega \cdot \partial_\varphi i(\varphi) - (X_{\mathcal{N}_\alpha} + \varepsilon X_P)(i(\varphi)) \\ &:= \begin{pmatrix} \omega \cdot \partial_\varphi \theta(\varphi) - \alpha - \varepsilon \partial_I P(i(\varphi)) \\ \omega \cdot \partial_\varphi I(\varphi) + \varepsilon \partial_\theta P(i(\varphi)) \\ \omega \cdot \partial_\varphi z(\varphi) - J(\Omega z(\varphi) + \varepsilon \nabla_z P(i(\varphi))) \end{pmatrix} \end{aligned} \tag{4.10}$$

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 (4.9) is reversible, i.e. $H_\alpha \circ \tilde{\rho} = H_\alpha$ where the involution $\tilde{\rho}$ is defined in (4.7). We look for reversible solutions of $\mathcal{F}(i,\alpha) = 0$, namely satisfying $\tilde{\rho} i(\varphi) = i(-\varphi)$ (see (4.7)), i.e.

$$\theta(-\varphi) = -\theta(\varphi)\,, \quad I(-\varphi) = I(\varphi)\,, \quad z(-\varphi) = (\rho z)(\varphi)\,. \tag{4.11}$$

The norm of the periodic component of the embedded torus

$$\mathfrak{I}(\varphi) := i(\varphi) - (\varphi, 0, 0) := (\Theta(\varphi), I(\varphi), z(\varphi))\,, \quad \Theta(\varphi) := \theta(\varphi) - \varphi\,, \tag{4.12}$$

is

$$\|\mathfrak{I}\|_s^{k_0,\gamma} := \|\Theta\|_{H^s_\varphi}^{k_0,\gamma} + \|I\|_{H^s_\varphi}^{k_0,\gamma} + \|z\|_s^{k_0,\gamma}\,, \tag{4.13}$$

where $\|z\|_s^{k_0,\gamma} = \|\eta\|_s^{k_0,\gamma} + \|\psi\|_s^{k_0,\gamma}$. We define

$$k_0 := k_0^* + 2, \tag{4.14}$$

where k_0^* is the index of non-degeneracy provided by Proposition 3.4, 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)

$$\mathtt{\Omega} := \Big\{ \omega \in \mathbb{R}^\nu \,:\, \mathrm{dist}\big(\omega, \vec{\omega}[\mathtt{h}_1, \mathtt{h}_2]\big) < \delta \Big\}, \quad \delta > 0 \tag{4.15}$$

of the unperturbed linear frequencies $\vec{\omega}[\mathtt{h}_1, \mathtt{h}_2]$ defined in (3.1).

Theorem 4.1. (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 $\mathbb{S}^+, k_0, \tau$ 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*

$$\alpha_\infty : \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] \mapsto \mathbb{R}^\nu\,, \quad \alpha_\infty(\omega, \mathtt{h}) = \omega + r_\varepsilon(\omega, \mathtt{h})\,, \quad \textit{with} \quad |r_\varepsilon|^{k_0,\gamma} \leq C\varepsilon\gamma^{-1}\,, \tag{4.16}$$

a family of embedded tori i_∞ defined for all $(\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ satisfying (4.11) *and*

$$\|i_\infty(\varphi) - (\varphi, 0, 0)\|_{s_0}^{k_0,\gamma} \leq C\varepsilon\gamma^{-1}\,, \tag{4.17}$$

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

$$\mu_j^\infty(\omega, \mathtt{h}) = \mathtt{m}_{\frac{1}{2}}^\infty(\omega, \mathtt{h})(j \tanh(\mathtt{h} j))^{\frac{1}{2}} + \mathfrak{r}_j^\infty(\omega, \mathtt{h}) \tag{4.18}$$

satisfying

$$|\mathtt{m}_{\frac{1}{2}}^\infty - 1|^{k_0,\gamma} \leq C\varepsilon\gamma^{-1}\,, \qquad \sup_{j \in \mathbb{N}^+ \setminus \mathbb{S}^+} j^{\frac{1}{2}} |\mathfrak{r}_j^\infty|^{k_0,\gamma} \leq C\varepsilon\gamma^{-\kappa_1} \tag{4.19}$$

such that for all $(\omega, \mathtt{h})$ in the Cantor like set

$$\begin{aligned}\mathcal{C}^{\gamma}_{\infty} := \Big\{ & (\omega, \mathtt{h}) \in \Omega \times [\mathtt{h}_1, \mathtt{h}_2] \, : \, |\omega \cdot \ell| \geq 8\gamma \langle \ell \rangle^{-\tau}, \, \forall \ell \in \mathbb{Z}^{\nu} \setminus \{0\} \, , \\ & |\omega \cdot \ell + \mu^{\infty}_j(\omega, \mathtt{h})| \geq 4\gamma j^{\frac{1}{2}} \langle \ell \rangle^{-\tau}, \, \forall \ell \in \mathbb{Z}^{\nu}, \, j \in \mathbb{N}^+ \setminus \mathbb{S}^+, \\ & |\omega \cdot \ell + \mu^{\infty}_j(\omega, \mathtt{h}) + \mu^{\infty}_{j'}(\omega, \mathtt{h})| \geq 4\gamma (j^{\frac{1}{2}} + j'^{\frac{1}{2}}) \langle \ell \rangle^{-\tau}, \ \forall \ell \in \mathbb{Z}^{\nu}, \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \, , \\ & |\omega \cdot \ell + \mu^{\infty}_j(\omega, \mathtt{h}) - \mu^{\infty}_{j'}(\omega, \mathtt{h})| \geq 4\gamma j^{-\mathtt{d}} j'^{-\mathtt{d}} \langle \ell \rangle^{-\tau}, \ \forall \ell \in \mathbb{Z}^{\nu}, \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+, \, (\ell, j, j') \neq (0, j, j) \Big\}\end{aligned} \tag{4.20}$$

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 4.1 is proved in Section 15. The very weak second Melnikov non-resonance conditions in (4.20) can be verified for most parameters if $\mathtt{d}$ is large enough, i.e. $\mathtt{d} > \frac{3}{4} k_0^*$, see Theorem 4.2 below.

4.2 Measure estimates

The aim is now to deduce Theorem 1.1 from Theorem 4.1.

By (4.16) the function $\alpha_\infty(\cdot, \mathtt{h})$ from Ω into the image $\alpha_\infty(\Omega, \mathtt{h})$ is invertible:

$$\beta = \alpha_\infty(\omega, \mathtt{h}) = \omega + r_\varepsilon(\omega, \mathtt{h}) \quad \Longleftrightarrow \quad \omega = \alpha_\infty^{-1}(\beta, \mathtt{h}) = \beta + \breve{r}_\varepsilon(\beta, \mathtt{h}) \quad \text{with} \quad |\breve{r}_\varepsilon|^{k_0, \gamma} \leq C\varepsilon\gamma^{-1} \, . \tag{4.21}$$

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 $\beta \in \alpha_\infty(\mathcal{C}^\gamma_\infty)$, Theorem 4.1 proves the existence of an embedded invariant torus filled by quasi-periodic solutions with Diophantine frequency $\omega = \alpha_\infty^{-1}(\beta, \mathtt{h})$ for the Hamiltonian

$$H_\beta = \beta \cdot I + \frac{1}{2}(z, \Omega z)_{L^2} + \varepsilon P \, .$$

Consider the curve of the unperturbed tangential 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 (1.37). In Theorem 4.2 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 $\mathcal{C}^\gamma_\infty$. Hence, for such values of $\mathtt{h}$ we have found an embedded invariant torus for the Hamiltonian H_ε in (4.8), 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 4.2. (Measure estimates) *Let*

$$\gamma = \varepsilon^a \, , \quad 0 < a < \min\{a_0, 1/(k_0 + \kappa_1)\} \, , \quad \tau > k_0^*(\nu + 4) \, , \quad \mathtt{d} > \frac{3k_0^*}{4} \, , \tag{4.22}$$

where k_0^* is the index of non-degeneracy given by Proposition 3.4 and $k_0 = k_0^* + 2$. Then the set

$$\mathcal{G}_\varepsilon := \big\{ \mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2] : \big(\alpha_\infty^{-1}(\vec{\omega}(\mathtt{h}), \mathtt{h}), \mathtt{h} \big) \in \mathcal{C}^\gamma_\infty \big\} \tag{4.23}$$

has a measure satisfying $|\mathcal{G}_\varepsilon| \to \mathtt{h}_2 - \mathtt{h}_1$ as $\varepsilon \to 0$.

The rest of this section is devoted to the proof of Theorem 4.2. By (4.21) the vector

$$\omega_\varepsilon(\mathtt{h}) := \alpha_\infty^{-1}(\vec{\omega}(\mathtt{h}), \mathtt{h}) = \vec{\omega}(\mathtt{h}) + \mathtt{r}_\varepsilon(\mathtt{h}) \, , \quad \mathtt{r}_\varepsilon(\mathtt{h}) := \breve{r}_\varepsilon(\vec{\omega}(\mathtt{h}), \mathtt{h}) \, , \tag{4.24}$$

satisfies

$$|\partial_{\mathtt{h}}^k \mathtt{r}_\varepsilon(\mathtt{h})| \leq C\varepsilon\gamma^{-k-1} \quad \forall 0 \leq k \leq k_0 \, . \tag{4.25}$$

We also denote, with a small abuse of notation, for all $j \in \mathbb{N}^+ \setminus \mathbb{S}^+$,

$$\mu_j^\infty(\mathtt{h}) := \mu_j^\infty(\omega_\varepsilon(\mathtt{h}), \mathtt{h}) := \mathtt{m}^\infty_{\frac{1}{2}}(\mathtt{h}) (j \tanh(\mathtt{h} j))^{\frac{1}{2}} + \mathfrak{r}_j^\infty(\mathtt{h}), \tag{4.26}$$

where

$$\mathtt{m}_{\frac{1}{2}}^{\infty}(\mathtt{h}) := \mathtt{m}_{\frac{1}{2}}^{\infty}(\omega_\varepsilon(\mathtt{h}),\mathtt{h}), \qquad \mathfrak{r}_j^{\infty}(\mathtt{h}) := \mathfrak{r}_j^{\infty}(\omega_\varepsilon(\mathtt{h}),\mathtt{h}). \tag{4.27}$$

By (4.19), (4.27) and (4.24)-(4.25), using that $\varepsilon\gamma^{-k_0-1} \le 1$ (which by (4.22) is satisfied for ε small), we get

$$|\partial_{\mathtt{h}}^k(\mathtt{m}_{\frac{1}{2}}^{\infty}(\mathtt{h}) - 1)| \le C\varepsilon\gamma^{-1-k}, \quad \sup_{j\in\mathbb{N}^+\setminus\mathbb{S}^+} j^{\frac{1}{2}}|\partial_{\mathtt{h}}^k \mathfrak{r}_j^{\infty}(\mathtt{h})| \le C\varepsilon\gamma^{-\kappa_1-k} \qquad \forall 0\le k\le k_0. \tag{4.28}$$

By (4.20), (4.24), (4.26), the Cantor set $\mathcal{G}_\varepsilon$ in (4.23) becomes

$$\begin{aligned}\mathcal{G}_\varepsilon = \Big\{ & \mathtt{h}\in[\mathtt{h}_1,\mathtt{h}_2] \colon |\omega_\varepsilon(\mathtt{h})\cdot\ell| \ge 8\gamma\langle\ell\rangle^{-\tau},\ \forall\ell\in\mathbb{Z}^\nu\setminus\{0\},\\ & |\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h})| \ge 4\gamma j^{\frac{1}{2}}\langle\ell\rangle^{-\tau},\ \forall\ell\in\mathbb{Z}^\nu,\ j\in\mathbb{N}^+\setminus\mathbb{S}^+,\\ & |\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h}) + \mu_{j'}^\infty(\mathtt{h})| \ge 4\gamma(j^{\frac{1}{2}} + j'^{\frac{1}{2}})\langle\ell\rangle^{-\tau},\ \forall\ell\in\mathbb{Z}^\nu,\ j,j'\in\mathbb{N}^+\setminus\mathbb{S}^+,\\ & |\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| \ge \frac{4\gamma\langle\ell\rangle^{-\tau}}{j^{\mathtt{d}}j'^{\mathtt{d}}}, \forall\ell\in\mathbb{Z}^\nu, j,j'\in\mathbb{N}^+\setminus\mathbb{S}^+, (\ell,j,j')\ne(0,j,j)\Big\}.\end{aligned} \tag{4.29}$$

We estimate the measure of the complementary set

$$\mathcal{G}_\varepsilon^c := [\mathtt{h}_1,\mathtt{h}_2]\setminus\mathcal{G}_\varepsilon := \Big(\bigcup_{\ell\ne0} R_\ell^{(0)}\Big) \cup \Big(\bigcup_{\ell,j} R_{\ell,j}^{(I)}\Big) \cup \Big(\bigcup_{\ell,j,j'} Q_{\ell jj'}^{(II)}\Big) \cup \Big(\bigcup_{(\ell,j,j')\ne(0,j,j)} R_{\ell jj'}^{(II)}\Big) \tag{4.30}$$

where the "resonant sets" are

$$R_\ell^{(0)} := \big\{\mathtt{h}\in[\mathtt{h}_1,\mathtt{h}_2] : |\omega_\varepsilon(\mathtt{h})\cdot\ell| < 8\gamma\langle\ell\rangle^{-\tau}\big\} \tag{4.31}$$

$$R_{\ell j}^{(I)} := \big\{\mathtt{h}\in[\mathtt{h}_1,\mathtt{h}_2] : |\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h})| < 4\gamma j^{\frac{1}{2}}\langle\ell\rangle^{-\tau}\big\} \tag{4.32}$$

$$Q_{\ell jj'}^{(II)} := \big\{\mathtt{h}\in[\mathtt{h}_1,\mathtt{h}_2] : |\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h}) + \mu_{j'}^\infty(\mathtt{h})| < 4\gamma(j^{\frac{1}{2}} + j'^{\frac{1}{2}})\langle\ell\rangle^{-\tau}\big\} \tag{4.33}$$

$$R_{\ell jj'}^{(II)} := \Big\{\mathtt{h}\in[\mathtt{h}_1,\mathtt{h}_2] : |\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| < \frac{4\gamma\langle\ell\rangle^{-\tau}}{j^{\mathtt{d}}j'^{\mathtt{d}}}\Big\} \tag{4.34}$$

with $j,j'\in\mathbb{N}^+\setminus\mathbb{S}^+$. We first note that some of these sets are empty.

Lemma 4.3. *For ε, $\gamma\in(0,\gamma_0)$ small, we have that*

1. *If $R_{\ell j}^{(I)}\ne\emptyset$ then $j^{\frac{1}{2}}\le C\langle\ell\rangle$.*
2. *If $R_{\ell jj'}^{(II)}\ne\emptyset$ then $|j^{\frac{1}{2}} - j'^{\frac{1}{2}}| \le C\langle\ell\rangle$. Moreover, $R_{0jj'}^{(II)} = \emptyset$, for all $j\ne j'$.*
3. *If $Q_{\ell jj'}^{(II)}\ne\emptyset$ then $j^{\frac{1}{2}} + j'^{\frac{1}{2}} \le C\langle\ell\rangle$.*

Proof. Let us consider the case of $R_{\ell jj'}^{(II)}$. If $R_{\ell jj'}^{(II)}\ne\emptyset$ there is $\mathtt{h}\in[\mathtt{h}_1,\mathtt{h}_2]$ such that

$$|\mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| < \frac{4\gamma\langle\ell\rangle^{-\tau}}{j^{\mathtt{d}}j'^{\mathtt{d}}} + |\omega_\varepsilon(\mathtt{h})\cdot\ell| \le C\langle\ell\rangle. \tag{4.35}$$

On the other hand, (4.26), (4.28), and (3.13) imply

$$|\mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| \ge \mathtt{m}_{\frac{1}{2}}^{\infty} c|\sqrt{j} - \sqrt{j'}| - C\varepsilon\gamma^{-\kappa_1} \ge \frac{c}{2}|\sqrt{j} - \sqrt{j'}| - 1. \tag{4.36}$$

Combining (4.35) and (4.36) we deduce $|j^{\frac{1}{2}} - j'^{\frac{1}{2}}| \le C\langle\ell\rangle$.

Next we prove that $R_{0jj'}^{(II)} = \emptyset$, $\forall j\ne j'$. Recalling (4.26), (4.28), and the definition $\Omega_j(\mathtt{h}) = \sqrt{j\tanh(\mathtt{h}j)}$, we have

$$\begin{aligned}|\mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})| &\ge \mathtt{m}_{\frac{1}{2}}^{\infty}(\mathtt{h})|\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})| - \frac{C\varepsilon\gamma^{-\kappa_1}}{j^{\frac{1}{2}}} - \frac{C\varepsilon\gamma^{-\kappa_1}}{(j')^{\frac{1}{2}}}\\ &\overset{(3.13)}{\ge} \frac{c}{2}|\sqrt{j} - \sqrt{j'}| - \frac{C\varepsilon\gamma^{-\kappa_1}}{j^{\frac{1}{2}}} - \frac{C\varepsilon\gamma^{-\kappa_1}}{(j')^{\frac{1}{2}}}.\end{aligned} \tag{4.37}$$

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

$$|\sqrt{j} - \sqrt{j'}| \geq c_0 \max\Big\{\frac{1}{\sqrt{j}}, \frac{1}{\sqrt{j'}}\Big\} \geq \frac{c_0}{2}\Big(\frac{1}{\sqrt{j}} + \frac{1}{\sqrt{j'}}\Big) \geq \frac{c_0}{j^{\frac14}(j')^{\frac14}} \qquad \forall j, j' \in \mathbb{N}^+, \ j \neq j'. \tag{4.38}$$

As a consequence of (4.37) and of the three inequalities in (4.38), 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 $Q^{(II)}_{\ell jj'}$ is elementary. $\square$

By Lemma 4.3, the last union in (4.30) becomes

$$\bigcup_{(\ell,j,j')\neq(0,j,j)} R^{(II)}_{\ell jj'} = \bigcup_{\substack{\ell\neq 0 \\ |\sqrt{j}-\sqrt{j'}|\leq C\langle\ell\rangle}} R^{(II)}_{\ell jj'}. \tag{4.39}$$

In order to estimate the measure of the sets (4.31)-(4.34) that are nonempty, the key point is to prove that the perturbed frequencies satisfy estimates similar to (3.9)-(3.12) in Proposition 3.4.

Lemma 4.4. (Perturbed transversality) *For ε small enough, for all $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$,*

$$\max_{k\leq k_0^*} |\partial_{\mathtt{h}}^k\{\omega_\varepsilon(\mathtt{h})\cdot\ell\}| \geq \frac{\rho_0}{2}\langle\ell\rangle \quad \forall \ell \in \mathbb{Z}^\nu\setminus\{0\}, \tag{4.40}$$

$$\max_{k\leq k_0^*} |\partial_{\mathtt{h}}^k\{\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h})\}| \geq \frac{\rho_0}{2}\langle\ell\rangle \quad \forall \ell \in \mathbb{Z}^\nu, \, j \in \mathbb{N}^+\setminus\mathbb{S}^+ : j^{\frac12} \leq C\langle\ell\rangle, \tag{4.41}$$

$$\max_{k\leq k_0^*} |\partial_{\mathtt{h}}^k\{\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})\}| \geq \frac{\rho_0}{2}\langle\ell\rangle \quad \forall \ell \in \mathbb{Z}^\nu\setminus\{0\}, \ j, j' \in \mathbb{N}^+\setminus\mathbb{S}^+ : |j^{\frac12} - j'^{\frac12}| \leq C\langle\ell\rangle, \tag{4.42}$$

$$\max_{k\leq k_0^*} |\partial_{\mathtt{h}}^k\{\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h}) + \mu_{j'}^\infty(\mathtt{h})\}| \geq \frac{\rho_0}{2}\langle\ell\rangle \quad \forall \ell \in \mathbb{Z}^\nu, \ j, j' \in \mathbb{N}^+\setminus\mathbb{S}^+ : j^{\frac12} + j'^{\frac12} \leq C\langle\ell\rangle, \tag{4.43}$$

where k_0^* is the index of non-degeneracy given by Proposition 3.4.

Proof. The most delicate estimate is (4.42). We split

$$\mu_j^\infty(\mathtt{h}) = \Omega_j(\mathtt{h}) + (\mu_j^\infty - \Omega_j)(\mathtt{h})$$

where $\Omega_j(\mathtt{h}) := j^{\frac12}(\tanh(j\mathtt{h}))^{\frac12}$. A direct calculation using (1.24) and (4.38) shows that, for $\mathtt{h} \in [\mathtt{h}_1, \mathtt{h}_2]$,

$$|\partial_{\mathtt{h}}^k\{\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})\}| \leq C_k|j^{\frac12} - j'^{\frac12}| \quad \forall k \geq 0. \tag{4.44}$$

Then, using (4.28), one has, for all $0 \leq k \leq k_0$,

$$\begin{aligned}
|\partial_{\mathtt{h}}^k\{(\mu_j^\infty - \mu_{j'}^\infty)(\mathtt{h}) - (\Omega_j - \Omega_{j'})(\mathtt{h})\}| &\leq |\partial_{\mathtt{h}}^k\{(\mathtt{m}_{\frac12}^\infty(\mathtt{h}) - 1)(\Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h}))\}| + |\partial_{\mathtt{h}}^k\mathfrak{r}_j^\infty(\mathtt{h})| + |\partial_{\mathtt{h}}^k\mathfrak{r}_{j'}^\infty(\mathtt{h})| \\
&\overset{(4.44)}{\leq} C_{k_0}\{\varepsilon\gamma^{-1-k}|j^{\frac12} - j'^{\frac12}| + \varepsilon\gamma^{-\kappa_1-k}(j^{-\frac12} + (j')^{-\frac12})\} \\
&\overset{(4.38)}{\leq} C'_{k_0}\varepsilon\gamma^{-\kappa_1-k}|j^{\frac12} - j'^{\frac12}|.
\end{aligned} \tag{4.45}$$

Recall that $k_0 = k_0^* + 2$ (see (4.14)). By (4.25) and (4.45), using $|j^{\frac12} - j'^{\frac12}| \leq C\langle\ell\rangle$, we get

$$\begin{aligned}
\max_{k\leq k_0^*} |\partial_{\mathtt{h}}^k\{\omega_\varepsilon(\mathtt{h})\cdot\ell + \mu_j^\infty(\mathtt{h}) - \mu_{j'}^\infty(\mathtt{h})\}| &\geq \max_{k\leq k_0^*} |\partial_{\mathtt{h}}^k\{\vec{\omega}(\mathtt{h})\cdot\ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})\}| - C\varepsilon\gamma^{-(1+k_0^*)}|\ell| \\
&\quad - C\varepsilon\gamma^{-(k_0^*+\kappa_1)}|j^{\frac12} - j'^{\frac12}| \\
&\geq \max_{k\leq k_0^*} |\partial_{\mathtt{h}}^k\{\vec{\omega}(\mathtt{h})\cdot\ell + \Omega_j(\mathtt{h}) - \Omega_{j'}(\mathtt{h})\}| - C\varepsilon\gamma^{-(k_0^*+\kappa_1)}\langle\ell\rangle \\
&\overset{(3.11)}{\geq} \rho_0\langle\ell\rangle - C\varepsilon\gamma^{-(k_0^*+\kappa_1)}\langle\ell\rangle \geq \rho_0\langle\ell\rangle/2
\end{aligned}$$

provided $\varepsilon\gamma^{-(k_0^*+\kappa_1)} \le \rho_0/(2C)$, which, by (4.22), is satisfied for ε small enough. $\square$

As an application of Rüssmann Theorem 17.1 in [58] we deduce the following

Lemma 4.5. (Estimates of the resonant sets) *The measure of the sets in* (4.31)-(4.34) *satisfies*

$$
\begin{aligned}
&|R^{(0)}_\ell| \lesssim \big(\gamma\langle\ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0^*}} \quad \forall \ell\neq 0\,, \qquad |R^{(I)}_{\ell j}| \lesssim \big(\gamma j^{\frac12}\langle\ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0^*}}\,,\\
&|R^{(II)}_{\ell jj'}| \lesssim \Big(\gamma\frac{\langle\ell\rangle^{-(\tau+1)}}{j^{\mathtt d}j'^{\mathtt d}}\Big)^{\frac{1}{k_0^*}} \quad \forall\ell\neq 0, \qquad |Q^{(II)}_{\ell jj'}| \lesssim \big(\gamma(j^{\frac12}+j'^{\frac12})\langle\ell\rangle^{-(\tau+1)}\big)^{\frac{1}{k_0^*}}\,.
\end{aligned}
$$

Proof. We prove the estimate of $R^{(II)}_{\ell jj'}$ in (4.34). The other cases are simpler. We write

$$
R^{(II)}_{\ell jj'} = \Big\{ \mathtt h\in[\mathtt h_1,\mathtt h_2] : |f_{\ell jj'}(\mathtt h)| < \frac{4\gamma}{\langle\ell\rangle^{\tau+1}j^{\mathtt d}j'^{\mathtt d}}\Big\}
$$

where $f_{\ell jj'}(\mathtt h) := (\omega_\varepsilon(\mathtt h)\cdot\ell + \mu_j^\infty(\mathtt h) - \mu_{j'}^\infty(\mathtt h))\langle\ell\rangle^{-1}$. By (4.39), we restrict to the case $|j^{\frac12}-j'^{\frac12}|\le C\langle\ell\rangle$ and $\ell\neq 0$. By (4.42),

$$
\max_{k\le k_0^*} |\partial_{\mathtt h}^k f_{\ell jj'}(\mathtt h)| \ge \rho_0/2\,, \quad \forall \mathtt h\in[\mathtt h_1,\mathtt h_2]\,.
$$

In addition, (4.24)-(4.28) and Lemma 4.3 imply that $\max_{k\le k_0}|\partial_{\mathtt h}^k f_{\ell jj'}(\mathtt h)|\le C$ for all $\mathtt h\in[\mathtt h_1,\mathtt h_2]$, provided $\varepsilon\gamma^{-(k_0+\kappa_1)}$ is small enough, namely, by (4.22), ε is small enough. In particular, $f_{\ell jj'}$ is of class $\mathcal C^{k_0-1} = \mathcal C^{k_0^*+1}$. Thus Theorem 17.1 in [58] applies, whence the lemma follows. $\square$

Proof of Theorem 4.2 completed. By Lemma 4.3 (in particular, recalling that $R^{(II)}_{\ell jj'}$ is empty for $\ell=0$ and $j\neq j'$, see (4.39)) and Lemma 4.5, the measure of the set $\mathcal G_\varepsilon^c$ in (4.30) is estimated by

$$
\begin{aligned}
|\mathcal G^c_\varepsilon| &\le \sum_{\ell\neq0}|R^{(0)}_\ell| + \sum_{\ell,j}|R^{(I)}_{\ell j}| + \sum_{(\ell,j,j')\neq(0,j,j)}|R^{(II)}_{\ell jj'}| + \sum_{\ell,j,j'}|Q^{(II)}_{\ell jj'}|\\
&\le \sum_{\ell\neq0}|R^{(0)}_\ell| + \sum_{j\le C\langle\ell\rangle^2}|R^{(I)}_{\ell j}| + \sum_{\substack{\ell\neq0\\ |\sqrt j-\sqrt{j'}|\le C\langle\ell\rangle}}|R^{(II)}_{\ell jj'}| + \sum_{j,j'\le C\langle\ell\rangle^2}|Q^{(II)}_{\ell jj'}|\\
&\lesssim \sum_\ell\Big(\frac{\gamma}{\langle\ell\rangle^{\tau+1}}\Big)^{\frac{1}{k_0^*}} + \sum_{j\le C\langle\ell\rangle^2}\Big(\frac{\gamma j^{\frac12}}{\langle\ell\rangle^{\tau+1}}\Big)^{\frac{1}{k_0^*}} + \sum_{|\sqrt j-\sqrt{j'}|\le C\langle\ell\rangle}\Big(\frac{\gamma}{\langle\ell\rangle^{\tau+1}j^{\mathtt d}j'^{\mathtt d}}\Big)^{\frac{1}{k_0^*}} + \sum_{j,j'\le C\langle\ell\rangle^2}\Big(\frac{\gamma(j^{\frac12}+j'^{\frac12})}{\langle\ell\rangle^{\tau+1}}\Big)^{\frac{1}{k_0^*}}\\
&\le C\gamma^{\frac{1}{k_0^*}}\Big\{\sum_{\ell\in\mathbb Z^\nu}\frac{1}{\langle\ell\rangle^{\frac{\tau}{k_0^*}-4}} + \sum_{|\sqrt j-\sqrt{j'}|\le C\langle\ell\rangle}\frac{1}{\langle\ell\rangle^{\frac{\tau+1}{k_0^*}}j^{\frac{\mathtt d}{k_0^*}}j'^{\frac{\mathtt d}{k_0^*}}}\Big\}\,.
\end{aligned}
\tag{4.46}
$$

The first series in (4.46) converges because $\frac{\tau}{k_0^*}-4>\nu$ by (4.22). For the second series in (4.46), 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

$$
\forall\ell,j,\quad \sum_{j'=j}^{j+p}\frac{1}{j'^{\frac{\mathtt d}{k_0^*}}} \le \sum_{j'=j}^{j+p}\frac{1}{j^{\frac{\mathtt d}{k_0^*}}} = \frac{p+1}{j^{\frac{\mathtt d}{k_0^*}}}\,, \quad p:=C^2\langle\ell\rangle^2+2C\sqrt j\langle\ell\rangle,
$$

the second series in (4.46) converges because $\frac{\tau+1}{k_0^*}-2>\nu$ and $2\frac{\mathtt d}{k_0^*}-\frac12>1$ by (4.22). By (4.46) 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 4.2 is concluded.

5 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* (see Theorem 5.6) of the linearized operator

$$d_{i,\alpha}\mathcal{F}(i_0,\alpha_0)[\widehat{\imath},\widehat{\alpha}] = \omega\cdot\partial_\varphi\widehat{\imath} - d_i X_{H_\alpha}(i_0(\varphi))[\widehat{\imath}] - (\widehat{\alpha},0,0)\,. \tag{5.1}$$

Note that $d_{i,\alpha}\mathcal{F}(i_0,\alpha_0) = d_{i,\alpha}\mathcal{F}(i_0)$ is independent of α_0, see (4.10) and recall that the perturbation P does not depend on α.

Since the linearized operator $d_i X_{H_\alpha}(i_0(\varphi))$ has the (θ,I,z)-components which are all coupled, it is particularly intricate to invert the operator (5.1). Then we implement the approach in [16], [8], [21] to reduce it, approximately, to a triangular form. We outline the steps of this strategy. The first observation is that, close to an invariant torus, there exists symplectic coordinates in which the linearized equations are a triangular system as in (1.27). We implement quantitatively this observation for any torus, which, in general, is non invariant. Thus we define the "error function"

$$Z(\varphi) := (Z_1,Z_2,Z_3)(\varphi) := \mathcal{F}(i_0,\alpha_0)(\varphi) = \omega\cdot\partial_\varphi i_0(\varphi) - X_{H_{\alpha_0}}(i_0(\varphi))\,. \tag{5.2}$$

If $Z=0$ then the torus i_0 is invariant for $X_{H_{\alpha_0}}$; in general, we say that i_0 is "approximately invariant", up to order $O(Z)$. Given a torus $i_0(\varphi)=(\theta_0(\varphi),I_0(\varphi),z_0(\varphi))$ satisfying (5.6) (condition which is satisfied by the approximate solutions obtained by the Nash-Moser iteration of Section 15), we first construct an isotropic torus $i_\delta(\varphi)=(\theta_0(\varphi),I_\delta(\varphi),z_0(\varphi))$ which is close to i_0, see Lemma 5.3. Note that, by (5.14), $\mathcal{F}(i_\delta,\alpha_0)$ is also $O(Z)$. Since i_δ is isotropic, the diffeomorphism $(\phi,y,w)\mapsto G_\delta(\phi,y,w)$ defined in (5.16) is symplectic. In these coordinates, the torus i_δ reads $(\phi,0,0)$, and the transformed Hamiltonian system becomes (5.19), where, by Lemma 5.4, the terms $\partial_\phi K_{00}, K_{10}-\omega, K_{01}$ are $O(Z)$. Thus, neglecting such terms, the problem of finding an approximate inverse of the linearized operator $d_{i,\alpha}\mathcal{F}(i_0,\alpha_0)$ is reduced to the task of inverting the operator $\mathbb{D}$ in (5.34). We solve system (5.35) in a *triangular* way. First we solve the equation for the y-component of system (5.35), simply by inverting the differential operator $\omega\cdot\partial_\varphi$, see (5.37) and recall that ω is Diophantine. Then in (5.38) we solve the equation for the w-component, thanks to the almost invertibility of the operator $\mathcal{L}_\omega$ in (5.26), which is proved in Theorem 14.10 and stated in this section as assumption (5.29)-(5.33). Finally the equation (5.39) for the ϕ-component is solved in (5.43), by modifying the counterterms according to (5.42) and by inverting $\omega\cdot\partial_\varphi$. In conclusion, in Theorem 5.6 we estimate quantitatively how the conjugation of $\mathbb{D}$ with the differential of G_δ (see (5.45)) is an almost approximate inverse of the linearized operator $d_{i,\alpha}\mathcal{F}(i_0,\alpha_0)$.

First of all, we state some preliminary estimates for the composition operator induced by the Hamiltonian vector field $X_P = (\partial_I P, -\partial_\theta P, J\nabla_z P)$ in (4.10).

Lemma 5.1. (Estimates of the perturbation P) *Let $\mathfrak{I}(\varphi)$ in* (4.12) *satisfy $\|\mathfrak{I}\|^{k_0,\gamma}_{3s_0+2k_0+5}\leq 1$. Then the following estimates hold:*

$$\|X_P(i)\|_s^{k_0,\gamma} \lesssim_s 1 + \|\mathfrak{I}\|^{k_0,\gamma}_{s+2s_0+2k_0+3}\,, \tag{5.3}$$

and for all $\widehat{\imath} := (\widehat{\theta},\widehat{I},\widehat{z})$

$$\|d_i X_P(i)[\widehat{\imath}]\|_s^{k_0,\gamma} \lesssim_s \|\widehat{\imath}\|^{k_0,\gamma}_{s+1} + \|\mathfrak{I}\|^{k_0,\gamma}_{s+2s_0+2k_0+4}\|\widehat{\imath}\|^{k_0,\gamma}_{s_0+1}\,, \tag{5.4}$$

$$\|d_i^2 X_P(i)[\widehat{\imath},\widehat{\imath}]\|_s^{k_0,\gamma} \lesssim_s \|\widehat{\imath}\|^{k_0,\gamma}_{s+1}\|\widehat{\imath}\|^{k_0,\gamma}_{s_0+1} + \|\mathfrak{I}\|^{k_0,\gamma}_{s+2s_0+2k_0+5}(\|\widehat{\imath}\|^{k_0,\gamma}_{s_0+1})^2\,. \tag{5.5}$$

Proof. The proof is the same as the one of Lemma 5.1 in [21], using also the estimates on the Dirichlet Neumann operator in Proposition A.1. □

Along this section we assume the following hypothesis, which is verified by the approximate solutions obtained at each step of the Nash-Moser Theorem 15.1.

- 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 $\mu := \mu(\tau,\nu)>0$, $\gamma\in(0,1)$,

$$\|\mathfrak{I}_0\|^{k_0,\gamma}_{s_0+\mu} + |\alpha_0-\omega|^{k_0,\gamma} \leq C\varepsilon\gamma^{-1}\,. \tag{5.6}$$

For some $\kappa := \kappa(\tau,\nu)>0$, we shall always assume the smallness condition $\varepsilon\gamma^{-\kappa}\ll 1$.

We now implement the symplectic procedure to reduce $d_{i,\alpha}\mathcal{F}(i_0, \alpha_0)$ approximately to a triangular form. An invariant torus i_0 with Diophantine flow is isotropic (see [30],[16]), namely the pull-back 1-form $i_0^*\Lambda$ is closed, where Λ is the 1-form in (4.4). This is equivalent to say that the 2-form $i_0^*\mathcal{W} = i_0^* d\Lambda = di_0^*\Lambda = 0$. For an approximately invariant torus i_0 the 1-form $i_0^*\Lambda$ is only "approximately closed": we consider

$$i_0^*\Lambda = \sum_{k=1}^{\nu} a_k(\varphi) d\varphi_k\,, \quad a_k(\varphi) := -\big([\partial_\varphi \theta_0(\varphi)]^T I_0(\varphi)\big)_k - \frac{1}{2}(\partial_{\varphi_k} z_0(\varphi), J z_0(\varphi))_{L^2(\mathbb{T}_x)} \tag{5.7}$$

and we show that

$$i_0^*\mathcal{W} = d\, i_0^*\Lambda = \sum_{1 \le k < j \le \nu} A_{kj}(\varphi) d\varphi_k \wedge d\varphi_j\,, \quad A_{kj}(\varphi) := \partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi)\,, \tag{5.8}$$

is of order $O(Z)$, see Lemma 5.2. By (4.10), (5.3), (5.6), the error function Z defined in (5.2) is estimated in terms of the approximate torus as

$$\|Z\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1} + \|\mathfrak{I}_0\|_{s+2}^{k_0,\gamma}. \tag{5.9}$$

Lemma 5.2. *Assume that ω belongs to $\mathtt{DC}(\gamma, \tau)$ defined in (2.13). Then the coefficients A_{kj} in (5.8) satisfy*

$$\|A_{kj}\|_s^{k_0,\gamma} \lesssim_s \gamma^{-1}\big(\|Z\|_{s+\tau(k_0+1)+k_0+1}^{k_0,\gamma} + \|Z\|_{s_0+1}^{k_0,\gamma} \|\mathfrak{I}_0\|_{s+\tau(k_0+1)+k_0+1}^{k_0,\gamma}\big)\,. \tag{5.10}$$

Proof. The A_{kj} satisfy the identity $\omega \cdot \partial_\varphi A_{kj} = \mathcal{W}\big(\partial_\varphi Z(\varphi)\underline{e}_k, \partial_\varphi i_0(\varphi)\underline{e}_j\big) + \mathcal{W}\big(\partial_\varphi i_0(\varphi)\underline{e}_k, \partial_\varphi Z(\varphi)\underline{e}_j\big)$ where $\underline{e}_k$ denotes the k-th versor of $\mathbb{R}^\nu$, see [16], Lemma 5. Then (5.10) follows by (5.6) and Lemma 2.5. ☐

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 $\Delta_\varphi := \sum_{k=1}^{\nu} \partial_{\varphi_k}^2$.

Lemma 5.3. (Isotropic torus) *The torus $i_\delta(\varphi) := (\theta_0(\varphi), I_\delta(\varphi), z_0(\varphi))$ defined by*

$$I_\delta := I_0 + [\partial_\varphi \theta_0(\varphi)]^{-T} \rho(\varphi)\,, \qquad \rho_j(\varphi) := \Delta_\varphi^{-1} \sum_{k=1}^{\nu} \partial_{\varphi_j} A_{kj}(\varphi) \tag{5.11}$$

is isotropic. There is $\sigma := \sigma(\nu, \tau, k_0)$ such that

$$\|I_\delta - I_0\|_s^{k_0,\gamma} \le \|I_0\|_{s+1}^{k_0,\gamma} \tag{5.12}$$

$$\|I_\delta - I_0\|_s^{k_0,\gamma} \lesssim_s \gamma^{-1}\big(\|Z\|_{s+\sigma}^{k_0,\gamma} + \|Z\|_{s_0+\sigma}^{k_0,\gamma} \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \tag{5.13}$$

$$\|\mathcal{F}(i_\delta, \alpha_0)\|_s^{k_0,\gamma} \lesssim_s \|Z\|_{s+\sigma}^{k_0,\gamma} + \|Z\|_{s_0+\sigma}^{k_0,\gamma} \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \tag{5.14}$$

$$\|d_i[i_\delta][\widehat{\imath}]\|_s^{k_0,\gamma} \lesssim_s \|\widehat{\imath}\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \|\widehat{\imath}\|_{s_0}^{k_0,\gamma}\,. \tag{5.15}$$

We denote by $\sigma := \sigma(\nu, \tau, k_0)$ possibly different (larger) "loss of derivatives" constants.

Proof. The Lemma follows as in [8] by (5.4) and (5.7)-(5.10). ☐

In order to find an approximate inverse of the linearized operator $d_{i,\alpha}\mathcal{F}(i_\delta)$, we introduce the symplectic diffeomorpshim $G_\delta : (\phi, y, w) \to (\theta, I, z)$ of the phase space $\mathbb{T}^\nu \times \mathbb{R}^\nu \times H_{\mathbb{S}^+}^\perp$ defined by

$$\begin{pmatrix} \theta \\ I \\ z \end{pmatrix} := G_\delta \begin{pmatrix} \phi \\ y \\ w \end{pmatrix} := \begin{pmatrix} \theta_0(\phi) \\ I_\delta(\phi) + [\partial_\phi \theta_0(\phi)]^{-T} y - \big[(\partial_\theta \tilde{z}_0)(\theta_0(\phi))\big]^T J w \\ z_0(\phi) + w \end{pmatrix} \tag{5.16}$$

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 5.3). In the new coordinates, i_δ is the trivial embedded torus $(\phi, y, w) = (\phi, 0, 0)$. Under the symplectic change of variables G_δ the Hamiltonian vector field X_{H_α} (the Hamiltonian H_α is defined in (4.9)) changes into

$$X_{K_\alpha} = (DG_\delta)^{-1} X_{H_\alpha} \circ G_\delta \qquad \text{where} \qquad K_\alpha := H_\alpha \circ G_\delta\,. \tag{5.17}$$

By (4.11) the transformation G_δ is also reversibility preserving and so K_α is reversible, $K_\alpha \circ \tilde{\rho} = K_\alpha$.

The Taylor expansion of K_α at the trivial torus $(\phi, 0, 0)$ is

$$
\begin{aligned}
K_\alpha(\phi, y, w) = K_{00}(\phi, \alpha) + K_{10}(\phi, \alpha) \cdot y + (K_{01}(\phi, \alpha), w)_{L^2(\mathbb{T}_x)} + \frac{1}{2} K_{20}(\phi) y \cdot y \\
+ \big(K_{11}(\phi) y, w\big)_{L^2(\mathbb{T}_x)} + \frac{1}{2}\big(K_{02}(\phi) w, w\big)_{L^2(\mathbb{T}_x)} + K_{\geq 3}(\phi, y, w)
\end{aligned}
\tag{5.18}
$$

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 (4.9) and (5.16), the only Taylor coefficients that depend on α are K_{00}, K_{10}, K_{01}.

The Hamilton equations associated to (5.18) are

$$
\begin{cases}
\dot\phi = K_{10}(\phi, \alpha) + K_{20}(\phi) y + K_{11}^T(\phi) w + \partial_y K_{\geq 3}(\phi, y, w) \\
\dot y = \partial_\phi K_{00}(\phi, \alpha) - [\partial_\phi K_{10}(\phi, \alpha)]^T y - [\partial_\phi K_{01}(\phi, \alpha)]^T w \\
\qquad - \partial_\phi\big(\tfrac{1}{2} K_{20}(\phi) y \cdot y + (K_{11}(\phi) y, w)_{L^2(\mathbb{T}_x)} + \tfrac{1}{2}(K_{02}(\phi) w, w)_{L^2(\mathbb{T}_x)} + K_{\geq 3}(\phi, y, w)\big) \\
\dot w = J\big(K_{01}(\phi, \alpha) + K_{11}(\phi) y + K_{02}(\phi) w + \nabla_w K_{\geq 3}(\phi, y, w)\big)
\end{cases}
\tag{5.19}
$$

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 $\mathbb{R}^\nu$,

$$
K_{11}^T(\phi) w = \sum\nolimits_{k=1}^{\nu} \big(K_{11}^T(\phi) w \cdot \underline{e}_k\big) \underline{e}_k = \sum\nolimits_{k=1}^{\nu} \big(w, K_{11}(\phi) \underline{e}_k\big)_{L^2(\mathbb{T}_x)} \underline{e}_k \in \mathbb{R}^\nu .
\tag{5.20}
$$

The coefficients K_{00}, K_{10}, K_{01} in the Taylor expansion (5.18) vanish on an exact solution (i.e. $Z = 0$).

Lemma 5.4. *We have*

$$
\begin{aligned}
&\|\partial_\phi K_{00}(\cdot, \alpha_0)\|_s^{k_0,\gamma} + \|K_{10}(\cdot, \alpha_0) - \omega\|_s^{k_0,\gamma} + \|K_{01}(\cdot, \alpha_0)\|_s^{k_0,\gamma} \lesssim_s \|Z\|_{s+\sigma}^{k_0,\gamma} + \|Z\|_{s_0+\sigma}^{k_0,\gamma} \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} . \\
&\|\partial_\alpha K_{00}\|_s^{k_0,\gamma} + \|\partial_\alpha K_{10} - \mathrm{Id}\|_s^{k_0,\gamma} + \|\partial_\alpha K_{01}\|_s^{k_0,\gamma} \lesssim_s \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} , \quad \|K_{20}\|_s^{k_0,\gamma} \lesssim_s \varepsilon\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) , \\
&\|K_{11} y\|_s^{k_0,\gamma} \lesssim_s \varepsilon\big(\|y\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \|y\|_{s_0}^{k_0,\gamma}\big) , \quad \|K_{11}^T w\|_s^{k_0,\gamma} \lesssim_s \varepsilon\big(\|w\|_{s+2}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \|w\|_{s_0+2}^{k_0,\gamma}\big) .
\end{aligned}
\tag{5.21}
$$

Proof. The lemma follows as in [16], [8], [21] by (5.3), (5.6), (5.12), (5.13), (5.14), (5.20). □

Under the linear change of variables

$$
DG_\delta(\varphi, 0, 0) \begin{pmatrix} \widehat\phi \\ \widehat y \\ \widehat w \end{pmatrix} := \begin{pmatrix} \partial_\phi \theta_0(\varphi) & 0 & 0 \\ \partial_\phi I_\delta(\varphi) & [\partial_\phi \theta_0(\varphi)]^{-T} & -[(\partial_\theta \tilde z_0)(\theta_0(\varphi))]^T J \\ \partial_\phi z_0(\varphi) & 0 & I \end{pmatrix} \begin{pmatrix} \widehat\phi \\ \widehat y \\ \widehat w \end{pmatrix}
\tag{5.22}
$$

the linearized operator $d_{i,\alpha}\mathcal{F}(i_\delta)$ is approximately transformed (see the proof of Theorem 5.6) into the one obtained when one linearizes the Hamiltonian system (5.19) at $(\phi, y, w) = (\varphi, 0, 0)$, differentiating also in α at α_0, and changing $\partial_t \rightsquigarrow \omega \cdot \partial_\varphi$, namely

$$
\begin{pmatrix} \widehat\phi \\ \widehat y \\ \widehat w \\ \widehat\alpha \end{pmatrix} \mapsto \begin{pmatrix} \omega \cdot \partial_\varphi \widehat\phi - \partial_\phi K_{10}(\varphi)[\widehat\phi] - \partial_\alpha K_{10}(\varphi)[\widehat\alpha] - K_{20}(\varphi) \widehat y - K_{11}^T(\varphi) \widehat w \\ \omega \cdot \partial_\varphi \widehat y + \partial_{\phi\phi} K_{00}(\varphi)[\widehat\phi] + \partial_\phi \partial_\alpha K_{00}(\varphi)[\widehat\alpha] + [\partial_\phi K_{10}(\varphi)]^T \widehat y + [\partial_\phi K_{01}(\varphi)]^T \widehat w \\ \omega \cdot \partial_\varphi \widehat w - J\{\partial_\phi K_{01}(\varphi)[\widehat\phi] + \partial_\alpha K_{01}(\varphi)[\widehat\alpha] + K_{11}(\varphi) \widehat y + K_{02}(\varphi) \widehat w\} \end{pmatrix} .
\tag{5.23}
$$

As in [8], by (5.22), (5.6), (5.12), the induced composition operator satisfies: for all $\widehat\imath := (\widehat\phi, \widehat y, \widehat w)$

$$
\|DG_\delta(\varphi, 0, 0)[\widehat\imath]\|_s^{k_0,\gamma} + \|DG_\delta(\varphi, 0, 0)^{-1}[\widehat\imath]\|_s^{k_0,\gamma} \lesssim_s \|\widehat\imath\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \|\widehat\imath\|_{s_0}^{k_0,\gamma} ,
\tag{5.24}
$$

$$
\|D^2 G_\delta(\varphi, 0, 0)[\widehat\imath_1, \widehat\imath_2]\|_s^{k_0,\gamma} \lesssim_s \|\widehat\imath_1\|_s^{k_0,\gamma} \|\widehat\imath_2\|_{s_0}^{k_0,\gamma} + \|\widehat\imath_1\|_{s_0}^{k_0,\gamma} \|\widehat\imath_2\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \|\widehat\imath_1\|_{s_0}^{k_0,\gamma} \|\widehat\imath_2\|_{s_0}^{k_0,\gamma} .
\tag{5.25}
$$

In order to construct an "almost-approximate" inverse of (5.23) we need that

$$\mathcal{L}_\omega := \Pi_{\mathbb{S}^+}^\perp \big(\omega \cdot \partial_\varphi - JK_{02}(\varphi)\big)_{|H_{\mathbb{S}^+}^\perp} \tag{5.26}$$

is "almost-invertible" up to remainders of size $O(N_{n-1}^{-\mathtt{a}})$ (see precisely (5.30)) where

$$N_n := K_n^p\,, \quad \forall n \geq 0\,, \tag{5.27}$$

and

$$K_n := K_0^{\chi^n}\,, \quad \chi := 3/2 \tag{5.28}$$

are the scales used in the nonlinear Nash-Moser iteration in Section 15. The almost invertibility of $\mathcal{L}_\omega$ is proved in Theorem 14.10 as the conclusion of the analysis of Sections 6-14, and it is stated here as an assumption (to avoid the involved definition of the set $\mathtt{\Lambda}_o$). Let $H_\perp^s(\mathbb{T}^{\nu+1}) := H^s(\mathbb{T}^{\nu+1}) \cap H_{\mathbb{S}^+}^\perp$ and recall that the phase space contains only functions even in x, see (4.2).

- **Almost-invertibility of $\mathcal{L}_\omega$.** There exists a subset $\mathtt{\Lambda}_o \subset \mathtt{DC}(\gamma,\tau)\times[\mathtt{h}_1,\mathtt{h}_2]$ such that, for all $(\omega,\mathtt{h}) \in \mathtt{\Lambda}_o$ the operator $\mathcal{L}_\omega$ in (5.26) may be decomposed as

$$\mathcal{L}_\omega = \mathcal{L}_\omega^< + \mathcal{R}_\omega + \mathcal{R}_\omega^\perp \tag{5.29}$$

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

$$\|\mathcal{R}_\omega h\|_s^{k_0,\gamma} \lesssim_S \varepsilon\gamma^{-2(M+1)} N_{n-1}^{-\mathtt{a}} \big(\|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,, \tag{5.30}$$

$$\|\mathcal{R}_\omega^\perp h\|_{s_0}^{k_0,\gamma} \lesssim_S K_n^{-b}\big(\|h\|_{s_0+b+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s_0+\mu(\mathtt{b})+\sigma+b}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,, \qquad \forall b > 0\,, \tag{5.31}$$

$$\|\mathcal{R}_\omega^\perp h\|_s^{k_0,\gamma} \lesssim_S \|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\,. \tag{5.32}$$

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

$$\|(\mathcal{L}_\omega^<)^{-1} g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\big(\|g\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|g\|_{s_0+\sigma}^{k_0,\gamma}\big)\,. \tag{5.33}$$

In order to find an almost-approximate inverse of the linear operator in (5.23) (and so of $d_{i,\alpha}\mathcal{F}(i_\delta)$), it is sufficient to invert the operator

$$\mathbb{D}[\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha}] := \begin{pmatrix} \omega\cdot\partial_\varphi\widehat{\phi} - \partial_\alpha K_{10}(\varphi)[\widehat{\alpha}] - K_{20}(\varphi)\widehat{y} - K_{11}^T(\varphi)\widehat{w} \\ \omega\cdot\partial_\varphi\widehat{y} + \partial_\phi\partial_\alpha K_{00}(\varphi)[\widehat{\alpha}] \\ (\mathcal{L}_\omega^<)\widehat{w} - J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}] - JK_{11}(\varphi)\widehat{y} \end{pmatrix} \tag{5.34}$$

obtained by neglecting in (5.23) the terms $\partial_\phi K_{10}$, $\partial_{\phi\phi}K_{00}$, $\partial_\phi K_{00}$, $\partial_\phi K_{01}$, which are $O(Z)$ by Lemma 5.4, and the small remainders $\mathcal{R}_\omega$, $\mathcal{R}_\omega^\perp$ appearing in (5.29). We look for an inverse of $\mathbb{D}$ by solving the system

$$\mathbb{D}[\widehat{\phi},\widehat{y},\widehat{w},\widehat{\alpha}] = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \tag{5.35}$$

where (g_1,g_2,g_3) satisfy the reversibility property

$$g_1(\varphi) = g_1(-\varphi)\,, \quad g_2(\varphi) = -g_2(-\varphi)\,, \quad g_3(\varphi) = -(\rho g_3)(-\varphi)\,. \tag{5.36}$$

We first consider the second equation in (5.35), 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

$$\widehat{y} := (\omega\cdot\partial_\varphi)^{-1}\big(g_2 - \partial_\alpha\partial_\phi K_{00}(\varphi)[\widehat{\alpha}]\big)\,. \tag{5.37}$$

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 (5.33), has a solution

$$\widehat{w} := (\mathcal{L}_\omega^<)^{-1}\big(g_3 + JK_{11}(\varphi)\widehat{y} + J\partial_\alpha K_{01}(\varphi)[\widehat{\alpha}]\big) . \tag{5.38}$$

Finally, we solve the first equation in (5.35), which, substituting (5.37), (5.38), becomes

$$\omega \cdot \partial_\varphi \widehat{\phi} = g_1 + M_1(\varphi)[\widehat{\alpha}] + M_2(\varphi)g_2 + M_3(\varphi)g_3 , \tag{5.39}$$

where

$$M_1(\varphi) := \partial_\alpha K_{10}(\varphi) - M_2(\varphi)\partial_\alpha\partial_\phi K_{00}(\varphi) + M_3(\varphi)J\partial_\alpha K_{01}(\varphi) , \tag{5.40}$$

$$M_2(\varphi) := K_{20}(\varphi)[\omega\cdot\partial_\varphi]^{-1} + K_{11}^T(\varphi)(\mathcal{L}_\omega^<)^{-1}JK_{11}(\varphi)[\omega\cdot\partial_\varphi]^{-1} , \quad M_3(\varphi) := K_{11}^T(\varphi)(\mathcal{L}_\omega^<)^{-1} . \tag{5.41}$$

In order to solve equation (5.39) we have to choose $\widehat{\alpha}$ such that the right hand side has zero average. By Lemma 5.4, (5.6), 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

$$\widehat{\alpha} := -\langle M_1\rangle^{-1}(\langle g_1\rangle + \langle M_2 g_2\rangle + \langle M_3 g_3\rangle) . \tag{5.42}$$

With this choice of $\widehat{\alpha}$, equation (5.39) has the solution

$$\widehat{\phi} := (\omega\cdot\partial_\varphi)^{-1}\big(g_1 + M_1(\varphi)[\widehat{\alpha}] + M_2(\varphi)g_2 + M_3(\varphi)g_3\big) . \tag{5.43}$$

In conclusion, we have obtained a solution $(\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\alpha})$ of the linear system (5.35).

Proposition 5.5. *Assume* (5.6) *(with $\mu = \mu(\mathtt{b}) + \sigma$) and* (5.33). *Then, for all $(\omega, \mathtt{h}) \in \Lambda_o$, for all $g := (g_1, g_2, g_3)$ even in x and satisfying* (5.36), *system* (5.35) *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* (5.43), (5.37), (5.38), (5.42), *which satisfies* (4.11) *and for any $s_0 \le s \le S$*

$$\|\mathbb{D}^{-1}g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\big(\|g\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\sigma}^{k_0,\gamma}\|g\|_{s_0+\sigma}^{k_0,\gamma}\big). \tag{5.44}$$

Proof. The lemma follows by (5.38), (5.40), (5.41), (5.42), (5.43), Lemma 5.4, (5.33), (5.6). □

Finally we prove that the operator

$$\mathbf{T}_0 := \mathbf{T}_0(i_0) := (D\widetilde{G}_\delta)(\varphi, 0, 0) \circ \mathbb{D}^{-1} \circ (DG_\delta)(\varphi, 0, 0)^{-1} \tag{5.45}$$

is an almost-approximate right inverse for $d_{i,\alpha}\mathcal{F}(i_0)$ where $\widetilde{G}_\delta(\phi, y, w, \alpha) := \big(G_\delta(\phi, y, w), \alpha\big)$ is the identity on the α-component. We denote the norm $\|(\phi, y, w, \alpha)\|_s^{k_0,\gamma} := \max\{\|(\phi, y, w)\|_s^{k_0,\gamma},\ |\alpha|^{k_0,\gamma}\}$.

Theorem 5.6. (Almost-approximate inverse) *Assume the inversion assumption* (5.29)-(5.33). *Then, there exists $\bar{\sigma} := \bar{\sigma}(\tau, \nu, k_0) > 0$ such that, if* (5.6) *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* (5.36), *the operator $\mathbf{T}_0$ defined in* (5.45) *satisfies, for all $s_0 \le s \le S$,*

$$\|\mathbf{T}_0 g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\big(\|g\|_{s+\bar{\sigma}}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar{\sigma}}^{k_0,\gamma}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\big) . \tag{5.46}$$

Moreover $\mathbf{T}_0$ is an almost-approximate inverse of $d_{i,\alpha}\mathcal{F}(i_0)$, namely

$$d_{i,\alpha}\mathcal{F}(i_0) \circ \mathbf{T}_0 - \mathrm{Id} = \mathcal{P}(i_0) + \mathcal{P}_\omega(i_0) + \mathcal{P}_\omega^\perp(i_0) \tag{5.47}$$

where, for all $s_0 \le s \le S$,

$$\begin{aligned}\|\mathcal{P}g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\Big(&\|\mathcal{F}(i_0,\alpha_0)\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\|g\|_{s+\bar{\sigma}}^{k_0,\gamma} \\ &+ \big\{\|\mathcal{F}(i_0,\alpha_0)\|_{s+\bar{\sigma}}^{k_0,\gamma} + \|\mathcal{F}(i_0,\alpha_0)\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar{\sigma}}^{k_0,\gamma}\big\}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\Big),\end{aligned} \tag{5.48}$$

$$\|\mathcal{P}_\omega g\|_s^{k_0,\gamma} \lesssim_S \varepsilon\gamma^{-2M-3}N_{n-1}^{-\mathtt{a}}\big(\|g\|_{s+\bar{\sigma}}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar{\sigma}}^{k_0,\gamma}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\big) , \tag{5.49}$$

$$\|\mathcal{P}_\omega^\perp g\|_{s_0}^{k_0,\gamma} \lesssim_{S,b} \gamma^{-1}K_n^{-b}\big(\|g\|_{s_0+\bar{\sigma}+b}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s_0+\mu(\mathtt{b})+\bar{\sigma}+b}^{k_0,\gamma}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\big) , \quad \forall b > 0 , \tag{5.50}$$

$$\|\mathcal{P}_\omega^\perp g\|_s^{k_0,\gamma} \lesssim_S \gamma^{-1}\big(\|g\|_{s+\bar{\sigma}}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})+\bar{\sigma}}^{k_0,\gamma}\|g\|_{s_0+\bar{\sigma}}^{k_0,\gamma}\big) . \tag{5.51}$$

Proof. Bound (5.46) follows from (5.45), (5.44), (5.24). By (4.10), since $X_{\mathcal{N}}$ does not depend on I, and i_δ differs by i_0 only in the I component (see (5.11)), we have

$$\mathcal{E}_0 := d_{i,\alpha}\mathcal{F}(i_0) - d_{i,\alpha}\mathcal{F}(i_\delta) = \varepsilon \int_0^1 \partial_I d_i X_P(\theta_0, I_\delta + s(I_0 - I_\delta), z_0)[I_0 - I_\delta, \Pi[\,\cdot\,]]ds \tag{5.52}$$

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 (5.16). Under the symplectic map G_δ, the nonlinear operator $\mathcal{F}$ in (4.10) is transformed into

$$\mathcal{F}(G_\delta(\mathtt{u}(\varphi)), \alpha) = DG_\delta(\mathtt{u}(\varphi))\big(\mathcal{D}_\omega \mathtt{u}(\varphi) - X_{K_\alpha}(\mathtt{u}(\varphi), \alpha)\big) \tag{5.53}$$

where $K_\alpha = H_\alpha \circ G_\delta$, see (5.17) and (5.19). Differentiating (5.53) at the trivial torus $\mathtt{u}_\delta(\varphi) = G_\delta^{-1}(i_\delta)(\varphi) = (\varphi, 0, 0)$, at $\alpha = \alpha_0$, we get

$$d_{i,\alpha}\mathcal{F}(i_\delta) = DG_\delta(\mathtt{u}_\delta)\big(\omega \cdot \partial_\varphi - d_{\mathtt{u},\alpha} X_{K_\alpha}(\mathtt{u}_\delta, \alpha_0)\big) D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1} + \mathcal{E}_1 \,, \tag{5.54}$$

$$\mathcal{E}_1 := D^2 G_\delta(\mathtt{u}_\delta)\big[DG_\delta(\mathtt{u}_\delta)^{-1}\mathcal{F}(i_\delta, \alpha_0),\, DG_\delta(\mathtt{u}_\delta)^{-1}\Pi[\,\cdot\,]\big] \tag{5.55}$$

In expanded form $\omega \cdot \partial_\varphi - d_{\mathtt{u},\alpha} X_{K_\alpha}(\mathtt{u}_\delta, \alpha_0)$ is provided by (5.23). By (5.34), (5.26), (5.29) and Lemma 5.4 we split

$$\omega \cdot \partial_\varphi - d_{\mathtt{u},\alpha} X_K(\mathtt{u}_\delta, \alpha_0) = \mathbb{D} + R_Z + \mathbb{R}_\omega + \mathbb{R}_\omega^\perp \tag{5.56}$$

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 (5.52), (5.54), (5.55), (5.56) we get the decomposition

$$d_{i,\alpha}\mathcal{F}(i_0) = DG_\delta(\mathtt{u}_\delta) \circ \mathbb{D} \circ D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1} + \mathcal{E} + \mathcal{E}_\omega + \mathcal{E}_\omega^\perp \tag{5.57}$$

where

$$\mathcal{E} := \mathcal{E}_0 + \mathcal{E}_1 + DG_\delta(\mathtt{u}_\delta) R_Z D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}, \quad \mathcal{E}_\omega := DG_\delta(\mathtt{u}_\delta)\mathbb{R}_\omega D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}, \tag{5.58}$$

$$\mathcal{E}_\omega^\perp := DG_\delta(\mathtt{u}_\delta)\mathbb{R}_\omega^\perp D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}. \tag{5.59}$$

Applying $\mathbf{T}_0$ defined in (5.45) to the right hand side in (5.57) (recall that $\mathtt{u}_\delta(\varphi) := (\varphi, 0, 0)$), since $\mathbb{D} \circ \mathbb{D}^{-1} = \mathrm{Id}$ (Proposition 5.5), we get

$$d_{i,\alpha}\mathcal{F}(i_0) \circ \mathbf{T}_0 - \mathrm{Id} = \mathcal{P} + \mathcal{P}_\omega + \mathcal{P}_\omega^\perp, \quad \mathcal{P} := \mathcal{E} \circ \mathbf{T}_0, \quad \mathcal{P}_\omega := \mathcal{E}_\omega \circ \mathbf{T}_0, \quad \mathcal{P}_\omega^\perp := \mathcal{E}_\omega^\perp \circ \mathbf{T}_0.$$

By (5.6), (5.21), (5.12), (5.13), (5.14), (5.24)-(5.25) we get the estimate

$$\|\mathcal{E}[\widehat{\imath}, \widehat{\alpha}]\|_s^{k_0,\gamma} \lesssim_s \|Z\|_{s_0+\sigma}^{k_0,\gamma} \|\widehat{\imath}\|_{s+\sigma}^{k_0,\gamma} + \|Z\|_{s+\sigma}^{k_0,\gamma} \|\widehat{\imath}\|_{s_0+\sigma}^{k_0,\gamma} + \|Z\|_{s_0+\sigma}^{k_0,\gamma} \|\widehat{\imath}\|_{s_0+\sigma}^{k_0,\gamma} \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}, \tag{5.60}$$

where $Z := \mathcal{F}(i_0, \alpha_0)$, recall (5.2). Then (5.48) follows from (5.46), (5.60), (5.6). Estimates (5.49), (5.50), (5.51) follow by (5.30)-(5.32), (5.46), (5.24), (5.12), (5.6). □

6 The linearized operator in the normal directions

In order to write an explicit expression of the linear operator $\mathcal{L}_\omega$ defined in (5.26) we have to express the operator $K_{02}(\phi)$ in terms of the original water waves Hamiltonian vector field.

Lemma 6.1. *The operator $K_{02}(\phi)$ is*

$$K_{02}(\phi) = \Pi_{\mathbb{S}^+}^{\perp} \partial_u \nabla_u H(T_\delta(\phi)) + \varepsilon R(\phi) \tag{6.1}$$

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*

$$T_\delta(\phi) := \varepsilon A(i_\delta(\phi)) = \varepsilon A(\theta_0(\phi), I_\delta(\phi), z_0(\phi)) = \varepsilon v(\theta_0(\phi), I_\delta(\phi)) + \varepsilon z_0(\phi) \tag{6.2}$$

with $A(\theta, I, z)$, $v(\theta, I)$ defined in (4.6)*. The operator $K_{02}(\phi)$ is even and reversible. The remainder $R(\phi)$ has the "finite dimensional" form*

$$R(\phi)[h] = \sum\nolimits_{j\in\mathbb{S}^+} \big(h\,,\,g_j\big)_{L^2_x} \chi_j\,, \quad \forall h \in H_{\mathbb{S}^+}^{\perp}\,, \tag{6.3}$$

for functions $g_j, \chi_j \in H_{\mathbb{S}^+}^{\perp}$ which satisfy the tame estimates: for some $\sigma := \sigma(\tau,\nu) > 0$, $\forall s \geq s_0$,

$$\|g_j\|_s^{k_0,\gamma} + \|\chi_j\|_s^{k_0,\gamma} \lesssim_s 1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{k_0,\gamma}\,, \quad \|d_i g_j[\widehat{\imath}]\|_s + \|d_i \chi_j[\widehat{\imath}]\|_s \lesssim_s \|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma}\,. \tag{6.4}$$

Proof. The lemma follows as in Lemma 6.1 in [21]. □

By Lemma 6.1 the linear operator $\mathcal{L}_\omega$ defined in (5.26) has the form

$$\mathcal{L}_\omega = \Pi_{\mathbb{S}^+}^{\perp}(\mathcal{L} + \varepsilon R)_{|H_{\mathbb{S}^+}^{\perp}} \qquad \text{where} \qquad \mathcal{L} := \omega\cdot\partial_\varphi - J\partial_u\nabla_u H(T_\delta(\varphi)) \tag{6.5}$$

is obtained linearizing the original water waves system (1.14), (1.6) at the torus $u = (\eta, \psi) = T_\delta(\varphi)$ defined in (6.2), 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).

In order to compute the linearization of the Dirichlet-Neumann operator, we recall the "shape derivative" formula, given for instance in [46], [47],

$$G'(\eta)[\hat\eta]\psi = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\{G(\eta+\epsilon\hat\eta)\psi - G(\eta)\psi\} = -G(\eta)(B\hat\eta) - \partial_x(V\hat\eta) \tag{6.6}$$

where

$$B := B(\eta,\psi) := \frac{\eta_x\psi_x + G(\eta)\psi}{1+\eta_x^2}\,, \qquad V := V(\eta,\psi) := \psi_x - B\eta_x\,. \tag{6.7}$$

It turns out that $(V, B) = \nabla_{x,y}\Phi$ is the velocity field evaluated at the free surface $(x, \eta(x))$. Using (6.6), the linearized operator of (1.14) is represented by the 2×2 operator matrix

$$\mathcal{L} := \omega\cdot\partial_\varphi + \begin{pmatrix} \partial_x V + G(\eta)B & -G(\eta) \\ (1 + BV_x) + BG(\eta)B & V\partial_x - BG(\eta) \end{pmatrix}. \tag{6.8}$$

Since the operator $G(\eta)$ is even according to Definition 2.19, 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.

Remark 6.2. It is convenient to first ignore the projection $\Pi_{\mathbb{S}_+}^{\perp}$ and consider the linearized operator $\mathcal{L}$ acting on the whole space $H^1(\mathbb{T})\times H^1(\mathbb{T})$. 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 subspace $H_{\mathbb{S}^+}^{\perp}$, see Section 13. 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 13. □

For the sequel we will always assume the following ansatz (satisfied by the approximate solutions obtained along the nonlinear Nash-Moser iteration of Section 15): for some constant $\mu_0 := \mu_0(\tau, \nu) > 0$, $\gamma \in (0,1)$,

$$\|\mathfrak{I}_0\|_{s_0+\mu_0}^{k_0,\gamma} \leq 1 \ , \qquad \text{and so, by (5.12), } \|\mathfrak{I}_\delta\|_{s_0+\mu_0}^{k_0,\gamma} \leq 2\,. \tag{6.9}$$

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

$$s_1 + \sigma_0 \leq s_0 + \mu_0\,, \quad \text{for some} \quad \sigma_0 := \sigma_0(\tau, \nu) > 0\,. \tag{6.10}$$

Thus by (6.9) we have

$$\|\mathfrak{I}_0\|_{s_1+\sigma_0}^{k_0,\gamma} \leq 1 \qquad \text{and so, by (5.12), } \|\mathfrak{I}_\delta\|_{s_1+\sigma_0}^{k_0,\gamma} \leq 2\,. \tag{6.11}$$

The constants μ_0 and σ_0 represent the *loss of derivatives* accumulated along the reduction procedure of Sections 7-12. What is important is that they are independent of the Sobolev index s. Along Sections 6-12, 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 (6.2) satisfies, $\forall s \geq s_0$,

$$\|u\|_s^{k_0,\gamma} = \|\eta\|_s^{k_0,\gamma} + \|\psi\|_s^{k_0,\gamma} \leq \varepsilon C(s)\big(1 + \|\mathfrak{I}_0\|_s^{k_0,\gamma}\big) \tag{6.12}$$

(the function A defined in (4.6) is smooth). Similarly

$$\|\Delta_{12} u\|_{s_1} \lesssim_{s_1} \varepsilon \|i_2 - i_1\|_{s_1} \tag{6.13}$$

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 (15.4). We recall that $\mathfrak{I}_0 := \mathfrak{I}_0(\omega, \mathtt{h})$ is defined for all $(\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ and that the functions B, V appearing in $\mathcal{L}$ in (6.8) are $\mathcal{C}^\infty$ in (φ, x) as the approximate torus $u = (\eta, \psi) = T_\delta(\varphi)$. This enables to use directly pseudo-differential operator theory as reminded in Section 2.3.

Starting from here, until the end of Section 13, our goal is to prove Proposition 13.3.

6.1 Linearized good unknown of Alinhac

Following [1], [21] we conjugate the linearized operator $\mathcal{L}$ in (6.8) by the multiplication operator

$$\mathcal{Z} := \begin{pmatrix} 1 & 0 \\ B & 1 \end{pmatrix}, \quad \mathcal{Z}^{-1} = \begin{pmatrix} 1 & 0 \\ -B & 1 \end{pmatrix}, \tag{6.14}$$

where $B = B(\varphi, x)$ is the function defined in (6.7), obtaining

$$\mathcal{L}_0 := \mathcal{Z}^{-1} \mathcal{L} \mathcal{Z} = \omega \cdot \partial_\varphi + \begin{pmatrix} \partial_x V & -G(\eta) \\ a & V \partial_x \end{pmatrix} \tag{6.15}$$

where a is the function

$$a := a(\varphi, x) := 1 + (\omega \cdot \partial_\varphi B) + V B_x\,. \tag{6.16}$$

All a, B, V are real valued periodic functions of (φ, x) — variable coefficients — and satisfy

$$B = \text{odd}(\varphi)\text{even}(x), \quad V = \text{odd}(\varphi)\text{odd}(x), \quad a = \text{even}(\varphi)\text{even}(x)\,.$$

The matrix $\mathcal{Z}$ in (6.14) amounts to introduce, as in Lannes [46]-[47], 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 6.3. *The maps* $\mathcal{Z}^{\pm 1} - \mathrm{Id}$ *are even, reversibility preserving and* $\mathcal{D}^{k_0}$*-tame with tame constants satisfying, for all* $s \geq s_0$,

$$\mathfrak{M}_{\mathcal{Z}^{\pm 1}-\mathrm{Id}}(s)\,,\,\mathfrak{M}_{(\mathcal{Z}^{\pm 1}-\mathrm{Id})^*}(s) \lesssim_s \varepsilon\big(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{6.17}$$

The operator $\mathcal{L}_0$ *is even and reversible. There is* $\sigma := \sigma(\tau,\nu) > 0$ *such that the functions*

$$\|a-1\|_s^{k_0,\gamma} + \|V\|_s^{k_0,\gamma} + \|B\|_s^{k_0,\gamma} \lesssim_s \varepsilon\big(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{6.18}$$

Moreover

$$\|\Delta_{12}a\|_{s_1} + \|\Delta_{12}V\|_{s_1} + \|\Delta_{12}B\|_{s_1} \lesssim_{s_1} \varepsilon\|i_1-i_2\|_{s_1+\sigma} \tag{6.19}$$

$$\|\Delta_{12}(\mathcal{Z}^{\pm 1})h\|_{s_1}\,,\,\|\Delta_{12}(\mathcal{Z}^{\pm 1})^*h\|_{s_1} \lesssim_{s_1} \varepsilon\|i_1-i_2\|_{s_1+\sigma}\|h\|_{s_1}\,. \tag{6.20}$$

Proof. The proof is the same as the one of Lemma 6.3 in [21]. □

We expand $\mathcal{L}_0$ in (6.15) as

$$\mathcal{L}_0 = \omega\cdot\partial_\varphi + \begin{pmatrix} V\partial_x & 0 \\ 0 & V\partial_x \end{pmatrix} + \begin{pmatrix} V_x & -G(\eta) \\ a & 0 \end{pmatrix}\,. \tag{6.21}$$

In the next section we deal with the first order operator $\omega\cdot\partial_\varphi + V\partial_x$.

7 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

$$\mathcal{B}^{-1}\big(\omega\cdot\partial_\varphi + V(\varphi,x)\partial_x\big)\mathcal{B} = \omega\cdot\partial_\varphi\,. \tag{7.1}$$

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 $\beta : \mathbb{T}^\nu_\varphi \times \mathbb{T}_x \to \mathbb{R}$ is odd in x, even in φ, and $\|\beta_x\|_{L^\infty} < 1/2$. We denote by $\mathcal{B}$ the corresponding composition operator, namely $(\mathcal{B}h)(\varphi,x) := h(\varphi, x+\beta(\varphi,x))$. The conjugated operator in the left hand side in (7.1) is

$$\mathcal{B}^{-1}\big(\omega\cdot\partial_\varphi + V(\varphi,x)\partial_x\big)\mathcal{B} = \omega\cdot\partial_\varphi + c(\varphi,y)\,\partial_y \tag{7.2}$$

where

$$c(\varphi,y) := \mathcal{B}^{-1}\big(\omega\cdot\partial_\varphi\beta + V(1+\beta_x)\big)(\varphi,y)\,. \tag{7.3}$$

In view of (7.2)-(7.3) we obtain (7.1) if $\beta(\varphi,x)$ solves the equation

$$\omega\cdot\partial_\varphi\beta(\varphi,x) + V(\varphi,x)(1+\beta_x(\varphi,x)) = 0\,, \tag{7.4}$$

which can be interpreted as a *quasi-periodic transport* equation.

Quasi-periodic characteristic equation. Instead of solving directly (7.4) we solve the equation satisfied by the inverse diffeomorphism

$$x+\beta(\varphi,x) = y \quad \Longleftrightarrow \quad x = y + \breve{\beta}(\varphi,y)\,, \quad \forall x,y\in\mathbb{R},\ \varphi\in\mathbb{T}^\nu\,. \tag{7.5}$$

It turns out that equation (7.4) for $\beta(\varphi,x)$ is equivalent to the following equation for $\breve{\beta}(\varphi,y)$:

$$\omega\cdot\partial_\varphi\breve{\beta}(\varphi,y) = V(\varphi, y+\breve{\beta}(\varphi,y)) \tag{7.6}$$

which is a quasi-periodic version of the *characteristic* equation $\dot{x} = V(\omega t, x)$.

Remark 7.1. We can give a geometric interpretation of equation (7.6) 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

$$\begin{pmatrix} \varphi \\ x \end{pmatrix} = \begin{pmatrix} \psi \\ y + \breve{\beta}(\psi, y) \end{pmatrix}, \quad \text{the system} \quad \frac{d}{dt}\begin{pmatrix} \varphi \\ x \end{pmatrix} = \begin{pmatrix} \omega \\ V(\varphi, x) \end{pmatrix}$$

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 (7.6) 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 $x(t) = c + \breve{\beta}(\omega t, c)$. $\square$

In Theorem 7.3 we solve equation (7.6), for $V(\varphi, x)$ small and ω Diophantine, by applying the Nash-Moser-Hörmander implicit function theorem in Appendix C. Rename $\breve{\beta} \to u$, $y \to x$, and write (7.6) as

$$F(u)(\varphi, x) := \omega \cdot \partial_\varphi u(\varphi, x) - V(\varphi, x + u(\varphi, x)) = 0\,. \tag{7.7}$$

The linearized operator at a given function $u(\varphi, x)$ is

$$F'(u)h := \omega \cdot \partial_\varphi h - q(\varphi, x)h, \quad q(\varphi, x) := V_x(\varphi, x + u(\varphi, x))\,. \tag{7.8}$$

In the next lemma we solve the linear problem $F'(u)h = f$.

Lemma 7.2. (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*

$$\|u\|^{k_0,\gamma}_{s_0+\varsigma} + \gamma^{-1}\|V\|^{k_0,\gamma}_{s_0+\varsigma} \leq \delta_0 \tag{7.9}$$

with δ_0 small enough. Then, given a periodic function f which is odd(φ)odd(x), the linearized equation

$$F'(u)h = f \tag{7.10}$$

has a unique periodic solution $h(\varphi, x)$ which is even(φ)odd(x) having zero average in φ, i.e.

$$\langle h \rangle_\varphi(x) := \frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} h(\varphi, x)\, d\varphi = 0 \quad \forall x \in \mathbb{T}. \tag{7.11}$$

This defines a right inverse of the linearized operator $F'(u)$, which we denote by $h = F'(u)^{-1}f$. It satisfies

$$\|F'(u)^{-1}f\|^{k_0,\gamma}_s \lesssim_s \gamma^{-1}\big(\|f\|^{k_0,\gamma}_{s+\varsigma} + \gamma^{-1}(\|V\|^{k_0,\gamma}_{s+\varsigma} + \|u\|^{k_0,\gamma}_{s+\varsigma}\|V\|^{k_0,\gamma}_{s_0+\varsigma})\|f\|^{k_0,\gamma}_{s_0}\big) \tag{7.12}$$

for all $s \geq s_0$, 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 (7.8). 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 (7.11); this gives uniqueness of the solution. Moreover $v = $ even(φ)even(x), $w_0 = $ even(φ)odd(x), $g = $ odd(x), whence h is even(φ)odd(x). Using (2.10), (2.11), (2.18), (2.19), (7.9), and (2.9) the proof of (7.12) is complete. $\square$

We now prove the existence of a solution of equation (7.7).

Theorem 7.3. (Solution of the quasi-periodic characteristic equation (7.7)**)** *Let ς be the constant defined in Lemma 7.2, 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*

$$\gamma^{-1} \|V\|^{k_0,\gamma}_{s_2+p} \le \delta, \tag{7.13}$$

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

$$\|u\|^{k_0,\gamma}_{s_2} \le C\gamma^{-1} \|V\|^{k_0,\gamma}_{s_2+p}. \tag{7.14}$$

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

$$\|u\|^{k_0,\gamma}_{s} \le C_s \gamma^{-1} \|V\|^{k_0,\gamma}_{s+p} \tag{7.15}$$

for some constant C_s depending on s, ς, s_0, independent of V, γ.

Proof. We apply Theorem C.1 of Appendix C. For $a, b \ge 0$, we define

$$E_a := \big\{ u \in \mathrm{Lip}(k_0, \mathtt{DC}(2\gamma,\tau), 2s_0 + a, \gamma) : u = \mathrm{even}(\varphi)\mathrm{odd}(x),\ \langle u \rangle_\varphi(x) = 0 \big\}, \quad \|u\|_{E_a} := \|u\|^{k_0,\gamma}_{2s_0+a}, \tag{7.16}$$

$$F_b := \big\{ g \in \mathrm{Lip}(k_0, \mathtt{DC}(2\gamma,\tau), 2s_0 + b, \gamma) : g = \mathrm{odd}(\varphi)\mathrm{odd}(x) \big\}, \qquad \|g\|_{F_b} := \|g\|^{k_0,\gamma}_{2s_0+b} \tag{7.17}$$

(s_0 is in the last term of (7.12), while $2s_0$ appears in the composition estimate (2.11)). We consider Fourier truncations at powers of 2 as smoothing operators, namely

$$S_n \ :\ u(\varphi, x) = \sum_{(\ell,j) \in \mathbb{Z}^{\nu+1}} u_{\ell j} e^{\mathrm{i}(\ell\cdot\varphi + jx)} \quad \mapsto \quad (S_n u)(\varphi, x) := \sum_{\langle \ell, j \rangle \le 2^n} u_{\ell j} e^{\mathrm{i}(\ell\cdot\varphi + jx)} \tag{7.18}$$

on both spaces E_a and F_b. Hence both E_a and F_b satisfy (C.1)-(C.5), and the operators R_n defined in (C.6) give the dyadic decomposition $2^n < \langle \ell, j \rangle \le 2^{n+1}$. Since S_n in (7.18) are "crude" Fourier truncations, (C.7) holds with "=" instead of "≤" and $C = 1$. As a consequence, every $g \in F_\beta$ satisfies the first inequality in (C.11) with $A = 1$ (it becomes, in fact, an equality), and, similarly, if $g \in F_{\beta+c}$ then (C.14) 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 (7.7) as the operator Φ of Theorem C.1. By Lemma 2.4, if $\|u\|^{k_0,\gamma}_{2s_0+1} \le \delta_{2.4}$ (where we denote by $\delta_{2.4}$ the constant δ of Lemma 2.4), then $\mathcal{V}(u)$ satisfies (2.11), namely for all $s \ge s_0$

$$\|\mathcal{V}(u)\|^{k_0,\gamma}_s \lesssim_s \|V\|^{k_0,\gamma}_{s+k_0} + \|u\|^{k_0,\gamma}_s \|V\|^{k_0,\gamma}_{s_0+k_0+1}, \tag{7.19}$$

and its second derivative $\mathcal{V}''(u)[v,w] = V_{xx}(\varphi, x + u(\varphi,x))vw$ satisfies

$$\begin{aligned} \|\mathcal{V}''(u)[v,w]\|^{k_0,\gamma}_s \lesssim_s\ & \|V\|^{k_0,\gamma}_{s_0+k_0+3} \Big(\|v\|^{k_0,\gamma}_s \|w\|^{k_0,\gamma}_{s_0} + \|v\|^{k_0,\gamma}_{s_0} \|w\|^{k_0,\gamma}_s \Big) \\ & + \big\{ \|V\|^{k_0,\gamma}_{s_0+k_0+3} \|u\|^{k_0,\gamma}_s + \|V\|^{k_0,\gamma}_{s+k_0+2} \big\} \|v\|^{k_0,\gamma}_{s_0} \|w\|^{k_0,\gamma}_{s_0}. \end{aligned} \tag{7.20}$$

We fix μ, U of Theorem C.1 as $\mu := 1$, $U := \{ u \in E_1 : \|u\|_{E_1} \le \delta_{2.4} \}$. 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 (7.24)). 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 (C.9) with $a_0 := 0$,

$$M_1(a) := C(a) \|V\|^{k_0,\gamma}_{s_0+k_0+3}, \quad M_2(a) := M_1(a), \quad M_3(a) := C(a) \|V\|^{k_0,\gamma}_{2s_0+k_0+2+a}. \tag{7.21}$$

We fix a_1, δ_1 of Theorem C.1 as $a_1 := \varsigma$, where $\varsigma = 3k_0 + 2\tau(k_0+1) + 2$ is the constant appearing in Lemma 7.2, and $\delta_1 := \frac{1}{2}\delta_{7.2}$, where $\delta_{7.2}$ is the constant δ_0 of Lemma 7.2. If $\gamma^{-1}\|V\|^{k_0,\gamma}_{s_0+\varsigma} \le \delta_1$ and $\|v\|_{E_{a_1}} \le \delta_1$, then, by Lemma 7.2, the right inverse $\Psi(v) := F'(v)^{-1}$ is well defined, and it satisfies

$$\|\Psi(v) g\|_{E_a} \le L_1(a) \|g\|_{F_{a+\varsigma}} + (L_2(a) \|v\|_{E_{a+\varsigma}} + L_3(a)) \|g\|_{F_0} \tag{7.22}$$

where

$$L_1(a) := C(a)\gamma^{-1}, \quad L_2(a) := C(a)\gamma^{-2}\|V\|^{k_0,\gamma}_{s_0+\varsigma}, \quad L_3(a) := C(a)\gamma^{-2}\|V\|^{k_0,\gamma}_{2s_0+a+\varsigma}. \tag{7.23}$$

We fix α, β, a_2 of Theorem C.1 as

$$\beta := 4\varsigma + 1, \quad \alpha := 3\varsigma + 1, \quad a_2 := 5\varsigma + 3, \tag{7.24}$$

so that (C.8) is satisfied. Bound (7.22) implies (C.10) 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 C.1 are satisfied. As a consequence, there exists a constant $\delta_{C.13}$ (given by (C.13) with $A = 1$) such that, if $\|g\|_{F_\beta} \leq \delta_{C.13}$, then the equation $\Phi(u) = \Phi(0) + g$ has a solution $u \in E_\alpha$, with bound (C.12). 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 $\|g\|_{F_\beta} \leq \delta_{C.13}$. Using (7.21), (7.23), (7.13), we verify that $\delta_{C.13} \geq C\gamma$. Thus, the smallness condition $\|g\|_{F_\beta} \leq \delta_{C.13}$ 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 (7.13), since $2s_0 + a_2 + \varsigma = s_2 + p$. Then (C.12), recalling (7.24), gives $\|u\|^{k_0,\gamma}_{s_2} \leq C\gamma^{-1}\|V\|^{k_0,\gamma}_{s_2+\varsigma}$, which implies (7.14) since $p \geq \varsigma$.

We finally prove estimate (7.15). 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 C.1 are satisfied. By (7.21), (7.23) and (7.13), we estimate the constants defined in (C.16)-(C.17) as

$$\mathcal{G}_1 \leq C_c\gamma^{-2}\|V\|^{k_0,\gamma}_{2s_0+a_2+c+\varsigma}, \quad \mathcal{G}_2 \leq C_c\gamma^{-1}, \quad z \leq C_c$$

for some constant C_c depending on c. Bound (C.15) implies (7.15) with $s = s_2 + c$ (the highest norm of V in (7.15) does not come from the term $\|V\|_{F_{\beta+c}}$ of (C.15), but from the factor $\mathcal{G}_1$). The proof is complete. □

The next lemma deals with the dependence of the solution u of (7.7) on V (actually it would be enough to estimate this Lipschitz dependence only in the "low" norm s_1 introduced in (6.10)).

Lemma 7.4. (Lipschitz dependence of u on V) *Let ς, s_2, p be as defined in Theorem 7.3. Let V_1, V_2 satisfy* (7.13)*, 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 7.3. Then for all $s \geq s_2 - \mu$ (where μ is the constant defined in (2.18)*)*

$$\|u_1 - u_2\|^{k_0,\gamma}_s \lesssim_s \gamma^{-1}\|V_1 - V_2\|^{k_0,\gamma}_{s+\mu+k_0} + \gamma^{-2}\max_{i=1,2}\|V_i\|^{k_0,\gamma}_{s+2\mu+p}\|V_1 - V_2\|^{k_0,\gamma}_{s_2+k_0}. \tag{7.25}$$

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 7.2)

$$v := (\omega \cdot \partial_\varphi)^{-1}a, \quad w := w_0 + g, \quad w_0 := (\omega \cdot \partial_\varphi)^{-1}(e^{-v}b), \quad g = g(x) := -\frac{\langle w_0 e^v\rangle_\varphi}{\langle e^v\rangle_\varphi}\,.$$

Then (7.25) follows by (2.11), (7.13), (7.14), (7.15), (2.18) and (2.19). □

In Theorem 7.3, 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 (7.7), namely the quasi-periodic characteristic equation (7.6), so that the periodic function β, defined by the inverse diffeomorphism in (7.5), solves the quasi-periodic transport equation (7.4).

By Theorem B.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 (B.6), the difference of the extended functions $\mathcal{E}_k(u_1) - \mathcal{E}_k(u_2)$ also satisfies the same estimate (7.25) 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

$$\beta_{ext} := \beta, \quad \breve{\beta}_{ext} := \breve{\beta}, \quad \mathcal{B}_{ext} := \mathcal{B}, \quad \mathcal{B}_{ext}^{-1} := \mathcal{B}^{-1}. \tag{7.26}$$

We have the following estimates on the transformations $\mathcal{B}$ and $\mathcal{B}^{-1}$.

Lemma 7.5. *Let $\beta, \breve{\beta}$ be defined in (7.26). There exists $\sigma := \sigma(\tau, \nu, k_0)$ such that, if (6.9) holds with $\mu_0 \geq \sigma$, then for any $s \geq s_2$,*

$$\|\beta\|_s^{k_0,\gamma}, \|\breve{\beta}\|_s^{k_0,\gamma} \lesssim_s \varepsilon \gamma^{-1} \big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) . \tag{7.27}$$

The operators $A = \mathcal{B}^{\pm 1} - \mathrm{Id}, (\mathcal{B}^{\pm 1} - \mathrm{Id})^*$ satisfy the estimates

$$\|Ah\|_s^{k_0,\gamma} \lesssim_s \varepsilon \gamma^{-1} \big(\|h\|_{s+k_0+1}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma} \|h\|_{s_0+k_0+2}^{k_0,\gamma}\big) \quad \forall s \geq s_2 . \tag{7.28}$$

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

$$\|\Delta_{12}\beta\|_{s_1}, \|\Delta_{12}\breve{\beta}\|_{s_1} \lesssim_{s_1} \varepsilon \gamma^{-1} \|i_1 - i_2\|_{s_1+\sigma} , \tag{7.29}$$

$$\|(\Delta_{12} A)[h]\|_{s_1} \lesssim_{s_1} \varepsilon \gamma^{-1} \|i_1 - i_2\|_{s_1+\sigma} \|h\|_{s_1+1} , \quad A \in \{\mathcal{B}^{\pm 1}, (\mathcal{B}^{\pm 1})^*\} , \tag{7.30}$$

where s_1 is introduced in (6.10).

Proof. Bound (7.27) for $\breve{\beta}$ follows, recalling that $\breve{\beta} = u$, by (7.15) and (6.18). Estimate (7.27) for β follows by that for $\breve{\beta}$, applying (2.12). We now prove estimate (7.28) 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 (7.28) follows by applying (2.11) to the operator $\mathcal{B}_\tau$, using the estimates on β, ansatz (6.9) and (2.10). The estimate for $\mathcal{B}^{-1} - \mathrm{Id}$ is obtained similarly. The estimate on the adjoint operators follows because

$$\mathcal{B}^* h(\varphi, y) = (1 + \breve{\beta}(\varphi, y)) h(\varphi, y + \breve{\beta}(\varphi, y)), \quad (\mathcal{B}^{-1})^* h(\varphi, x) = (1 + \beta(\varphi, x)) h(\varphi, x + \beta(\varphi, x)) .$$

Estimates (7.29), (7.30) follow by Lemma 7.4, and by (6.18)-(6.19). $\square$

We now conjugate the whole operator $\mathcal{L}_0$ in (6.15) by the diffeomorphism $\mathcal{B}$.

Lemma 7.6. *Let $\beta, \breve{\beta}, \mathcal{B}, \mathcal{B}^{-1}$ be defined in (7.26). 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 (6.15) to*

$$\mathcal{L}_1 := \mathcal{B}^{-1} \mathcal{L}_0 \mathcal{B} = \omega \cdot \partial_\varphi + \begin{pmatrix} a_1 & -a_2 \partial_y \mathcal{H} T_{\mathtt{h}} + \mathcal{R}_1 \\ a_3 & 0 \end{pmatrix} , \tag{7.31}$$

$$T_{\mathtt{h}} := \tanh(\mathtt{h} |D_y|) := \mathrm{Op}\big(\tanh(\mathtt{h} \chi(\xi) |\xi|)\big), \tag{7.32}$$

where a_1, a_2, a_3 are the functions

$$a_1(\varphi, y) := (\mathcal{B}^{-1} V_x)(\varphi, y), \quad a_2(\varphi, y) := 1 + (\mathcal{B}^{-1} \beta_x)(\varphi, y) , \quad a_3(\varphi, y) := (\mathcal{B}^{-1} a)(\varphi, y), \tag{7.33}$$

and $\mathcal{R}_1$ is a pseudo-differential operator of order $OPS^{-\infty}$. Formula (7.33) 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

$$\beta = even(\varphi) odd(x); \qquad a_1 = odd(\varphi) even(y); \qquad a_2, a_3 = even(\varphi) even(y). \tag{7.34}$$

There exists $\sigma = \sigma(\tau, \nu, k_0) > 0$ such that for any $m, \alpha \geq 0$, assuming (6.9) 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:

$$\|a_1\|_s^{k_0,\gamma} + \|a_2 - 1\|_s^{k_0,\gamma} + \|a_3 - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big), \tag{7.35}$$

$$|\!|\!|\mathcal{R}_1|\!|\!|_{-m,s,\alpha}^{k_0,\gamma} \lesssim_{m,s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+m+\alpha}^{k_0,\gamma}\big). \tag{7.36}$$

Finally, given two tori i_1, i_2, we have

$$\|\Delta_{12} a_1\|_{s_1} + \|\Delta_{12} a_2\|_{s_1} + \|\Delta_{12} a_3\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}, \tag{7.37}$$

$$|\!|\!|\Delta_{12}\mathcal{R}_1|\!|\!|_{-m,s_1,\alpha} \lesssim_{m,s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma+m+\alpha}. \tag{7.38}$$

Proof. By (6.21) and (7.2)-(7.4) we have that

$$\mathcal{L}_1 := \mathcal{B}^{-1}\mathcal{L}_0\mathcal{B} = \omega \cdot \partial_\varphi + \begin{pmatrix} a_1 & -\mathcal{B}^{-1}G(\eta)\mathcal{B} \\ a_3 & 0 \end{pmatrix} \tag{7.39}$$

where the functions a_1 and a_3 are defined in (7.33). We now conjugate the Dirichlet-Neumann operator $G(\eta)$ under the diffeomorphism $\mathcal{B}$. Following Proposition A.1, we write

$$G(\eta) = |D_x| \tanh(\mathtt{h}|D_x|) + \mathcal{R}_G = \partial_x \mathcal{H} T_{\mathtt{h}} + \mathcal{R}_G, \qquad T_{\mathtt{h}} := \tanh(\mathtt{h}|D_x|), \tag{7.40}$$

where $\mathcal{R}_G$ is an integral operator in $OPS^{-\infty}$. We decompose

$$\tanh(\mathtt{h}|D_x|) = \mathrm{Id} + \mathrm{Op}(r_{\mathtt{h}}), \quad r_{\mathtt{h}}(\xi) := -\frac{2}{1 + e^{2\mathtt{h}|\xi|\chi(\xi)}} \in S^{-\infty}, \tag{7.41}$$

and, since $\mathcal{B}^{-1}\, \partial_x\, \mathcal{B} = a_2 \partial_y$ where the function a_2 is defined in (7.33), we have

$$\begin{aligned} \mathcal{B}^{-1}\partial_x \mathcal{H} T_{\mathtt{h}} \mathcal{B} &= (\mathcal{B}^{-1}\partial_x \mathcal{B})(\mathcal{B}^{-1}\mathcal{H}\mathcal{B})(\mathcal{B}^{-1} T_{\mathtt{h}} \mathcal{B}) = a_2 \partial_y \{\mathcal{H} + (\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H})\}(\mathcal{B}^{-1} T_{\mathtt{h}} \mathcal{B}) \\ &= a_2 \partial_y \mathcal{H} T_{\mathtt{h}} + a_2 \partial_y \mathcal{H}[\mathcal{B}^{-1}\mathrm{Op}(r_{\mathtt{h}})\mathcal{B} - \mathrm{Op}(r_{\mathtt{h}})] + a_2 \partial_y (\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H})(\mathcal{B}^{-1} T_{\mathtt{h}} \mathcal{B}). \end{aligned} \tag{7.42}$$

Therefore by (7.40)-(7.42) we get

$$-\mathcal{B}^{-1} G(\eta)\mathcal{B} = -a_2 \partial_y \mathcal{H} T_{\mathtt{h}} + \mathcal{R}_1, \tag{7.43}$$

where $\mathcal{R}_1$ is the operator in $OPS^{-\infty}$ defined by

$$\begin{aligned} \mathcal{R}_1 &:= \mathcal{R}_1^{(1)} + \mathcal{R}_1^{(2)} + \mathcal{R}_1^{(3)} & \mathcal{R}_1^{(1)} &:= -\mathcal{B}^{-1}\mathcal{R}_G\mathcal{B}, \\ \mathcal{R}_1^{(2)} &:= -a_2\partial_y \mathcal{H}[\mathcal{B}^{-1}\mathrm{Op}(r_{\mathtt{h}})\mathcal{B} - \mathrm{Op}(r_{\mathtt{h}})], & \mathcal{R}_1^{(3)} &:= -a_2 \partial_y(\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H})\mathcal{B}^{-1} T_{\mathtt{h}} \mathcal{B}. \end{aligned} \tag{7.44}$$

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 (7.40) and in (7.41), are in $OPS^{-\infty}$. The operator $\mathcal{B}^{-1}\mathcal{H}\mathcal{B} - \mathcal{H}$ is in $OPS^{-\infty}$ by Lemma 2.17.

In conclusion, (7.39) and (7.43) imply (7.31)-(7.33), for all λ in the Cantor set $\mathtt{DC}(\gamma, \tau) \times [\mathtt{h}_1, \mathtt{h}_2]$. By formulas (7.44), $\mathcal{R}_1$ is defined on the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$.

Estimates (7.35), (7.37) for a_1, a_2, a_3 on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ follow by (6.18), (6.19) and Lemma 7.5. Estimates (7.36), (7.38) follow applying Lemmata 2.15 and 2.17 and Proposition A.1, and by using Lemma 7.5. □

Remark 7.7. We stress that the conjugation identity (7.31) 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 (7.33) and (7.44) 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 (7.31) 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 (7.31)-(7.34) the space variable y by x.

8 Change of the space variable

We consider a φ-independent diffeomorphism of the torus $\mathbb{T}$ of the form

$$y = x + \alpha(x) \qquad \text{with inverse} \qquad x = y + \breve{\alpha}(y) \tag{8.1}$$

where α is a $\mathcal{C}^\infty(\mathbb{T}_x)$ real valued function, independent of φ, satisfying $\|\alpha_x\|_{L^\infty} \leq 1/2$. We also make the following ansatz on α that will be verified when we choose it in Section 11, see formula (11.23): 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 $k \in \mathbb{N}^{\nu+1}$, $|k| \leq k_0$, and it satisfies the estimate

$$\|\alpha\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,,\ \forall s \geq s_0\,, \quad \|\Delta_{12}\alpha\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,, \tag{8.2}$$

for some $\sigma = \sigma(k_0, \tau, \nu) > 0$. By (8.2) and Lemma 2.4, arguing as in the proof of Lemma 7.5, one gets

$$\|\breve{\alpha}\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,,\ \forall s \geq s_0\,, \quad \|\Delta_{12}\breve{\alpha}\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,, \tag{8.3}$$

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 (7.31) by the composition operator

$$(\mathcal{A}u)(\varphi, x) := u(\varphi, x + \alpha(x)), \quad (\mathcal{A}^{-1}u)(\varphi, y) := u(\varphi, y + \breve{\alpha}(y))\,. \tag{8.4}$$

By (7.31), using that the operator $\mathcal{A}$ is φ-independent, recalling expansion (7.41) and arguing as in (7.42) to compute the conjugation $\mathcal{A}^{-1}\big(-a_2\partial_x \mathcal{H} T_{\mathtt{h}}\big)\mathcal{A}$, one has

$$\mathcal{L}_2 := \mathcal{A}^{-1}\mathcal{L}_1\mathcal{A} = \omega\cdot\partial_\varphi + \begin{pmatrix} a_4 & -a_5\partial_y \mathcal{H} T_{\mathtt{h}} + \mathcal{R}_2 \\ a_6 & 0 \end{pmatrix}\,, \tag{8.5}$$

where a_4, a_5, a_6 are the functions

$$a_4(\varphi, y) := (\mathcal{A}^{-1}a_1)(\varphi, y) = a_1(\varphi, y + \breve{\alpha}(y))\,, \tag{8.6}$$

$$a_5(\varphi, y) := \big(\mathcal{A}^{-1}(a_2(1 + \alpha_x))\big)(\varphi, y) = \{a_2(\varphi, x)(1 + \alpha_x(x))\}|_{x = y + \breve{\alpha}(y)} \tag{8.7}$$

$$a_6(\varphi, y) := (\mathcal{A}^{-1}a_3)(\varphi, y) = a_3(\varphi, y + \breve{\alpha}(y)) \tag{8.8}$$

and $\mathcal{R}_2$ is the operator in $OPS^{-\infty}$ given by

$$\mathcal{R}_2 := -a_5\partial_y \mathcal{H}\big[\mathcal{A}^{-1}\mathrm{Op}(r_{\mathtt{h}})\mathcal{A} - \mathrm{Op}(r_{\mathtt{h}})\big] - a_5\partial_y(\mathcal{A}^{-1}\mathcal{H}\mathcal{A} - \mathcal{H})(\mathcal{A}^{-1}T_{\mathtt{h}}\mathcal{A}) + \mathcal{A}^{-1}\mathcal{R}_1\mathcal{A}\,. \tag{8.9}$$

Lemma 8.1. *There exists a constant $\sigma = \sigma(k_0, \tau, \nu) > 0$ such that, if (6.9) 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*

$$\begin{aligned} &\|Ah\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(\|h\|_{s+k_0+1}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|h\|_{s_0+k_0+2}^{k_0,\gamma}\big)\,, \quad \forall s \geq s_0\,, \\ &\|(\Delta_{12}A)h\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\|h\|_{s_1+1}\,. \end{aligned} \tag{8.10}$$

The real valued functions a_4, a_5, a_6 in (8.6)-(8.8) satisfy

$$a_4 = odd(\varphi)even(y), \qquad a_5, a_6 = even(\varphi)even(y)\,, \tag{8.11}$$

and

$$\begin{aligned} &\|a_4\|_s^{k_0,\gamma}\,, \|a_5 - 1\|_s^{k_0,\gamma}\,, \|a_6 - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) \\ &\|\Delta_{12}a_4\|_{s_1}\,, \|\Delta_{12}a_5\|_{s_1}\,, \|\Delta_{12}a_6\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,. \end{aligned} \tag{8.12}$$

The remainder $\mathcal{R}_2$ defined in (8.9) is an even and reversible pseudo-differential operator in $OPS^{-\infty}$. Moreover, for any $m, \alpha \geq 0$, and assuming (6.9) with $\sigma + m + \alpha \leq \mu_0$, the following estimates hold:

$$\begin{aligned} &|\mathcal{R}_2|_{-m,s,\alpha}^{k_0,\gamma} \lesssim_{m,s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+m+\alpha}^{k_0,\gamma}\big)\,, \quad \forall s \geq s_0 \\ &|\Delta_{12}\mathcal{R}_2|_{-m,s_1,\alpha} \lesssim_{m,s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma+m+\alpha}\,. \end{aligned} \tag{8.13}$$

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 (8.10) can be proved by using (8.2), (8.3), arguing as in the proof of Lemma 7.5.

Estimate (8.12) follows by definitions (8.6)-(8.8), by estimates (8.2), (8.3), (8.10), (7.35), (7.37), and by applying Lemma 2.4. Estimates (8.13) of the remainder $\mathcal{R}_2$ follow by using the same arguments we used in Lemma 7.6 to get estimates (7.36), (7.38) for the remainder $\mathcal{R}_1$. □

In the sequel we rename in (8.5)-(8.9) the space variable y by x.

9 Symmetrization of the order $1/2$

The aim of this section is to conjugate the operator $\mathcal{L}_2$ defined in (8.5) 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 (9.10)-(9.14)). In the complex variables $(u, \bar{u})$ that we will introduce in Section 10, this amounts to the symmetrization of the linear operator at the highest order, see (10.1)-(10.3).

We first conjugate $\mathcal{L}_2$ by the real, even and reversibility preserving transformation

$$\mathcal{M}_2 := \begin{pmatrix} \Lambda_{\mathtt{h}} & 0 \\ 0 & \Lambda_{\mathtt{h}}^{-1} \end{pmatrix}, \tag{9.1}$$

where $\Lambda_{\mathtt{h}}$ is the Fourier multiplier, acting on the periodic functions,

$$\Lambda_{\mathtt{h}} := \pi_0 + |D|^{\frac{1}{4}} T_{\mathtt{h}}^{\frac{1}{4}}, \qquad \text{with inverse} \qquad \Lambda_{\mathtt{h}}^{-1} = \pi_0 + |D|^{-\frac{1}{4}} T_{\mathtt{h}}^{-\frac{1}{4}}, \tag{9.2}$$

with $T_{\mathtt{h}} = \tanh(\mathtt{h}|D|)$ and π_0 defined in (2.33). The conjugated operator is

$$\mathcal{L}_3 := \mathcal{M}_2^{-1} \mathcal{L}_2 \mathcal{M}_2 = \omega \cdot \partial_\varphi + \begin{pmatrix} \Lambda_{\mathtt{h}}^{-1} a_4 \Lambda_{\mathtt{h}} & \Lambda_{\mathtt{h}}^{-1}(-a_5 \partial_x \mathcal{H} T_{\mathtt{h}} + \mathcal{R}_2) \Lambda_{\mathtt{h}}^{-1} \\ \Lambda_{\mathtt{h}} a_6 \Lambda_{\mathtt{h}} & 0 \end{pmatrix} =: \omega \cdot \partial_\varphi + \begin{pmatrix} A_3 & B_3 \\ C_3 & 0 \end{pmatrix}. \tag{9.3}$$

We develop the operators in (9.3) up to order $-1/2$. First we write

$$A_3 = \Lambda_{\mathtt{h}}^{-1} a_4 \Lambda_{\mathtt{h}} = a_4 + \mathcal{R}_{A_3} \qquad \text{where} \qquad \mathcal{R}_{A_3} := [\Lambda_{\mathtt{h}}^{-1}, a_4] \Lambda_{\mathtt{h}} \in OPS^{-1} \tag{9.4}$$

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

$$\begin{aligned} C_3 = \Lambda_{\mathtt{h}} a_6 \Lambda_{\mathtt{h}} &= a_6 \Lambda_{\mathtt{h}}^2 + [\Lambda_{\mathtt{h}}, a_6] \Lambda_{\mathtt{h}} = a_6 (\pi_0 + |D|^{\frac{1}{4}} T_{\mathtt{h}}^{\frac{1}{4}})^2 + [\Lambda_{\mathtt{h}}, a_6] \Lambda_{\mathtt{h}} \\ &= a_6 |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} + \pi_0 + \mathcal{R}_{C_3} \qquad \text{where} \qquad \mathcal{R}_{C_3} := (a_6 - 1)\pi_0 + [\Lambda_{\mathtt{h}}, a_6] \Lambda_{\mathtt{h}}. \end{aligned} \tag{9.5}$$

Using that $|D| = \mathcal{H} \partial_x$, (9.2) and $|D| \pi_0 = 0$ on the periodic functions, we write B_3 in (9.3) as

$$\begin{aligned} B_3 &= \Lambda_{\mathtt{h}}^{-1}(-a_5 \partial_x \mathcal{H} T_{\mathtt{h}} + \mathcal{R}_2) \Lambda_{\mathtt{h}}^{-1} = -a_5 |D| T_{\mathtt{h}} \Lambda_{\mathtt{h}}^{-2} - [\Lambda_{\mathtt{h}}^{-1}, a_5] |D| T_{\mathtt{h}} \Lambda_{\mathtt{h}}^{-1} + \Lambda_{\mathtt{h}}^{-1} \mathcal{R}_2 \Lambda_{\mathtt{h}}^{-1} \\ &= -a_5 |D| T_{\mathtt{h}} \big(\pi_0 + |D|^{-\frac{1}{4}} T_{\mathtt{h}}^{-\frac{1}{4}}\big)^2 - [\Lambda_{\mathtt{h}}^{-1}, a_5] |D| T_{\mathtt{h}} \Lambda_{\mathtt{h}}^{-1} + \Lambda_{\mathtt{h}}^{-1} \mathcal{R}_2 \Lambda_{\mathtt{h}}^{-1} \\ &= -a_5 |D|^{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} + \mathcal{R}_{B_3} \qquad \text{where} \qquad \mathcal{R}_{B_3} := -[\Lambda_{\mathtt{h}}^{-1}, a_5] |D| T_{\mathtt{h}} \Lambda_{\mathtt{h}}^{-1} + \Lambda_{\mathtt{h}}^{-1} \mathcal{R}_2 \Lambda_{\mathtt{h}}^{-1}. \end{aligned} \tag{9.6}$$

Lemma 9.1. *The operators $\Lambda_{\mathtt{h}} \in OPS^{\frac{1}{4}}$, $\Lambda_{\mathtt{h}}^{-1} \in OPS^{-\frac{1}{4}}$ and $\mathcal{R}_{A_3}, \mathcal{R}_{B_3}, \mathcal{R}_{C_3} \in OPS^{-\frac{1}{2}}$. Furthermore, there exists $\sigma(k_0, \tau, \nu) > 0$ such that for any $\alpha > 0$, assuming (6.9) with $\mu_0 \geq \sigma + \alpha$, then for all $s \geq s_0$,*

$$|\Lambda_{\mathtt{h}}|_{\frac{1}{4}, s, \alpha}^{k_0, \gamma}, |\Lambda_{\mathtt{h}}^{-1}|_{-\frac{1}{4}, s, \alpha}^{k_0, \gamma} \lesssim_\alpha 1, \tag{9.7}$$

$$|\mathcal{R}|_{-\frac{1}{2}, s, \alpha}^{k_0, \gamma} \lesssim_{s, \alpha} \varepsilon \gamma^{-1} \big(1 + \|\mathfrak{I}_0\|_{s + \sigma + \alpha}^{k_0, \gamma}\big), \quad |\Delta_{12} \mathcal{R}|_{-\frac{1}{2}, s_1, \alpha} \lesssim_{s_1, \alpha} \varepsilon \gamma^{-1} \|\Delta_{12} i\|_{s_1 + \sigma + \alpha} \tag{9.8}$$

for all $\mathcal{R} \in \{\mathcal{R}_{A_3}, \mathcal{R}_{B_3}, \mathcal{R}_{C_3}\}$. The operator $\mathcal{L}_3$ in (9.3) 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 (9.4), (9.6), (9.5), by Lemmata 2.10 and 2.11, recalling (2.39) and using (8.12), (8.13). $\square$

Consider now a transformation $\mathcal{M}_3$ of the form

$$\mathcal{M}_3 := \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathcal{M}_3^{-1} = \begin{pmatrix} p^{-1} & 0 \\ 0 & 1 \end{pmatrix}, \tag{9.9}$$

where $p(\varphi, x)$ is a real-valued periodic function, with $p-1$ small (see (9.14)). The conjugated operator is

$$\mathcal{L}_4 := \mathcal{M}_3^{-1}\mathcal{L}_3\mathcal{M}_3 = \omega\cdot\partial_\varphi + \begin{pmatrix} p^{-1}(\omega\cdot\partial_\varphi p) + p^{-1}A_3 p & p^{-1}B_3 \\ C_3 p & 0 \end{pmatrix} = \omega\cdot\partial_\varphi + \begin{pmatrix} A_4 & B_4 \\ C_4 & 0 \end{pmatrix} \tag{9.10}$$

where, recalling (9.4), (9.6), (9.5), one has

$$A_4 = \breve{a}_4 + \mathcal{R}_{A_4}, \quad \breve{a}_4 := a_4 + p^{-1}(\omega\cdot\partial_\varphi p), \quad \mathcal{R}_{A_4} := p^{-1}\mathcal{R}_{A_3} p \tag{9.11}$$

$$B_4 = -p^{-1}a_5|D|^{\frac12}T_{\mathtt{h}}^{\frac12} + \mathcal{R}_{B_4}, \quad \mathcal{R}_{B_4} := p^{-1}\mathcal{R}_{B_3} \tag{9.12}$$

$$C_4 = a_6 p|D|^{\frac12}T_{\mathtt{h}}^{\frac12} + \pi_0 + \mathcal{R}_{C_4}, \quad \mathcal{R}_{C_4} := a_6[|D|^{\frac12}T_{\mathtt{h}}^{\frac12}, p] + \pi_0(p-1) + \mathcal{R}_{C_3} p \tag{9.13}$$

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 (9.12) and C_4 in (9.13) are opposite if $a_6 p = p^{-1}a_5$. Therefore we fix the real valued function

$$p := \sqrt{\frac{a_5}{a_6}}, \qquad a_6 p = p^{-1}a_5 = \sqrt{a_5 a_6} =: a_7 \,. \tag{9.14}$$

Lemma 9.2. *There exists $\sigma := \sigma(\tau, \nu, k_0) > 0$ such that for any $\alpha > 0$, assuming* (6.9) *with $\mu_0 \geq \sigma + \alpha$, then for any $s \geq s_0$ the following holds. The transformation $\mathcal{M}_3$ defined in* (9.9) *is real, even and reversibility preserving and satisfies*

$$|\mathcal{M}_3^{\pm1} - \mathrm{Id}|_{0,s,0}^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{9.15}$$

The real valued functions $\breve{a}_4, a_7$ defined in (9.11), (9.14) *satisfy*

$$\breve{a}_4 = \mathrm{odd}(\varphi)\mathrm{even}(x)\,, \qquad a_7 = \mathrm{even}(\varphi)\mathrm{even}(x)\,, \tag{9.16}$$

and, for any $s \geq s_0$,

$$\|\breve{a}_4\|_s^{k_0,\gamma}, \|a_7 - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{9.17}$$

The remainders $\mathcal{R}_{A_4}, \mathcal{R}_{B_4}, \mathcal{R}_{C_4} \in OPS^{-\frac12}$ defined in (9.11)-(9.13) *satisfy*

$$|\mathcal{R}|_{-\frac12,s,\alpha}^{k_0,\gamma} \lesssim_{s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\alpha}^{k_0,\gamma}\big)\,, \qquad \mathcal{R} \in \{\mathcal{R}_{A_4}, \mathcal{R}_{B_4}, \mathcal{R}_{C_4}\}\,. \tag{9.18}$$

Let i_1, i_2 be given embedded tori. Then

$$|\Delta_{12}\mathcal{M}_3^{\pm1}|_{0,s_1,0} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,, \tag{9.19}$$

$$\|\Delta_{12}\breve{a}_4\|_{s_1}, \|\Delta_{12} a_7\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}\,, \tag{9.20}$$

$$|\Delta_{12}\mathcal{R}|_{-\frac12,s_1,\alpha} \lesssim_{s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma+\alpha}\,, \qquad \mathcal{R} \in \{\mathcal{R}_{A_4}, \mathcal{R}_{B_4}, \mathcal{R}_{C_4}\}\,. \tag{9.21}$$

The operator $\mathcal{L}_4$ in (9.10) *is real, even and reversible.*

Proof. By (8.11), 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 (9.16). Since p is even(φ)even(x), the transformation $\mathcal{M}_3$ is real, even and reversibility preserving.

By definition (9.14), Lemma 2.6, the interpolation estimate (2.10) and applying estimates (8.12) on a_5 and a_6, one gets that p satisfies the estimates

$$\|p^{\pm1} - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \|\Delta_{12} p^{\pm1}\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma} \tag{9.22}$$

for some $\sigma = \sigma(\tau, \nu, k_0) > 0$. Hence estimates (9.15), (9.19) for $\mathcal{M}_3^{\pm1}$ follow by definition (9.9), using estimates (2.39), (9.22). Estimates (9.17), (9.20) for $\breve{a}_4, a_7$ follow by definitions (9.11), (9.14) and applying estimates (8.12) on a_4, a_5 and a_6, estimates (9.22) on p, Lemma 2.6 and the interpolation estimate (2.10). Estimates (9.18), (9.21) follow by definitions (9.11)-(9.13), estimate (2.39), Lemmata 2.10 and 2.11, bounds (8.12) on a_4, a_5, a_6, (9.22) on p, and Lemma 9.1. $\square$

10 Symmetrization of the lower orders

To symmetrize the linear operator $\mathcal{L}_4$ in (9.10), with p fixed in (9.14), at lower orders, it is convenient to introduce the complex coordinates $(u,\bar{u}) := \mathcal{C}^{-1}(\eta,\psi)$, with $\mathcal{C}$ defined in (2.60), namely $u = \eta + \mathrm{i}\psi$, $\bar{u} = \eta - \mathrm{i}\psi$. In these complex coordinates the linear operator $\mathcal{L}_4$ becomes, using (2.61) and (9.14),

$$\mathcal{L}_5 := \mathcal{C}^{-1}\mathcal{L}_4\mathcal{C} = \omega\cdot\partial_\varphi + \mathrm{i}a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma + a_8\mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_5 + \mathcal{Q}_5\,, \qquad a_8 := \frac{\breve{a}_4}{2}\,, \tag{10.1}$$

where the real valued functions a_7, $\breve{a}_4$ are defined in (9.14), (9.11) and satisfy (9.16),

$$\Sigma := \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \qquad \Pi_0 := \frac12\begin{pmatrix}\pi_0 & \pi_0\\ -\pi_0 & -\pi_0\end{pmatrix}, \qquad \mathbb{I}_2 := \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \tag{10.2}$$

π_0 is defined in (2.33), and

$$\begin{gathered}\mathcal{P}_5 := \begin{pmatrix}P_5 & 0\\ 0 & \overline{P_5}\end{pmatrix}, \quad \mathcal{Q}_5 := \begin{pmatrix}0 & Q_5\\ \overline{Q_5} & 0\end{pmatrix},\\ P_5 := \frac12\big\{\mathcal{R}_{A_4} + \mathrm{i}(\mathcal{R}_{C_4} - \mathcal{R}_{B_4})\big\}\,, \qquad Q_5 := a_8 + \frac12\big\{\mathcal{R}_{A_4} + \mathrm{i}(\mathcal{R}_{C_4} + \mathcal{R}_{B_4})\big\}\,.\end{gathered} \tag{10.3}$$

By the estimates of Lemma 9.2 we have

$$\|a_7 - 1\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \|\Delta_{12}a_7\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma} \tag{10.4}$$

$$\|a_8\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,, \quad \|\Delta_{12}a_8\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma}\,, \tag{10.5}$$

$$|\mathcal{P}_5|_{-\frac12,s,\alpha}^{k_0,\gamma}\,,\ |\mathcal{Q}_5|_{0,s,\alpha}^{k_0,\gamma} \lesssim_{s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\alpha}^{k_0,\gamma}\big) \tag{10.6}$$

$$|\Delta_{12}\mathcal{P}_5|_{-\frac12,s_1,\alpha}\,,\ |\Delta_{12}\mathcal{Q}_5|_{0,s_1,\alpha} \lesssim_{s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma+\alpha}\,. \tag{10.7}$$

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 (14.8).

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

$$\mathcal{L}_5^{(m)} := \omega\cdot\partial_\varphi + \mathrm{i}a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma + a_8\mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_5^{(m)} + \mathcal{Q}_5^{(m)}\,, \tag{10.8}$$

where

$$\begin{gathered}\mathcal{P}_5^{(m)} = \begin{pmatrix}P_5^{(m)} & 0\\ 0 & \overline{P_5^{(m)}}\end{pmatrix}, \quad \mathcal{Q}_5^{(m)} = \begin{pmatrix}0 & Q_5^{(m)}\\ \overline{Q_5^{(m)}} & 0\end{pmatrix},\\ P_5^{(m)} = \mathrm{Op}(p_m) \in OPS^{-\frac12}\,, \quad Q_5^{(m)} = \mathrm{Op}(q_m) \in OPS^{-\frac{m}{2}}\,.\end{gathered} \tag{10.9}$$

For any $\alpha \in \mathbb{N}$, assuming (6.9) with $\mu_0 \geq \aleph_4(m,\alpha) + \sigma$, where the increasing constants $\aleph_4(m,\alpha)$ are defined inductively by

$$\aleph_4(0,\alpha) := \alpha\,, \qquad \aleph_4(m+1,\alpha) := \aleph_4(m,\alpha+1) + \frac{m}{2} + 2\alpha + 4\,, \tag{10.10}$$

we have

$$|\mathcal{P}_5^{(m)}|_{-\frac12,s,\alpha}^{k_0,\gamma}\,,\ |\mathcal{Q}_5^{(m)}|_{-\frac{m}{2},s,\alpha} \lesssim_{m,s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_4(m,\alpha)+\sigma}^{k_0,\gamma}\big)\,, \tag{10.11}$$

$$|\Delta_{12}\mathcal{P}_5^{(m)}|_{-\frac12,s_1,\alpha}\,,\ |\Delta_{12}\mathcal{Q}_5^{(m)}|_{-\frac{m}{2},s_1,\alpha} \lesssim_{m,s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\aleph_4(m,\alpha)+\sigma}\,. \tag{10.12}$$

For $m \geq 1$, there exist real, even, reversibility preserving, invertible maps Φ_{m-1} of the form

$$\Phi_{m-1} := \mathbb{I}_2 + \Psi_{m-1}\,, \qquad \Psi_{m-1} := \begin{pmatrix} 0 & \psi_{m-1}(\varphi,x,D) \\ \overline{\psi_{m-1}(\varphi,x,D)} & 0 \end{pmatrix}\,, \tag{10.13}$$

with $\psi_{m-1}(\varphi,x,D)$ in $OPS^{-\frac{m-1}{2}-\frac{1}{2}}$, such that

$$\mathcal{L}_5^{(m)} = \Phi_{m-1}^{-1}\mathcal{L}_5^{(m-1)}\Phi_{m-1}\,. \tag{10.14}$$

Initialization. The real, even and reversible operator $\mathcal{L}_5^{(0)} = \mathcal{L}_5$ in (10.1) satisfies the assumptions (10.8)-(10.12) for $m = 0$ by (10.6)-(10.7).

Inductive step. We conjugate $\mathcal{L}_5^{(m)}$ in (10.8) by a real operator of the form (see (10.13))

$$\Phi_m := \mathbb{I}_2 + \Psi_m\,,\ \Psi_m := \begin{pmatrix} 0 & \psi_m(\varphi,x,D) \\ \overline{\psi_m(\varphi,x,D)} & 0 \end{pmatrix}\,,\ \psi_m(\varphi,x,D) := \mathrm{Op}(\psi_m) \in OPS^{-\frac{m}{2}-\frac{1}{2}}\,. \tag{10.15}$$

We compute

$$\begin{aligned}\mathcal{L}_5^{(m)}\Phi_m &= \Phi_m\big(\omega\cdot\partial_\varphi + \mathrm{i}a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma + a_8\mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_5^{(m)}\big) \\ &\quad + \big[\mathrm{i}a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma + a_8\mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_5^{(m)}, \Psi_m\big] + (\omega\cdot\partial_\varphi\Psi_m) + \mathcal{Q}_5^{(m)} + \mathcal{Q}_5^{(m)}\Psi_m\,.\end{aligned} \tag{10.16}$$

In the next lemma we choose Ψ_m to decrease the order of the off-diagonal operator $\mathcal{Q}_5^{(m)}$.

Lemma 10.1. *Let*

$$\psi_m(\varphi,x,\xi) := \begin{cases} -\dfrac{\chi(\xi)q_m(\varphi,x,\xi)}{2\mathrm{i}a_7(\varphi,x)|\xi|^{\frac12}\tanh^{\frac12}(\mathtt{h}|\xi|)} & \text{if} \quad |\xi| > \frac13\,, \\ 0 & \text{if} \quad |\xi| \leq \frac13\,, \end{cases} \qquad \psi_m \in S^{-\frac{m}{2}-\frac12}\,, \tag{10.17}$$

where the cut-off function χ is defined in (2.16). *Then the operator Ψ_m in* (10.15) *solves*

$$\mathrm{i}\big[a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma, \Psi_m\big] + \mathcal{Q}_5^{(m)} = \mathcal{Q}_{\psi_m} \tag{10.18}$$

where

$$\mathcal{Q}_{\psi_m} := \begin{pmatrix} 0 & q_{\psi_m}(\varphi,x,D) \\ \overline{q_{\psi_m}(\varphi,x,D)} & 0 \end{pmatrix}, \quad q_{\psi_m} \in S^{-\frac{m}{2}-1}\,. \tag{10.19}$$

Moreover, there exists $\sigma(k_0,\tau,\nu) > 0$ such that, for any $\alpha > 0$, if (6.9) *holds with $\mu_0 \geq \aleph_4(m,\alpha+1)+\alpha+\frac{m}{2}+\sigma+4$, then*

$$|q_{\psi_m}(\varphi,x,D)|_{-\frac{m}{2}-1,s,\alpha}^{k_0,\gamma} \lesssim_{s,\alpha} \varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|_{s+\aleph_4(m,\alpha+1)+\frac{m}{2}+\alpha+\sigma+4}^{k_0,\gamma}\big)\,. \tag{10.20}$$

The map Ψ_m is real, even, reversibility preserving and

$$|\psi_m(\varphi,x,D)|_{-\frac{m}{2}-\frac12,s,\alpha}^{k_0,\gamma} \lesssim_{m,s,\alpha} \varepsilon\gamma^{-1}\big(1+\|\mathfrak{I}_0\|_{s+\sigma+\aleph_4(m,\alpha)}^{k_0,\gamma}\big)\,, \tag{10.21}$$

$$|\Delta_{12}\psi_m(\varphi,x,D)|_{-\frac{m}{2}-\frac12,s_1,\alpha} \lesssim_{m,s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\sigma+\aleph_4(m,\alpha)}\,, \tag{10.22}$$

$$|\Delta_{12}q_{\psi_m}(\varphi,x,D)|_{-\frac{m}{2}-1,s_1,\alpha} \lesssim_{m,s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1+\aleph_4(m,\alpha+1)+\frac{m}{2}+\alpha+\sigma+4}\,. \tag{10.23}$$

Proof. We first note that in (10.17) the denominator $a_7|\xi|^{\frac12}\tanh(\mathtt{h}|\xi|)^{\frac12} \geq c|\xi|^{\frac12}$ with $c > 0$ for all $|\xi| \geq 1/3$, since $a_7 - 1 = O(\varepsilon\gamma^{-1})$ by (9.17) and (6.9). Thus the symbol ψ_m is well defined and estimate (10.21) follows by (10.17), (2.46) and (10.11), (9.17), Lemma 2.6, (6.9). Recalling the definition (10.2) of Σ, the vector valued commutator $\mathrm{i}[a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma, \Psi_m]$ is

$$\mathrm{i}\big[a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma, \Psi_m\big] = \begin{pmatrix} 0 & A \\ \overline{A} & 0 \end{pmatrix}\,, \qquad A := \mathrm{i}\big(a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\mathrm{Op}(\psi_m) + \mathrm{Op}(\psi_m)a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\big)\,. \tag{10.24}$$

By (10.24), in order to solve (10.18) with a remainder $\mathcal{Q}_{\psi_m} \in OPS^{-\frac{m}{2}-1}$ as in (10.19), we have to solve

$$
\mathrm{i} a_7 |D|^{\frac12} T_{\mathtt{h}}^{\frac12} \mathrm{Op}(\psi_m) + \mathrm{i}\mathrm{Op}(\psi_m) a_7 |D|^{\frac12} T_{\mathtt{h}}^{\frac12} + \mathrm{Op}(q_m) = \mathrm{Op}(q_{\psi_m}) \in OPS^{-\frac{m}{2}-1} . \tag{10.25}
$$

By (2.42), applied with $N=1$, $A = a_7 |D|^{\frac12} T_{\mathtt{h}}^{\frac12}$, $B = \mathrm{Op}(\psi_m)$, and (2.31), we have the expansion

$$
a_7 |D|^{\frac12} T_{\mathtt{h}}^{\frac12} \mathrm{Op}(\psi_m) + \mathrm{Op}(\psi_m) a_7 |D|^{\frac12} T_{\mathtt{h}}^{\frac12} = \mathrm{Op}\big(2 a_7 |\xi|^{\frac12} \tanh^{\frac12}(\mathtt{h}|\xi|) \psi_m\big) + \mathrm{Op}(\mathfrak{q}_{\psi_m}) \tag{10.26}
$$

where, using that $a_7 \chi(\xi) |\xi|^{\frac12} \tanh^{\frac12}(\mathtt{h}\chi(\xi)|\xi|) \in S^{\frac12}$ and $\psi_m \in S^{-\frac{m}{2}-\frac12}$, the symbol

$$
\mathfrak{q}_{\psi_m} = r_{1,AB} + r_{1,BA} + 2 a_7 |\xi|^{\frac12} \big(\tanh^{\frac12}(\mathtt{h}\chi(\xi)|\xi|)\chi(\xi) - \tanh^{\frac12}(\mathtt{h}|\xi|)\big) \psi_m \in S^{-\frac{m}{2}-1} , \tag{10.27}
$$

recalling that $1-\chi(\xi) \in S^{-\infty}$ by (2.16). The symbol ψ_m in (10.17) is the solution of

$$
2\mathrm{i} a_7 |\xi|^{\frac12} \tanh^{\frac12}(\mathtt{h}|\xi|) \psi_m + \chi(\xi) q_m = 0 , \tag{10.28}
$$

and therefore, by (10.26)-(10.28), the remainder q_{ψ_m} in (10.25) is

$$
q_{\psi_m} = \mathrm{i}\mathfrak{q}_{\psi_m} + (1-\chi(\xi)) q_m \in S^{-\frac{m}{2}-1} . \tag{10.29}
$$

This proves (10.18)-(10.19). We now prove (10.20). We first estimate (10.27). By (2.45) (applied with $N=1$, $A = a_7 |D|^{\frac12} T_{\mathtt{h}}^{\frac12}$, $B = \mathrm{Op}(\psi_m)$, $m = 1/2$, $m' = -\frac{m}{2} - \frac12$ and also by inverting the role of A and B), and the estimates (10.21), (10.4), (6.9) we have $|\mathfrak{q}_{\psi_m}(\varphi,x,D)|^{k_0,\gamma}_{-\frac{m}{2}-1,s,\alpha} \lesssim_{m,s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha+1)+\frac{m}{2}+\alpha+4}\big)$ and the estimate (10.20) for $q_{\psi_m}(\varphi,x,D)$ follows by (10.29) using (10.11), recalling that $1-\chi(\xi) \in S^{-\infty}$ and by applying (2.46) with $g(D) = 1-\chi(D)$ and $A = q_m(\varphi,x,D)$. Bounds (10.22)-(10.23) follow by similar arguments and by a repeated use of the triangular inequality.

Finally, the map Ψ_m defined by (10.15), (10.17) is real, even and reversibility preserving because $\mathcal{Q}_5^{(m)}$ is real, even, reversible and a_7 is even(φ)even(x) (see (9.16)). ☐

For $\varepsilon\gamma^{-1}$ small enough, by (10.21) and (6.9) the operator Φ_m is invertible, and, by Lemma 2.13,

$$
|\Phi_m^{-1} - \mathbb{I}_2|^{k_0,\gamma}_{0,s,\alpha} \lesssim_{s,\alpha} |\Psi_m|^{k_0,\gamma}_{0,s,\alpha} \lesssim_{s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m,\alpha)}\big) . \tag{10.30}
$$

By (10.16) and (10.18), the conjugated operator is

$$
\mathcal{L}_5^{(m+1)} := \Phi_m^{-1} \mathcal{L}_5^{(m)} \Phi_m = \omega\cdot\partial_\varphi + \mathrm{i} a_7 |D|^{\frac12} T_{\mathtt{h}}^{\frac12} \Sigma + a_8 \mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_5^{(m)} + \breve{\mathcal{P}}_{m+1} \tag{10.31}
$$

where $\breve{\mathcal{P}}_{m+1} := \Phi_m^{-1} \mathcal{P}^*_{m+1}$ and

$$
\mathcal{P}^*_{m+1} := \mathcal{Q}_{\psi_m} + \big[\mathrm{i}\Pi_0, \Psi_m\big] + \big[a_8 \mathbb{I}_2 + \mathcal{P}_5^{(m)}, \Psi_m\big] + (\omega\cdot\partial_\varphi \Psi_m) + \mathcal{Q}_5^{(m)} \Psi_m . \tag{10.32}
$$

Thus (10.14) at order $m+1$ is proved. Note that $\breve{\mathcal{P}}_{m+1}$ and Π_0 are the only operators in (10.31) containing off-diagonal terms.

Lemma 10.2. *The operator $\breve{\mathcal{P}}_{m+1} \in OPS^{-\frac{m}{2}-\frac12}$. Furthermore, for any $\alpha > 0$, assuming (6.9) with $\mu_0 \geq \sigma + \aleph_4(m+1,\alpha)$, the following estimates hold:*

$$
|\breve{\mathcal{P}}_{m+1}|^{k_0,\gamma}_{-\frac{m}{2}-\frac12,s,\alpha} \lesssim_{m,s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma+\aleph_4(m+1,\alpha)}\big) , \quad \forall s \geq s_0 , \tag{10.33}
$$

$$
|\Delta_{12}\breve{\mathcal{P}}_{m+1}|_{-\frac{m}{2}-\frac12,s_1,\alpha} \lesssim_{m,s_1,\alpha} \varepsilon\gamma^{-1} \|\Delta_{12} i\|_{s_1+\sigma+\aleph_4(m+1,\alpha)} \tag{10.34}
$$

where the constant $\aleph_4(m+1,\alpha)$ is defined in (10.10).

Proof. Use Lemma 10.1, (10.9), (10.15), (2.44), (10.5), (10.11), (10.12), (2.38), (10.32), (10.30). ☐

The operator $\mathcal{L}_5^{(m+1)}$ in (10.31) has the same form (10.8) as $\mathcal{L}_5^{(m)}$ with diagonal operators $\mathcal{P}_5^{(m+1)}$ and off-diagonal operators $\mathcal{Q}_5^{(m+1)}$ like in (10.9), with $\mathcal{P}_5^{(m+1)} + \mathcal{Q}_5^{(m+1)} = \mathcal{P}_5^{(m)} + \breve{\mathcal{P}}_{m+1}$, satisfying (10.11)-(10.12) at the step $m+1$ thanks to (10.33)-(10.34) and (10.11)-(10.12) at the step m. This proves the inductive claim. Applying it $2M$ times (the constant M will be fixed in (14.8)), we derive the following lemma.

Lemma 10.3. *For any $\alpha > 0$, assuming* (6.9) *with $\mu_0 \geq \aleph_5(M,\alpha) + \sigma$ where the constant $\aleph_5(M,\alpha) := \aleph_4(2M,\alpha)$ is defined recursively by* (10.10)*, the following holds. The real, even, reversibility preserving, invertible map*

$$\mathbf{\Phi}_M := \Phi_0 \circ \ldots \circ \Phi_{2M-1} \tag{10.35}$$

where Φ_m, $m = 0, \ldots, 2M-1$, are defined in (10.15)*, satisfies*

$$|\!|\!|\mathbf{\Phi}_M^{\pm 1} - \mathbb{I}_2|\!|\!|_{0,s,0}^{k_0,\gamma}\,,\ |\!|\!|(\mathbf{\Phi}_M^{\pm 1} - \mathbb{I}_2)^*|\!|\!|_{0,s,0}^{k_0,\gamma} \lesssim_{s,M} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_5(M,0)}^{k_0,\gamma}\big)\,, \quad \forall s \geq s_0\,, \tag{10.36}$$

$$|\!|\!|\Delta_{12}\mathbf{\Phi}_M^{\pm 1}|\!|\!|_{0,s_1,0}\,,\ |\!|\!|\Delta_{12}(\mathbf{\Phi}_M^{\pm 1})^*|\!|\!|_{0,s_1,0} \lesssim_{M,s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma+\aleph_5(M,0)}\,. \tag{10.37}$$

The map $\mathbf{\Phi}_M$ conjugates $\mathcal{L}_5$ to the real, even and reversible operator

$$\mathcal{L}_6 := \mathbf{\Phi}_M^{-1}\mathcal{L}_5\mathbf{\Phi}_M = \omega\cdot\partial_\varphi + \mathrm{i}a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma + a_8\mathbb{I}_2 + \mathrm{i}\Pi_0 + \mathcal{P}_6 + \mathcal{Q}_6 \tag{10.38}$$

where the functions a_7, a_8 are defined in (9.14), (10.1)*, and*

$$\mathcal{P}_6 := \begin{pmatrix} P_6 & 0 \\ 0 & \overline{P_6}\end{pmatrix} \in OPS^{-\frac12}\,, \quad \mathcal{Q}_6 := \begin{pmatrix} 0 & Q_6 \\ \overline{Q_6} & 0\end{pmatrix} \in OPS^{-M} \tag{10.39}$$

given by $\mathcal{P}_6 := \mathcal{P}_5^{(2M)}$, $\mathcal{Q}_6 := \mathcal{Q}_5^{(2M)}$ in (10.8)-(10.9) *for $m = 2M$, satisfy*

$$|\!|\!|\mathcal{P}_6|\!|\!|_{-\frac12,s,\alpha}^{k_0,\gamma} + |\!|\!|\mathcal{Q}_6|\!|\!|_{-M,s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma+\aleph_5(M,\alpha)}^{k_0,\gamma}\big)\,, \quad \forall s \geq s_0\,, \tag{10.40}$$

$$|\!|\!|\Delta_{12}\mathcal{P}_6|\!|\!|_{-\frac12,s_1,\alpha} + |\!|\!|\Delta_{12}\mathcal{Q}_6|\!|\!|_{-M,s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma+\aleph_5(M,\alpha)}\,. \tag{10.41}$$

Proof. We use (10.11), (10.12), (10.15), (10.21), (2.44), (10.30) and Lemma 2.12. □

11 Reduction of the order $1/2$

We have obtained the operator $\mathcal{L}_6$ in (10.38), where $\mathcal{P}_6$ is in $OPS^{-\frac12}$ 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 $\mathrm{i}a_7(\varphi,x)|D|^{\frac12}T_{\mathtt{h}}^{\frac12}\Sigma$. To this end, we study how the operator $\mathcal{L}_6$ transforms under the action of the flow $\Phi(\tau) := \Phi(\tau,\varphi)$

$$\begin{cases}\partial_\tau\Phi(\tau) = \mathrm{i}A(\varphi)\Phi(\tau)\\ \Phi(0) = \mathrm{Id}\,,\end{cases} \qquad A(\varphi) := \beta(\varphi,x)|D|^{\frac12} \tag{11.1}$$

where the function $\beta(\varphi,x)$ is a real valued smooth function, which will be defined in (11.19). Since $\beta(\varphi,x)$ is real valued, usual energy estimates imply that the flow $\Phi(\tau,\varphi)$ is a bounded operator on Sobolev spaces satisfying tame estimates, see Section 2.7.

Let $\Phi := \Phi(\varphi) := \Phi(1,\varphi)$. Note that $\Phi^{-1} = \overline{\Phi}$ (see Section 2.7) and

$$\Phi\pi_0 = \pi_0 = \Phi^{-1}\pi_0\,. \tag{11.2}$$

We write the operator $\mathcal{L}_6$ in (10.38) 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 (10.2), Q_6 in (10.39), and

$$P_6^{(0)} := P_6^{(0)}(\varphi,x,D) := \mathrm{i}a_7|D|^{\frac12}T_{\mathtt{h}}^{\frac12} + a_8 + P_6 \tag{11.3}$$

with P_6 defined in (10.39). Conjugating $\mathcal{L}_6$ with the real operator

$$\boldsymbol{\Phi} := \begin{pmatrix} \Phi & 0 \\ 0 & \overline{\Phi} \end{pmatrix} \tag{11.4}$$

we get, since $\boldsymbol{\Phi}^{-1}\Pi_0\boldsymbol{\Phi} = \Pi_0\boldsymbol{\Phi}$ by (11.2),

$$\mathcal{L}_7 := \boldsymbol{\Phi}^{-1}\mathcal{L}_6\boldsymbol{\Phi} = \omega\cdot\partial_\varphi + \boldsymbol{\Phi}^{-1}\big(\omega\cdot\partial_\varphi\boldsymbol{\Phi}\big) + \mathrm{i}\Pi_0\boldsymbol{\Phi} + \begin{pmatrix} \Phi^{-1}P_6^{(0)}\Phi & \Phi^{-1}Q_6\overline{\Phi} \\ \overline{\Phi}^{-1}\overline{Q_6}\Phi & \overline{\Phi}^{-1}\overline{P_6^{(0)}}\overline{\Phi} \end{pmatrix}. \tag{11.5}$$

Let us study the operator

$$L_7 := \omega\cdot\partial_\varphi + \Phi^{-1}\big(\omega\cdot\partial_\varphi\Phi\big) + \Phi^{-1}P_6^{(0)}\Phi\,. \tag{11.6}$$

Analysis of the term $\Phi^{-1}P_6^{(0)}\Phi$. Recalling (11.1), 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

$$\begin{aligned}\Phi(1,\varphi)^{-1}P_6^{(0)}\Phi(1,\varphi) = P_6^{(0)} - \mathrm{i}[A,P_6^{(0)}] + \sum_{n=2}^{2M}\frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}^n_{A(\varphi)}P_6^{(0)} \\ + \frac{(-\mathrm{i})^{2M+1}}{(2M)!}\int_0^1(1-\tau)^{2M}\Phi(\tau,\varphi)^{-1}\mathrm{Ad}^{2M+1}_{A(\varphi)}P_6^{(0)}\Phi(\tau,\varphi)\,d\tau\,.\end{aligned} \tag{11.7}$$

The order M of the expansion will be fixed in (14.8). We remark that (11.7) 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 (11.1) and (11.3),

$$-\mathrm{i}\big[A,P_6^{(0)}\big] = \big[\beta|D|^{\frac12},a_7|D|^{\frac12}\big] + \big[\beta|D|^{\frac12},a_7|D|^{\frac12}(T_{\mathtt h}^{\frac12}-\mathrm{Id})\big] - \mathrm{i}\big[\beta|D|^{\frac12},a_8+P_6\big]\,. \tag{11.8}$$

Moreover, by (2.47), (2.48) one has

$$\begin{aligned}[\beta|D|^{\frac12},a_7|D|^{\frac12}] &= \mathrm{Op}\Big(-\mathrm{i}\{\beta\chi(\xi)|\xi|^{\frac12},a_7\chi(\xi)|\xi|^{\frac12}\} + \mathtt{r}_2(\beta\chi(\xi)|\xi|^{\frac12},a_7\chi(\xi)|\xi|^{\frac12})\Big) \\ &= \mathrm{i}\big((\partial_x\beta)a_7-\beta(\partial_x a_7)\big)\mathrm{Op}\Big(\frac12\chi^2(\xi)\mathrm{sign}(\xi)+\chi(\xi)\partial_\xi\chi(\xi)|\xi|\Big) + \mathrm{Op}\big(\mathtt{r}_2(\beta\chi(\xi)|\xi|^{\frac12},a_7\chi(\xi)|\xi|^{\frac12})\big)\end{aligned} \tag{11.9}$$

where the symbol $\mathtt{r}_2(\beta\chi(\xi)|\xi|^{\frac12},a_7\chi(\xi)|\xi|^{\frac12}) \in S^{-1}$ is defined according to (2.49). Therefore (11.8), (11.9) imply the expansion

$$-\mathrm{i}\big[A,P_6^{(0)}\big] = -\frac12\big((\partial_x\beta)a_7-\beta(\partial_x a_7)\big)\mathcal{H} + R_{A,P_6^{(0)}} \tag{11.10}$$

where the remainder

$$\begin{aligned}R_{A,P_6^{(0)}} &:= \mathrm{i}\big((\partial_x\beta)a_7-\beta(\partial_x a_7)\big)\mathrm{Op}\Big(\chi(\xi)\partial_\xi\chi(\xi)|\xi| + \frac12(\chi^2(\xi)-\chi(\xi))\mathrm{sign}(\xi)\Big) \\ &+ \mathrm{Op}\big(\mathtt{r}_2(\beta\chi(\xi)|\xi|^{\frac12},a_7\chi(\xi)|\xi|^{\frac12})\big) + \big[\beta|D|^{\frac12},a_7|D|^{\frac12}(T_{\mathtt h}^{\frac12}-\mathrm{Id})\big] - \mathrm{i}\big[\beta|D|^{\frac12},a_8+P_6\big]\end{aligned} \tag{11.11}$$

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

$$
\begin{aligned}
\Phi(1,\varphi)^{-1}\circ\omega\cdot\partial_\varphi\circ\Phi(1,\varphi) &= \omega\cdot\partial_\varphi - \mathrm{i}\mathrm{Ad}_{A(\varphi)}\omega\cdot\partial_\varphi + \frac{(-\mathrm{i})^2}{2}\mathrm{Ad}^2_{A(\varphi)}\omega\cdot\partial_\varphi + \sum_{n=3}^{2M+1}\frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}^n_{A(\varphi)}\omega\cdot\partial_\varphi \\
&\quad + \frac{(-\mathrm{i})^{2M+2}}{(2M+1)!}\int_0^1 (1-\tau)^{2M+1}\Phi(\tau,\varphi)^{-1}\big(\mathrm{Ad}^{2M+2}_{A(\varphi)}\omega\cdot\partial_\varphi\big)\Phi(\tau,\varphi)\,d\tau. \qquad (11.12)
\end{aligned}
$$

We compute the commutator

$$
\mathrm{Ad}_{A(\varphi)}\omega\cdot\partial_\varphi = \big[A(\varphi),\omega\cdot\partial_\varphi\big] = -(\omega\cdot\partial_\varphi A(\varphi)) \stackrel{(11.1)}{=} -(\omega\cdot\partial_\varphi\beta(\varphi,x))|D|^{1/2} \qquad (11.13)
$$

and, using (2.47), (2.48),

$$
\begin{aligned}
\mathrm{Ad}^2_{A(\varphi)}\omega\cdot\partial_\varphi &= \big[(\omega\cdot\partial_\varphi A(\varphi)),A(\varphi)\big] = \big[(\omega\cdot\partial_\varphi\beta)|D|^{\frac12},\beta|D|^{\frac12}\big] \\
&= \mathrm{Op}\Big(-\mathrm{i}\big\{(\omega\cdot\partial_\varphi\beta)\chi(\xi)|\xi|^{\frac12},\beta\chi(\xi)|\xi|^{\frac12}\big\} + \mathtt{r}_2((\omega\cdot\partial_\varphi\beta)\chi(\xi)|\xi|^{\frac12},\beta\chi(\xi)|\xi|^{\frac12})\Big)\,.
\end{aligned}
$$

According to (2.48) 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

$$
\frac{(-\mathrm{i})^2}{2}\mathrm{Ad}^2_{A(\varphi)}\omega\cdot\partial_\varphi = \frac14\big(\beta\,\omega\cdot\partial_\varphi\beta_x - \beta_x\,\omega\cdot\partial_\varphi\beta\big)\mathcal{H} + R_{A,\omega\cdot\partial_\varphi} \qquad (11.14)
$$

where

$$
\begin{aligned}
R_{A,\omega\cdot\partial_\varphi} := & -\frac{\mathrm{i}}{4}\left(\beta\,\omega\cdot\partial_\varphi\beta_x - \beta_x\,\omega\cdot\partial_\varphi\beta\right)\mathrm{Op}\Big((\chi(\xi)^2-\chi(\xi))\mathrm{sign}(\xi) + 2\chi(\xi)\partial_\xi\chi(\xi)|\xi|\Big) \\
& -\frac12\mathrm{Op}\left(\mathtt{r}_2\big((\omega\cdot\partial_\varphi\beta)\chi(\xi)|\xi|^{\frac12},\beta\chi(\xi)|\xi|^{\frac12}\big)\right). \qquad (11.15)
\end{aligned}
$$

is an operator in OPS^{-1} (the first line of (11.15) 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 (11.12), (11.13) and (11.14), we get

$$
\begin{aligned}
\Phi(1,\varphi)^{-1}\circ\omega\cdot\partial_\varphi\circ\Phi(1,\varphi) &= \omega\cdot\partial_\varphi + \mathrm{i}(\omega\cdot\partial_\varphi\beta)(\varphi,x)|D|^{\frac12} + \frac14\big(\beta(\omega\cdot\partial_\varphi\beta_x) - \beta_x(\omega\cdot\partial_\varphi\beta)\big)\mathcal{H} + R_{A,\omega\cdot\partial_\varphi} \\
&\quad - \sum_{n=3}^{2M+1}\frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}^{n-1}_{A(\varphi)}\big(\omega\cdot\partial_\varphi A(\varphi)\big) \qquad (11.16) \\
&\quad - \frac{(-\mathrm{i})^{2M+2}}{(2M+1)!}\int_0^1(1-\tau)^{2M+1}\Phi(\tau,\varphi)^{-1}\big(\mathrm{Ad}^{2M+1}_{A(\varphi)}\big(\omega\cdot\partial_\varphi A(\varphi)\big)\big)\Phi(\tau,\varphi)\,d\tau\,.
\end{aligned}
$$

This is an expansion in operators with decreasing orders (and size).

In conclusion, by (11.6), (11.7), (11.3), (11.10), (11.16), the term of order $|D|^{\frac12}$ in L_7 in (11.6) is

$$
\mathrm{i}\big((\omega\cdot\partial_\varphi\beta) + a_7 T_{\mathtt{h}}^{\frac12}\big)|D|^{\frac12}\,. \qquad (11.17)
$$

Choice of the functions $\beta(\varphi,x)$ **and** $\alpha(x)$**.** We choose the function $\beta(\varphi,x)$ such that

$$
(\omega\cdot\partial_\varphi\beta)(\varphi,x) + a_7(\varphi,x) = \langle a_7\rangle_\varphi(x)\,, \quad \langle a_7\rangle_\varphi(x) := \frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} a_7(\varphi,x)\,d\varphi. \qquad (11.18)
$$

For all $\omega\in\mathtt{DC}(\gamma,\tau)$, the solution of (11.18) is the periodic function

$$
\beta(\varphi,x) := -(\omega\cdot\partial_\varphi)^{-1}\big(a_7(\varphi,x) - \langle a_7\rangle_\varphi(x)\big)\,, \qquad (11.19)
$$

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 11.1. *The real valued function* β *defined in* (11.19) *is odd*(φ)*even*(x). *Moreover there exists* $\sigma(k_0, \tau, \nu) > 0$ *such that, if* (6.9) *holds with* $\mu_0 \geq \sigma$, *then* β *satisfies the following estimates:*

$$\|\beta\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-2}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big), \quad \|\omega \cdot \partial_\varphi \beta\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-1}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) \tag{11.20}$$

$$\|\Delta_{12}\beta\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-2}\|\Delta_{12} i\|_{s_1+\sigma}, \quad \|\omega \cdot \partial_\varphi \Delta_{12}\beta\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1+\sigma}. \tag{11.21}$$

Proof. The function a_7 is even(φ)even(x) (see (9.16)), and therefore, by (11.19), β is odd(φ)even(x). Estimates (11.20)-(11.21) follow by (11.18), (11.19), (10.4) and Lemma 2.5. □

By (9.14), (8.7), (8.8) 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 (8.1)) so that

$$\langle a_7 \rangle_\varphi(x) = \mathtt{m}_{\frac{1}{2}} \tag{11.22}$$

is independent of x, for some real constant $\mathtt{m}_{\frac{1}{2}}$. This is equivalent to solve the equation

$$\langle \sqrt{a_2 a_3}\,\rangle_\varphi(x)\,\sqrt{1 + \alpha_x(x)} = \mathtt{m}_{\frac{1}{2}}$$

whose solution is

$$\mathtt{m}_{\frac{1}{2}} := \Big(\frac{1}{2\pi}\int_{\mathbb{T}} \frac{dx}{\langle \sqrt{a_2 a_3}\,\rangle_\varphi^2(x)}\Big)^{-\frac{1}{2}}, \qquad \alpha(x) := \partial_x^{-1}\Big(\frac{\mathtt{m}_{\frac{1}{2}}^2}{\langle \sqrt{a_2 a_3}\,\rangle_\varphi^2(x)} - 1\Big). \tag{11.23}$$

Lemma 11.2. *The real valued function* $\alpha(x)$ *defined in* (11.23) *is odd*(x) *and* (8.2) *holds. Moreover*

$$|\mathtt{m}_{\frac{1}{2}} - 1|^{k_0,\gamma} \lesssim \varepsilon\gamma^{-1}, \quad |\Delta_{12}\mathtt{m}_{\frac{1}{2}}| \lesssim \varepsilon\gamma^{-1}\|\Delta_{12} i\|_{s_1}. \tag{11.24}$$

Proof. Since a_2, a_3 are even(x) by (7.34), the function $\alpha(x)$ defined in (11.23) is odd(x). Estimates (11.24) follow by the definition of $\mathtt{m}_{\frac{1}{2}}$ in (11.23) and (7.35), (7.37), (6.9), applying also Lemma 2.6 and (2.10). Similarly α satisfies (8.2) by (7.35), (7.37), (11.24), Lemma 2.6 and (2.10). □

By (11.18) and (11.22) the term in (11.17) reduces to

$$\mathrm{i}\big(\omega \cdot \partial_\varphi \beta(\varphi, x) + a_7(\varphi, x) T_{\mathtt{h}}^{\frac{1}{2}}\big)|D|^{\frac{1}{2}} = \mathrm{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}}|D|^{\frac{1}{2}} + \mathtt{R}_\beta \tag{11.25}$$

where $\mathtt{R}_\beta$ is the $OPS^{-\infty}$ operator defined by

$$\mathtt{R}_\beta := \mathrm{i}(\omega \cdot \partial_\varphi \beta)(\mathrm{Id} - T_{\mathtt{h}}^{\frac{1}{2}})|D|^{\frac{1}{2}}. \tag{11.26}$$

Finally, the operator L_7 in (11.6) is, in view of (11.7), (11.3), (11.10), (11.16), (11.25),

$$L_7 = \omega \cdot \partial_\varphi + \mathrm{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}}|D|^{\frac{1}{2}} + a_8 + a_9\mathcal{H} + P_7 + T_7 \tag{11.27}$$

where a_9 is the real valued function

$$a_9 := a_9(\varphi, x) := -\frac{1}{2}\big(\beta_x\, a_7 - \beta(\partial_x a_7)\big) - \frac{1}{4}\big(\beta_x\, \omega \cdot \partial_\varphi \beta - \beta\, \omega \cdot \partial_\varphi \beta_x\big), \tag{11.28}$$

P_7 is the operator in $OPS^{-1/2}$ given by

$$P_7 := R_{A, P_6^{(0)}} + R_{A, \omega\cdot\partial_\varphi} - \sum_{n=3}^{2M+1} \frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}_{A(\varphi)}^{n-1}\big(\omega \cdot \partial_\varphi A(\varphi)\big) + \sum_{n=2}^{2M} \frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}_{A(\varphi)}^n P_6^{(0)} + P_6 + \mathtt{R}_\beta \tag{11.29}$$

(the operators $R_{A,P_6^{(0)}}, R_{A,\omega\cdot\partial_\varphi}, P_6, \mathtt{R}_\beta$ are defined respectively in (11.11), (11.15), (10.39), (11.26)), and

$$
\begin{aligned}
T_7 := & -\frac{(-\mathrm{i})^{2M+2}}{(2M+1)!}\int_0^1 (1-\tau)^{2M+1}\Phi(\tau,\varphi)^{-1}\big(\mathrm{Ad}_{A(\varphi)}^{2M+1}\big(\omega\cdot\partial_\varphi A(\varphi)\big)\big)\Phi(\tau,\varphi)\,d\tau \\
& + \frac{(-\mathrm{i})^{2M+1}}{(2M)!}\int_0^1 (1-\tau)^{2M}\Phi(\tau,\varphi)^{-1}\mathrm{Ad}_{A(\varphi)}^{2M+1}P_6^{(0)}\Phi(\tau,\varphi)\,d\tau
\end{aligned}
\tag{11.30}
$$

(T_7 stands for "tame remainders", namely remainders satisfying tame estimates together with their derivatives, see (11.39), without controlling their pseudo-differential structure). In conclusion, we have the following lemma.

Lemma 11.3. *Let $\beta(\varphi,x)$ and $\alpha(x)$ be the functions defined in* (11.19) *and* (11.23)*. Then $\mathcal{L}_7 := \boldsymbol{\Phi}^{-1}\mathcal{L}_6\boldsymbol{\Phi}$ in* (11.5) *is the real, even and reversible operator*

$$
\mathcal{L}_7 = \omega\cdot\partial_\varphi + \mathrm{i}\mathtt{m}_{\frac12}T_{\mathtt{h}}^{\frac12}|D|^{\frac12}\Sigma + \mathrm{i}\Pi_0 + (a_8 + a_9\mathcal{H})\mathbb{I}_2 + \mathcal{P}_7 + \mathcal{T}_7 \tag{11.31}
$$

where $\mathtt{m}_{\frac12}$ is the real constant defined in (11.23)*, a_8, a_9 are the real valued functions in* (10.1), (11.28)*,*

$$
a_8 = odd(\varphi)even(x)\,, \qquad a_9 = odd(\varphi)odd(x)\,, \tag{11.32}
$$

and $\mathcal{P}_7$, $\mathcal{T}_7$ are the real operators

$$
\mathcal{P}_7 := \begin{pmatrix} P_7 & 0 \\ 0 & \overline{P_7}\end{pmatrix} \in OPS^{-\frac12}\,, \qquad \mathcal{T}_7 := \mathrm{i}\Pi_0(\boldsymbol{\Phi}-\mathbb{I}_2) + \boldsymbol{\Phi}^{-1}\mathcal{Q}_6\boldsymbol{\Phi} + \begin{pmatrix} T_7 & 0 \\ 0 & \overline{T_7}\end{pmatrix}\,, \tag{11.33}
$$

where P_7 is defined in (11.29) *and T_7 in* (11.30)*.*

Proof. Formula (11.31) follows by (11.5) and (11.27). By Lemma 11.1 the real function β is odd(φ)even(x). Thus, by Sections 2.5 and 2.7, the flow map $\boldsymbol{\Phi}$ in (11.4) 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 (9.16) and a_9 defined in (11.28) is odd(φ)odd(x). □

Note that formulas (11.28) and (11.33) (via (11.29), (11.30)) 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 (11.31) 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 11.4. *There exists $\sigma(k_0,\tau,\nu) > 0$ such that, if* (6.9) *holds with $\mu_0 \geq \sigma$, then*

$$
\|a_9\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-2}(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma})\,, \quad \forall s \geq s_0\,, \qquad \|\Delta_{12}a_9\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-2}\|\Delta_{12}i\|_{s_1+\sigma}\,. \tag{11.34}
$$

For any $s \geq s_0$ there exists $\delta(s) > 0$ small enough such that if $\varepsilon\gamma^{-2} \leq \delta(s)$, then

$$
\|(\Phi^{\pm 1} - \mathrm{Id})h\|_s^{k_0,\gamma}, \|(\Phi^* - \mathrm{Id})h\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-2}\big(\|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\|h\|_{s_0+\sigma}^{k_0,\gamma}\big)\,, \tag{11.35}
$$

$$
\|\Delta_{12}\Phi^{\pm1}h\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-2}\|\Delta_{12}i\|_{s_1+\sigma}\|h\|_{s_1+\frac12}\,. \tag{11.36}
$$

The pseudo-differential operator $\mathcal{P}_7$ defined in (11.33) *is in $OPS^{-\frac12}$. Moreover for any $M, \alpha > 0$, there exists a constant $\aleph_6(M,\alpha) > 0$ such that assuming* (6.9) *with $\mu_0 \geq \aleph_6(M,\alpha) + \sigma$, the following estimates hold:*

$$
|\!|\!| \mathcal{P}_7 |\!|\!|_{-\frac12, s, \alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-2}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_6(M,\alpha)+\sigma}^{k_0,\gamma}\big)\,, \tag{11.37}
$$

$$
|\!|\!| \Delta_{12}\mathcal{P}_7 |\!|\!|_{-\frac12, s_1, \alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-2}\|\Delta_{12}i\|_{s_1+\aleph_6(M,\alpha)+\sigma}\,. \tag{11.38}
$$

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 (6.9) *with $\mu_0 \geq \aleph_6'(M,\beta_0) + \sigma$, for any $m_1, m_2 \geq 0$, 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*

$$
\mathfrak{M}_{\langle D\rangle^{m_1}\partial_\varphi^\beta\mathcal{T}_7\langle D\rangle^{m_2}}(s) \lesssim_{M,S} \varepsilon\gamma^{-2}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_6'(M,\beta_0)+\sigma}\big)\,, \qquad \forall s_0 \leq s \leq S \tag{11.39}
$$

$$
\|\langle D\rangle^{m_1}\Delta_{12}\partial_\varphi^\beta\mathcal{T}_7\langle D\rangle^{m_2}\|_{\mathcal{L}(H^{s_1})} \lesssim_{M,S} \varepsilon\gamma^{-2}\|\Delta_{12}i\|_{s_1+\aleph_6'(M,\beta_0)+\sigma}\,. \tag{11.40}
$$

Proof. Estimates (11.34) for a_9 defined in (11.28) follow by (10.4), (11.20), (11.21), (2.10) and (6.9).
PROOF OF (11.35)-(11.36). It follows by applying Proposition 2.37, Lemma 2.38, estimates (11.20)-(11.21) 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 (11.37)-(11.38). First we prove (11.37), estimating each term in the definition (11.29) of P_7. The operator $A=\beta(\varphi,x)|D|^{\frac12}$ in (11.1) satisfies, by (2.46) and (11.20),

$$|A|_{\frac12,s,\alpha}^{k_0,\gamma}\lesssim_{s,\alpha}\|\beta\|_s^{k_0,\gamma}\lesssim_{s,\alpha}\varepsilon\gamma^{-2}\big(1+\|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big)\,. \tag{11.41}$$

The operator $P_6^{(0)}$ in (11.3) satisfies, by (10.4), (10.5), (2.46), (10.40),

$$|P_6^{(0)}|_{\frac12,s,\alpha}^{k_0,\gamma}\lesssim_{M,s,\alpha}1+\|\mathfrak{I}_0\|_{s+\aleph_5(M,\alpha)+\sigma}^{k_0,\gamma}\,. \tag{11.42}$$

The estimate of the term $-\sum_{n=3}^{2M+1}\frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}_{A(\varphi)}^{n-1}\big(\omega\cdot\partial_\varphi A(\varphi)\big)+\sum_{n=2}^{2M}\frac{(-\mathrm{i})^n}{n!}\mathrm{Ad}_{A(\varphi)}^{n}P_6^{(0)}$ in (11.29) then follows by (11.41), (11.42) and by applying Lemma 2.10 and the estimate (2.51). The term $\mathtt{R}_\beta\in OPS^{-\infty}$ defined in (11.26) can be estimated by (2.46) (applied with $A:=\omega\cdot\partial_\varphi\beta$, $g(D):=(T_{\mathtt{h}}^{\frac12}-\mathrm{Id})|D|^{\frac12}\in OPS^{-\infty}$) and using (11.20), (7.41). The estimate of the terms $R_{A,P_6^{(0)}},R_{A,\omega\cdot\partial_\varphi}$ in (11.29) follows by their definition given in (11.11), (11.15) and by estimates (10.4), (10.5), (10.40), (11.20), (2.10), (2.46), and Lemmata 2.10, 2.11. Since P_6 satisfies (10.40), estimate (11.37) is proved. Estimate (11.38) can be proved by similar arguments.
PROOF OF (11.39), (11.40). We estimate the term $\mathbf{\Phi}^{-1}\mathcal{Q}_6\mathbf{\Phi}$ in (11.33). 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

$$\partial_\lambda^k\partial_\varphi^\beta(\mathbf{\Phi}^{-1}\mathcal{Q}_6\mathbf{\Phi})=\sum_{\substack{\beta_1+\beta_2+\beta_3=\beta\\ k_1+k_2+k_3=k}}C(\beta_1,\beta_2,\beta_3,k_1,k_2,k_3)(\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\mathbf{\Phi}^{-1})(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}\mathcal{Q}_6)(\partial_\lambda^{k_3}\partial_\varphi^{\beta_3}\mathbf{\Phi})\,. \tag{11.43}$$

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

$$\langle D\rangle^{m_1}(\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\mathbf{\Phi}^{-1})(\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}\mathcal{Q}_6)(\partial_\lambda^{k_3}\partial_\varphi^{\beta_3}\mathbf{\Phi})\langle D\rangle^{m_2}\,. \tag{11.44}$$

We write

$$(11.44)=\Big(\langle D\rangle^{m_1}\partial_\lambda^{k_1}\partial_\varphi^{\beta_1}\mathbf{\Phi}^{-1}\langle D\rangle^{-\frac{|\beta_1|+|k_1|}{2}-m_1}\Big) \tag{11.45}$$

$$\circ\Big(\langle D\rangle^{\frac{|\beta_1|+|k_1|}{2}+m_1}\partial_\lambda^{k_2}\partial_\varphi^{\beta_2}\mathcal{Q}_6\langle D\rangle^{\frac{|\beta_3|+|k_3|}{2}+m_2}\Big) \tag{11.46}$$

$$\circ\Big(\langle D\rangle^{-m_2-\frac{|\beta_3|+|k_3|}{2}}\partial_\lambda^{k_3}\partial_\varphi^{\beta_3}\mathbf{\Phi}\langle D\rangle^{m_2}\Big)\,. \tag{11.47}$$

The terms (11.45)-(11.47) can be estimated separately. To estimate the terms (11.45) and (11.47), we apply (2.86) of Proposition 2.37, (2.88) of Lemma 2.38, and (11.20)-(11.21). The pseudo-differential operator in (11.46) is estimated in $|\ |_{0,s,0}$ norm by using (2.40), (2.44), (2.46), bounds (10.40), (10.41) 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 2.28. As a consequence, each operator in (11.44), and hence the whole operator (11.43), satisfies (11.39).

The estimates of the terms in (11.30) can be done arguing similarly, using also the estimates (2.51), (11.41)-(11.42). 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 2.36 (with $A=\mathbb{I}_2$, $B=\mathbf{\Phi}$) and (11.35), (11.20), (11.21). □

12 Reduction of the lower orders

In this section we complete the reduction of the operator $\mathcal{L}_7$ in (11.31) to constant coefficients, up to a regularizing remainder of order $|D|^{-M}$. We write

$$\mathcal{L}_7=\begin{pmatrix}L_7 & 0\\ 0 & \overline{L_7}\end{pmatrix}+\mathrm{i}\Pi_0+\mathcal{T}_7\,, \tag{12.1}$$

where

$$L_7 := \omega \cdot \partial_\varphi + \mathtt{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}} + a_8 + a_9 \mathcal{H} + P_7 \,, \tag{12.2}$$

the real valued functions a_8, a_9 are introduced in (10.1), (11.28), satisfy (11.32), and the operator $P_7 \in OPS^{-\frac{1}{2}}$ in (11.29) is even and reversible. We first conjugate the operator L_7.

12.1 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 (12.2). 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 (11.32), thus with zero average, this step removes completely the terms of order zero. Consider the transformation

$$W_0 := \mathrm{Id} + f_0(\varphi, x) + g_0(\varphi, x)\mathcal{H} \,, \tag{12.3}$$

where f_0, g_0 are real valued functions to be determined. Since $\mathcal{H}^2 = -\mathrm{Id} + \pi_0$ on the periodic functions where π_0 is defined in (2.33), one has

$$\begin{aligned} L_7 W_0 = W_0\big(\omega \cdot \partial_\varphi + \mathtt{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}\big) + (\omega \cdot \partial_\varphi f_0 + a_8 + a_8 f_0 - a_9 g_0) \\ + (\omega \cdot \partial_\varphi g_0 + a_9 + a_8 g_0 + a_9 f_0)\mathcal{H} + \breve{P}_7 \end{aligned} \tag{12.4}$$

where $\breve{P}_7 \in OPS^{-\frac{1}{2}}$ is the operator

$$\breve{P}_7 := a_9[\mathcal{H}, f_0] + a_9[\mathcal{H}, g_0]\mathcal{H} + [\mathtt{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}, W_0] + P_7 W_0 + a_9 g_0 \pi_0 \,. \tag{12.5}$$

In order to eliminate the zero order terms in (12.4) we choose the functions f_0, g_0 such that

$$\begin{cases} \omega \cdot \partial_\varphi f_0 + a_8 + a_8 f_0 - a_9 g_0 = 0 \\ \omega \cdot \partial_\varphi g_0 + a_9 + a_8 g_0 + a_9 f_0 = 0 \,. \end{cases} \tag{12.6}$$

Writing $z_0 = 1 + f_0 + \mathrm{i} g_0$, the real system (12.6) is equivalent to the complex scalar equation

$$\omega \cdot \partial_\varphi z_0 + (a_8 + \mathrm{i} a_9) z_0 = 0 \,. \tag{12.7}$$

Since a_8, a_9 are odd functions in φ, we choose, for all $\omega \in \mathtt{DC}(\gamma, \tau)$, the periodic function

$$z_0 := \exp(p_0), \quad p_0 := -(\omega \cdot \partial_\varphi)^{-1}(a_8 + \mathrm{i} a_9), \tag{12.8}$$

which solves (12.7). Thus the real functions

$$\begin{aligned} f_0 &:= \mathrm{Re}(z_0) - 1 = \exp(-(\omega \cdot \partial_\varphi)^{-1} a_8) \cos((\omega \cdot \partial_\varphi)^{-1} a_9) - 1, \\ g_0 &:= \mathrm{Im}(z_0) = -\exp(-(\omega \cdot \partial_\varphi)^{-1} a_8) \sin((\omega \cdot \partial_\varphi)^{-1} a_9) \end{aligned} \tag{12.9}$$

solve (12.6), and, for $\omega \in \mathtt{DC}(\gamma, \tau)$, equation (12.4) reduces to

$$L_7 W_0 = W_0(\omega \cdot \partial_\varphi + \mathtt{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}}) + \breve{P}_7 \,, \quad \breve{P}_7 \in OPS^{-\frac{1}{2}} \,. \tag{12.10}$$

We extend the function p_0 in (12.8) to the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ by using $(\omega \cdot \partial_\varphi)^{-1}_{ext}$ introduced in Lemma 2.5. Thus the functions z_0, f_0, g_0 in (12.8), (12.9) are defined on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ as well.

Lemma 12.1. *The real valued functions f_0, g_0 in (12.9) satisfy*

$$f_0 = even(\varphi) even(x) \,, \quad g_0 = even(\varphi) odd(x) \,. \tag{12.11}$$

Moreover, there exists $\sigma(k_0, \tau, \nu) > 0$ such that, if (6.9) holds with $\mu_0 \geq \sigma$, then

$$\|f_0\|_s^{k_0,\gamma} \,, \|g_0\|_s^{k_0,\gamma} \lesssim_s \varepsilon\gamma^{-3}\big(1 + \|\mathfrak{I}_0\|_{s+\sigma}^{k_0,\gamma}\big) \,, \quad \|\Delta_{12} f_0\|_{s_1}, \|\Delta_{12} g_0\|_{s_1} \lesssim_{s_1} \varepsilon\gamma^{-3} \|\Delta_{12} i\|_{s_1+\sigma} \,. \tag{12.12}$$

The operator W_0 defined in (12.3) is even, reversibility preserving, invertible and for any $\alpha > 0$, assuming (6.9) with $\mu_0 \geq \alpha + \sigma$, the following estimates hold:

$$|W_0^{\pm 1} - \mathrm{Id}|_{0,s,\alpha}^{k_0,\gamma} \lesssim_{s,\alpha} \varepsilon\gamma^{-3}\big(1 + \|\mathfrak{I}_0\|_{s+\alpha+\sigma}^{k_0,\gamma}\big) \,, \quad |\Delta_{12} W_0^{\pm 1}|_{0,s_1,\alpha} \lesssim_{s_1,\alpha} \varepsilon\gamma^{-3} \|\Delta_{12} i\|_{s_1+\alpha+\sigma} \,. \tag{12.13}$$

Proof. The parities in (12.11) follow by (12.9) and (11.32). Therefore W_0 in (12.3) is even and reversibility preserving. Estimates (12.12) follow by (12.9), (10.5), (11.34), (2.10), (2.17), (2.19). The operator W_0 defined in (12.3) is invertible by Lemma 2.13, (12.12), (6.9), for $\varepsilon\gamma^{-3}$ small enough. Estimates (12.13) then follow by (12.12), using (2.39), (2.46) and Lemma 2.13. $\square$

For $\omega \in \mathtt{DC}(\gamma,\tau)$, by (12.10) we obtain the even and reversible operator

$$L_7^{(1)} := W_0^{-1} L_7 W_0 = \omega\cdot\partial_\varphi + \mathtt{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}} + P_7^{(1)}\,, \qquad P_7^{(1)} := W_0^{-1}\breve{P}_7\,, \tag{12.14}$$

where $\breve{P}_7$ is the operator in $OPS^{-\frac{1}{2}}$ defined in (12.5).

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 (12.5) is defined on $\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$, and $\omega\cdot\partial_\varphi + \mathtt{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}} |D|^{\frac{1}{2}} + P_7^{(1)}$ in (12.14) is an extension of $L_7^{(1)}$ to $\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$, still denoted $L_7^{(1)}$.

Lemma 12.2. *For any $M,\alpha>0$, there exists a constant $\aleph_7^{(1)}(M,\alpha)>0$ such that if* (6.9) *holds with $\mu_0 \geq \aleph_7^{(1)}(M,\alpha)$, the remainder $P_7^{(1)}\in OPS^{-\frac{1}{2}}$, defined in* (12.14)*, satisfies*

$$\begin{aligned} |P_7^{(1)}|_{-\frac{1}{2},s,\alpha}^{k_0,\gamma} &\lesssim_{M,s,\alpha} \varepsilon\gamma^{-3}\big(1+\|\mathfrak{I}_0\|_{s+\aleph_7^{(1)}(M,\alpha)}^{k_0,\gamma}\big)\,,\\ |\Delta_{12}P_7^{(1)}|_{-\frac{1}{2},s_1,\alpha} &\lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-3}\|\Delta_{12} i\|_{s_1+\aleph_7^{(1)}(M,\alpha)}\,. \end{aligned} \tag{12.15}$$

Proof. Estimates (12.15) follow by the definition of $P_7^{(1)}$ given in (12.14), by estimates (12.12), (12.13), (11.24), (11.34), (11.37), (11.38), by applying (2.39), (2.44), (2.46), (2.50) and using also Lemma 2.16. The fact that $P_7^{(1)}$ has size $\varepsilon\gamma^{-3}$ is due to the term $[\mathtt{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}}|D|^{\frac{1}{2}}, W_0] = [\mathtt{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}}|D|^{\frac{1}{2}}, W_0 - \mathrm{Id}]$, because $\mathtt{m}_{\frac{1}{2}} = 1 + O(\varepsilon\gamma^{-1})$ and $W_0 - \mathrm{Id} = O(\varepsilon\gamma^{-3})$. $\square$

We underline that the operator $L_7^{(1)}$ in (12.14) does not contain terms of order zero.

12.2 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 (12.14). The constant M will be fixed in (14.8). 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

  $$L_7^{(m)} := \omega\cdot\partial_\varphi + \Lambda_m(D) + P_7^{(m)}\,, \quad P_7^{(m)} \in OPS^{-\frac{m}{2}}\,, \tag{12.16}$$

  where

  $$\Lambda_m(D) := \mathtt{i}\mathtt{m}_{\frac{1}{2}} T_{\mathtt{h}}^{\frac{1}{2}}|D|^{\frac{1}{2}} + r_m(D)\,, \qquad r_m(D)\in OPS^{-\frac{1}{2}}\,. \tag{12.17}$$

  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 (6.9) holds with $\mu_0\geq\aleph_7^{(m)}(M,\alpha)$, then the following estimates hold:

  $$|r_m(D)|_{-\frac{1}{2},s,\alpha}^{k_0,\gamma} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+1)}\,, \quad |\Delta_{12} r_m(D)|_{-\frac{1}{2},s_1,\alpha} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+1)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m)}(M,\alpha)}\,, \tag{12.18}$$

  $$|P_7^{(m)}|_{-\frac{m}{2},s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+2)}\big(1+\|\mathfrak{I}_0\|_{s+\aleph_7^{(m)}(M,\alpha)}^{k_0,\gamma}\big)\,, \tag{12.19}$$

  $$|\Delta_{12}P_7^{(m)}|_{-\frac{m}{2},s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+2)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m)}(M,\alpha)}\,. \tag{12.20}$$

Note that by (12.17), using (11.24), (12.18) and (2.40) (applied for $g(D) = T_{\mathtt h}^{\frac12}|D|^{\frac12}$) one gets

$$|\Lambda_m(D)|^{k_0,\gamma}_{\frac12,s,\alpha} \lesssim_{M,\alpha} 1\,, \quad |\Delta_{12}\Lambda_m(D)|_{\frac12,s_1,\alpha} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+1)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m)}(M,\alpha)}\,. \tag{12.21}$$

For $m \geq 2$ there exist real, even, reversibility preserving, invertible maps $W^{(0)}_{m-1}$, $W^{(1)}_{m-1}$ of the form

$$\begin{aligned} W^{(0)}_{m-1} &:= \mathrm{Id} + w^{(0)}_{m-1}(\varphi,x,D) \qquad \text{with} \qquad w^{(0)}_{m-1}(\varphi,x,\xi) \in S^{-\frac{m-1}{2}}\,, \\ W^{(1)}_{m-1} &:= \mathrm{Id} + w^{(1)}_{m-1}(x,D) \qquad \text{with} \qquad w^{(1)}_{m-1}(x,\xi) \in S^{-\frac{m-1}{2}+\frac12} \end{aligned} \tag{12.22}$$

such that, for all $\omega \in \mathtt{DC}(\gamma,\tau)$,

$$L_7^{(m)} = (W^{(1)}_{m-1})^{-1}(W^{(0)}_{m-1})^{-1} L_7^{(m-1)} W^{(0)}_{m-1} W^{(1)}_{m-1}\,. \tag{12.23}$$

Initialization. For $m = 1$, the even and reversible operator $L_7^{(1)}$ in (12.14) has the form (12.16)-(12.17) with

$$r_1(D) = 0, \quad \Lambda_1(D) = \mathtt{im}_{\frac12} T_{\mathtt h}^{\frac12}|D|^{\frac12}\,. \tag{12.24}$$

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 (12.18) is trivial and (12.19)-(12.20) are (12.15).

Inductive step. In the next two subsections, we prove the above inductive claim, see (12.60)-(12.62) and Lemma 12.6. 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 (12.16), see (12.27). The resulting conjugated operator is $L_7^{(m,1)}$ in (12.34).
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 (12.34), see (12.48) and (12.52).

12.2.1 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 (12.16). We conjugate the operator $L_7^{(m)}$ in (12.16) by a transformation of the form (see (12.22))

$$W_m^{(0)} := \mathrm{Id} + w_m^{(0)}(\varphi,x,D)\,, \quad \text{with} \quad w_m^{(0)}(\varphi,x,\xi) \in S^{-\frac{m}{2}}\,, \tag{12.25}$$

which we shall fix in (12.29). We compute

$$\begin{aligned} L_7^{(m)} W_m^{(0)} = W_m^{(0)}\big(\omega\cdot\partial_\varphi + \Lambda_m(D)\big) + (\omega\cdot\partial_\varphi w_m^{(0)})(\varphi,x,D) + P_7^{(m)} \\ + \big[\Lambda_m(D), w_m^{(0)}(\varphi,x,D)\big] + P_7^{(m)} w_m^{(0)}(\varphi,x,D)\,. \end{aligned} \tag{12.26}$$

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 the commutator $[\Lambda_m(D), w_m^{(0)}(\varphi,x,D)]$ is in $OPS^{-\frac{m}{2}-\frac12}$ and $P_7^{(m)} w_m^{(0)}(\varphi,x,D)$ is in $OPS^{-m} \subseteq OPS^{-\frac{m}{2}-\frac12}$. Thus the term of order $-m/2$ in (12.26) 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

$$\omega\cdot\partial_\varphi w_m^{(0)}(\varphi,x,\xi) + p_7^{(m)}(\varphi,x,\xi) = \langle p_7^{(m)}\rangle_\varphi(x,\xi) \tag{12.27}$$

where

$$\langle p_7^{(m)}\rangle_\varphi(x,\xi) := \frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} p_7^{(m)}(\varphi,x,\xi)\,d\varphi\,. \tag{12.28}$$

For all $\omega \in \mathtt{DC}(\gamma,\tau)$, we choose the solution of (12.27) given by the periodic function

$$w_m^{(0)}(\varphi,x,\xi) := (\omega\cdot\partial_\varphi)^{-1}\Big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - p_7^{(m)}(\varphi,x,\xi)\Big)\,. \tag{12.29}$$

We extend the symbol $w_m^{(0)}$ in (12.29) 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 (12.25) is extended accordingly. We still denote by $w_m^{(0)}, W_m^{(0)}$ these extensions.

Lemma 12.3. *The operator $W_m^{(0)}$ defined in* (12.25), (12.29) *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* (12.18)-(12.21) *such that, if* (6.9) *holds with $\mu_0 \geq \aleph_7^{(m,1)}(M,\alpha)$, then for any $s \geq s_0$*

$$|\mathrm{Op}(w_m^{(0)})|^{k_0,\gamma}_{-\frac{m}{2},s,\alpha} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\aleph_7^{(m,1)}(M,\alpha)}\big) \tag{12.30}$$

$$|\Delta_{12}\mathrm{Op}(w_m^{(0)})|_{-\frac{m}{2},s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m,1)}(M,\alpha)}\,. \tag{12.31}$$

As a consequence, the transformation $W_m^{(0)}$ defined in (12.25), (12.29) *is invertible and*

$$|(W_m^{(0)})^{\pm 1} - \mathrm{Id}|^{k_0,\gamma}_{0,s,\alpha} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\aleph_7^{(m,1)}(M,\alpha)}\big) \tag{12.32}$$

$$|\Delta_{12}(W_m^{(0)})^{\pm 1}|_{0,s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m,1)}(M,\alpha)}\,. \tag{12.33}$$

Proof. We begin with proving (12.30). By (2.35)-(2.36) one has

$$|\mathrm{Op}(w_m^{(0)})|^{k_0,\gamma}_{-\frac{m}{2},s,\,\alpha} \lesssim_{k_0,\nu} \max_{\beta\in[0,\alpha]} \sup_{\xi\in\mathbb{R}} \langle\xi\rangle^{\frac{m}{2}+\beta}\big\|\partial_\xi^\beta w_m^{(0)}(\cdot,\cdot,\cdot,\xi)\big\|^{k_0,\gamma}_s.$$

By (12.29) and (2.17), for each $\xi \in \mathbb{R}$ one has

$$\|\partial_\xi^\beta w_m^{(0)}(\cdot,\cdot,\cdot,\xi)\|^{k_0,\gamma}_s \lesssim_{k_0,\nu} \gamma^{-1}\big\|\partial_\xi^\beta\big(\langle p_7^{(m)}\rangle_\varphi(\cdot,\xi) - p_7^{(m)}(\cdot,\cdot,\xi)\big)\big\|^{k_0,\gamma}_{s+\mu}$$

where μ is defined in (2.18) with $k+1 = k_0$. Hence $|\mathrm{Op}(w_m^{(0)})|^{k_0,\gamma}_{-\frac{m}{2},s,\,\alpha} \lesssim_{k_0,\nu} \gamma^{-1}|P_7^{(m)}|^{k_0,\gamma}_{-\frac{m}{2},s+\mu,\alpha}$ and (12.30) follows by (12.19). The other bounds are proved similarly, using the explicit formula (12.29), estimates (12.19)-(12.20) and (2.17), (2.44), and Lemma 2.13. □

By (12.26) and (12.27) we get the even and reversible operator

$$L_7^{(m,1)} := (W_m^{(0)})^{-1} L_7^{(m)} W_m^{(0)} = \omega\cdot\partial_\varphi + \Lambda_m(D) + \langle p_7^{(m)}\rangle_\varphi(x,D) + P_7^{(m,1)} \tag{12.34}$$

where

$$P_7^{(m,1)} := (W_m^{(0)})^{-1}\Big(\big[\Lambda_m(D), w_m^{(0)}(\varphi,x,D)\big] + P_7^{(m)} w_m^{(0)}(\varphi,x,D) - w_m^{(0)}(\varphi,x,D)\langle p_7^{(m)}\rangle_\varphi(x,D)\Big) \tag{12.35}$$

is in $OPS^{-\frac{m}{2}-\frac{1}{2}}$, as we prove in Lemma 12.4 below. Thus the term of order $-\frac{m}{2}$ in (12.34) is $\langle p_7^{(m)}\rangle_\varphi(x,D)$, which does not depend on φ any more.

Lemma 12.4. *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* (12.35) *is in $OPS^{-\frac{m}{2}-\frac{1}{2}}$. 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 12.3, such that, if* (6.9) *holds with $\mu_0 \geq \aleph_7^{(m,2)}(M,\alpha)$, then for any $s \geq s_0$*

$$|P_7^{(m,1)}|^{k_0,\gamma}_{-\frac{m}{2}-\frac{1}{2},s,\alpha} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)}(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\aleph_7^{(m,2)}(M,\alpha)})\,, \tag{12.36}$$

$$|\Delta_{12} P_7^{(m,1)}|_{-\frac{m}{2}-\frac{1}{2},s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)}\|\Delta_{12} i\|_{s_1+\aleph_7^{(m,2)}(M,\alpha)}\,. \tag{12.37}$$

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 (12.28) 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 (12.35) 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 (12.36)-(12.37) follow by the explicit expression in (12.35), Lemma 12.3, estimates (12.18)-(12.21), and (2.41), (2.44), (2.50). □

12.2.2 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 (12.34). To this aim we conjugate $L_7^{(m,1)}$ in (12.34) by a transformation of the form

$$W_m^{(1)} := \mathrm{Id} + w_m^{(1)}(x,D), \quad \text{where} \quad w_m^{(1)}(x,\xi)\in S^{-\frac{m}{2}+\frac12} \tag{12.38}$$

is a φ-independent symbol, which we shall fix in (12.50) (for $m=1$) and (12.54) (for $m\geq 2$). We denote the space average of the function $\langle p_7^{(m)}\rangle_\varphi(x,\xi)$ defined in (12.28) by

$$\langle p_7^{(m)}\rangle_{\varphi,x}(\xi) := \frac{1}{2\pi}\int_{\mathbb{T}}\langle p_7^{(m)}\rangle_\varphi(x,\xi)\,dx = \frac{1}{(2\pi)^{\nu+1}}\int_{\mathbb{T}^{\nu+1}} p_7^{(m)}(\varphi,x,\xi)\,d\varphi\,dx\,. \tag{12.39}$$

By (12.34), we compute

$$\begin{aligned} L_7^{(m,1)}W_m^{(1)} &= W_m^{(1)}\Big(\omega\cdot\partial_\varphi + \Lambda_m(D) + \langle p_7^{(m)}\rangle_{\varphi,x}\Big) + \big[\Lambda_m(D), w_m^{(1)}(x,D)\big] + \langle p_7^{(m)}\rangle_\varphi(x,D) - \langle p_7^{(m)}\rangle_{\varphi,x}(D) \\ &\quad + \langle p_7^{(m)}\rangle_\varphi(x,D)w_m^{(1)}(x,D) - w_m^{(1)}(x,D)\langle p_7^{(m)}\rangle_{\varphi,x}(D) + P_7^{(m,1)}W_m^{(1)}\,. \end{aligned} \tag{12.40}$$

By formulas (2.42), (2.43) (with $N=1$) and (2.47), (2.48),

$$\langle p_7^{(m)}\rangle_\varphi(x,D)w_m^{(1)}(x,D) = \mathrm{Op}\Big(\langle p_7^{(m)}\rangle_\varphi(x,\xi)w_m^{(1)}(x,\xi)\Big) + r_{\langle p_7^{(m)}\rangle_\varphi, w_m^{(1)}}(x,D)\,, \tag{12.41}$$

$$w_m^{(1)}(x,D)\langle p_7^{(m)}\rangle_{\varphi,x}(D) = \mathrm{Op}\Big(w_m^{(1)}(x,\xi)\langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\Big) + r_{w_m^{(1)},\langle p_7^{(m)}\rangle_{\varphi,x}}(x,D)\,, \tag{12.42}$$

$$\big[\Lambda_m(D), w_m^{(1)}(x,D)\big] = \mathrm{Op}\Big(-\mathrm{i}\partial_\xi\Lambda_m(\xi)\partial_x w_m^{(1)}(x,\xi)\Big) + \mathtt{r}_2(\Lambda_m, w_m^{(1)})(x,D) \tag{12.43}$$

where $r_{\langle p_7^{(m)}\rangle_\varphi, w_m^{(1)}}$, $r_{w_m^{(1)},\langle p_7^{(m)}\rangle_{\varphi,x}}\in S^{-m-\frac12}\subset S^{-\frac{m}{2}-\frac12}$, $\mathtt{r}_2(\Lambda_m,w_m^{(1)})(x,D)\in S^{-\frac{m}{2}-1}\subset S^{-\frac{m}{2}-\frac12}$. Let $\chi_0\in\mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ be a cut-off function satisfying

$$\chi_0(\xi)=\chi_0(-\xi)\ \ \forall\xi\in\mathbb{R}\,,\quad \chi_0(\xi)=0\ \ \forall|\xi|\leq\frac45\,,\quad \chi_0(\xi)=1\ \ \forall|\xi|\geq\frac78\,. \tag{12.44}$$

By (12.40)-(12.43), one has

$$\begin{aligned} L_7^{(m,1)}W_m^{(1)} &= W_m^{(1)}\big(\omega\cdot\partial_\varphi + \Lambda_m(D) + \langle p_7^{(m)}\rangle_{\varphi,x}(D)\big) \\ &\quad + \mathrm{Op}\Big(-\mathrm{i}\partial_\xi\Lambda_m(\xi)\partial_x w_m^{(1)}(x,\xi) + \chi_0(\xi)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big) \qquad (12.45) \\ &\qquad + \chi_0(\xi)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)w_m^{(1)}(x,\xi)\Big) \qquad (12.46) \\ &\quad + \mathrm{Op}\Big(\big(1-\chi_0(\xi)\big)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)\big(1+w_m^{(1)}(x,\xi)\big)\Big) \\ &\quad + \mathtt{r}_2(\Lambda_m,w_m^{(1)})(x,D) + r_{\langle p_7^{(m)}\rangle_\varphi, w_m^{(1)}}(x,D) - r_{w_m^{(1)},\langle p_7^{(m)}\rangle_{\varphi,x}}(x,D) + P_7^{(m,1)}W_m^{(1)}\,. \qquad (12.47) \end{aligned}$$

The terms containing $1-\chi_0(\xi)$ are in $S^{-\infty}$, by definition (12.44). The term in (12.45) is of order $-\frac{m}{2}$ and the term in (12.46) is of order $-m+\frac{1}{2}$, which equals $-\frac{m}{2}$ for $m=1$, and is strictly less than $-\frac{m}{2}$ for $m\geq 2$. Hence we split the two cases $m=1$ and $m\geq 2$.

First case: $m=1$. We look for $w_m^{(1)}(x,\xi)=w_1^{(1)}(x,\xi)$ such that

$$-\,\mathrm{i}\partial_\xi\Lambda_1(\xi)\partial_x w_1^{(1)}(x,\xi)+\chi_0(\xi)\Big(\langle p_7^{(1)}\rangle_\varphi(x,\xi)-\langle p_7^{(1)}\rangle_{\varphi,x}(\xi)\Big)(1+w_1^{(1)}(x,\xi))=0\,. \tag{12.48}$$

By (12.24) and recalling (2.31), (2.16), for $|\xi|\geq 4/5$ one has $\Lambda_1(\xi)=\mathtt{im}_{\frac{1}{2}}\tanh^{\frac{1}{2}}(\mathtt{h}|\xi|)|\xi|^{\frac{1}{2}}$. Since, by (11.24), $|\mathtt{m}_{\frac{1}{2}}|\geq 1/2$ for $\varepsilon\gamma^{-1}$ small enough, we have

$$\inf_{|\xi|\geq\frac{4}{5}}|\xi|^{\frac{1}{2}}|\partial_\xi\Lambda_1(\xi)|\geq\delta>0\,, \tag{12.49}$$

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 (12.48) given by the periodic function

$$w_1^{(1)}(x,\xi):=\exp\big(g_1(x,\xi)\big)-1,\quad g_1(x,\xi):=\begin{cases}\dfrac{\chi_0(\xi)\partial_x^{-1}\big(\langle p_7^{(1)}\rangle_\varphi(x,\xi)-\langle p_7^{(1)}\rangle_{\varphi,x}(\xi)\big)}{\mathrm{i}\partial_\xi\Lambda_1(\xi)} & \text{if } |\xi|\geq\frac{4}{5}\\ 0 & \text{if } |\xi|\leq\frac{4}{5}\,.\end{cases} \tag{12.50}$$

Note that, by the definition of the cut-off function χ_0 given in (12.44), recalling (12.24), (12.49) and applying estimates (2.40), (11.24), the Fourier multiplier $\frac{\chi_0(\xi)}{\partial_\xi\Lambda_1(\xi)}$ is a symbol in $S^{\frac{1}{2}}$ and satisfies

$$\Big|\mathrm{Op}\Big(\frac{\chi_0(\xi)}{\partial_\xi\Lambda_1(\xi)}\Big)\Big|_{\frac{1}{2},s,\alpha}^{k_0,\gamma}\lesssim_\alpha 1\,,\quad\Big|\Delta_{12}\mathrm{Op}\Big(\frac{\chi_0(\xi)}{\partial_\xi\Lambda_1(\xi)}\Big)\Big|_{\frac{1}{2},s_1,\alpha}\lesssim_\alpha\varepsilon\gamma^{-1}\|\Delta_{12}i\|_{s_1}\,. \tag{12.51}$$

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

$$-\,\mathrm{i}\partial_\xi\Lambda_m(\xi)\partial_x w_m^{(1)}(x,\xi)+\chi_0(\xi)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi)-\langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)=0\,. \tag{12.52}$$

Recalling (12.17)-(12.18) and (12.49), one has that

$$\begin{aligned}\inf_{|\xi|\geq\frac{4}{5}}|\xi|^{\frac{1}{2}}|\partial_\xi\Lambda_m(\xi)|&\geq\inf_{|\xi|\geq\frac{4}{5}}|\xi|^{\frac{1}{2}}|\partial_\xi\Lambda_1(\xi)|-\sup_{\xi\in\mathbb{R}}|\xi|^{\frac{1}{2}}|\partial_\xi r_m(\xi)|\geq\delta-|r_m(D)|_{-\frac{1}{2},0,1}\\&\geq\delta-C\varepsilon\gamma^{-(m+1)}\geq\delta/2\end{aligned} \tag{12.53}$$

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 (12.52) given by the periodic function

$$w_m^{(1)}(x,\xi):=\begin{cases}\dfrac{\chi_0(\xi)\partial_x^{-1}\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi)-\langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)}{\mathrm{i}\partial_\xi\Lambda_m(\xi)} & \text{if } |\xi|\geq\frac{4}{5}\\ 0 & \text{if } |\xi|\leq\frac{4}{5}\,.\end{cases} \tag{12.54}$$

By the definition of the cut-off function χ_0 in (12.44), recalling (12.24), (12.17), (12.53), and applying estimates (2.40), (11.24), (12.18), the Fourier multiplier $\frac{\chi_0(\xi)}{\partial_\xi\Lambda_m(\xi)}$ is a symbol in $S^{\frac{1}{2}}$ and satisfies

$$\Big|\mathrm{Op}\Big(\frac{\chi_0(\xi)}{\partial_\xi\Lambda_m(\xi)}\Big)\Big|_{\frac{1}{2},s,\alpha}^{k_0,\gamma}\lesssim_{M,\alpha}1\,,\quad\Big|\Delta_{12}\mathrm{Op}\Big(\frac{\chi_0(\xi)}{\partial_\xi\Lambda_m(\xi)}\Big)\Big|_{\frac{1}{2},s_1,\alpha}\lesssim_{M,\alpha}\varepsilon\gamma^{-(m+1)}\|\Delta_{12}i\|_{s_1+\aleph_7^{(m)}(M,\alpha)}\,. \tag{12.55}$$

By (12.53), the function $w_m^{(1)}(x,\xi)$ is a well-defined symbol in $S^{-\frac{m}{2}+\frac{1}{2}}$.

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 (12.47).

Lemma 12.5. *The operators $W_m^{(1)}$ defined in* (12.38), (12.50) *for $m = 1$, and* (12.54) *for $m \geq 2$, 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 12.4, such that, if* (6.9) *holds with $\mu_0 \geq \aleph_7^{(m,3)}(M, \alpha)$, then for any $s \geq s_0$*

$$|\mathrm{Op}(w_m^{(1)})|_{-\frac{m}{2}+\frac{1}{2},s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(m,3)}(M,\alpha)}^{k_0,\gamma}\big) \tag{12.56}$$

$$|\Delta_{12}\mathrm{Op}(w_m^{(1)})|_{-\frac{m}{2}+\frac{1}{2},s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)}\|\Delta_{12}i\|_{s_1+\aleph_7^{(m,3)}(M,\alpha)}\,. \tag{12.57}$$

As a consequence, the transformation $W_m^{(1)}$ is invertible and

$$|(W_m^{(1)})^{\pm 1} - \mathrm{Id}|_{0,s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(m,3)}(M,\alpha)}^{k_0,\gamma}\big) \tag{12.58}$$

$$|\Delta_{12}(W_m^{(1)})^{\pm 1}|_{0,s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)}\|\Delta_{12}i\|_{s_1+\aleph_7^{(m,3)}(M,\alpha)}\,. \tag{12.59}$$

Proof. The lemma follows by the explicit expressions in (12.38), (12.50), (12.54), (12.39), by estimates (2.40), (2.41), (2.46), Lemmata 2.10, 2.11, 2.13 and estimates (12.19), (12.20), (12.51), (12.55). □

In conclusion, by (12.47), (12.48) and (12.52), we obtain the even and reversible operator

$$L_7^{(m+1)} := (W_m^{(1)})^{-1} L_7^{(m,1)} W_m^{(1)} = \omega\cdot\partial_\varphi + \Lambda_{m+1}(D) + P_7^{(m+1)} \tag{12.60}$$

where

$$\begin{aligned}\Lambda_{m+1}(D) &:= \Lambda_m(D) + \langle p_7^{(m)}\rangle_{\varphi,x}(D) = \mathtt{im}_{\frac12} T_{\mathtt{h}}^{\frac12}|D|^{\frac12} + r_{m+1}(D)\,,\\ r_{m+1}(D) &:= r_m(D) + \langle p_7^{(m)}\rangle_{\varphi,x}(D)\,,\end{aligned} \tag{12.61}$$

and

$$\begin{aligned}P_7^{(m+1)} := (W_m^{(1)})^{-1}\Big\{ & \mathtt{r}_2(\Lambda_m, w_m^{(1)})(x,D) + r_{\langle p_7^{(m)}\rangle_\varphi, w_m^{(1)}}(x,D) - r_{w_m^{(1)}, \langle p_7^{(m)}\rangle_{\varphi,x}}(x,D) + P_7^{(m,1)} W_m^{(1)} \\ & + \chi_{(m\geq 2)}\mathrm{Op}\,\Big(\chi_0(\xi)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)w_m^{(1)}(x,\xi)\Big) \\ & + \mathrm{Op}\Big(\big(1-\chi_0(\xi)\big)\big(\langle p_7^{(m)}\rangle_\varphi(x,\xi) - \langle p_7^{(m)}\rangle_{\varphi,x}(\xi)\big)\big(1 + w_m^{(1)}(x,\xi)\big)\Big)\Big\}\end{aligned} \tag{12.62}$$

with $\chi_{(m\geq 2)}$ equal to 1 if $m \geq 2$, and zero otherwise.

Lemma 12.6. *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 12.5, such that, if* (6.9) *holds with $\mu_0 \geq \aleph_7^{(m+1)}(M, \alpha)$, then for any $s \geq s_0$*

$$|r_{m+1}(D)|_{-\frac12,s,\alpha}^{k_0,\gamma} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+2)}\,, \quad |\Delta_{12} r_{m+1}(D)|_{-\frac12,s_1,\alpha} \lesssim_{M,\alpha} \varepsilon\gamma^{-(m+2)}\|\Delta_{12}i\|_{s_1+\aleph_7^{(m+1)}(M,\alpha)} \tag{12.63}$$

$$|P_7^{(m+1)}|_{-\frac{m}{2}-\frac12,s,\alpha}^{k_0,\gamma} \lesssim_{M,s,\alpha} \varepsilon\gamma^{-(m+3)}\big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(m+1)}(M,\alpha)}^{k_0,\gamma}\big)\,, \tag{12.64}$$

$$|\Delta_{12}P_7^{(m+1)}|_{-\frac{m}{2}-\frac12,s_1,\alpha} \lesssim_{M,s_1,\alpha} \varepsilon\gamma^{-(m+3)}\|\Delta_{12}i\|_{s_1+\aleph_7^{(m+1)}(M,\alpha)}\,. \tag{12.65}$$

Proof. Since the operator $\langle p_7^{(m)}\rangle_\varphi(x,D)$ is even and reversible by Lemma 12.4, the average $\langle p_7^{(m)}\rangle_{\varphi,x}(D)$ defined in (12.39) is even and reversible as well (we use Remark 2.22). 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 (12.61) are even and reversible.

Estimates (12.63)-(12.65) for $r_{m+1}(D)$ and $P_7^{(m+1)}$ defined respectively in (12.61) and (12.62) follow by the explicit expressions of $\langle p_7^{(m)}\rangle_{\varphi,x}(\xi)$ in (12.39) and $w_m^{(1)}$ in (12.50) and (12.54) (for $m = 1$ and $m \geq 2$ respectively), by applying (2.41), (2.40), (12.58)-(12.59), (12.36)-(12.37), (2.46), Lemmata 2.10, 2.11, and the inductive estimates (12.18)-(12.21). □

Thus, the proof of the inductive claims (12.18)-(12.23) is complete.

12.2.3 Conclusion of the reduction of $L_7^{(1)}$

Composing all the previous transformations, we obtain the even and reversibility preserving map

$$W := W_0 \circ W_1^{(0)} \circ W_1^{(1)} \circ \ldots \circ W_{2M-1}^{(0)} \circ W_{2M-1}^{(1)} , \tag{12.66}$$

where W_0 is defined in (12.3) and for $m = 1, \ldots, 2M-1$, $W_m^{(0)}, W_m^{(1)}$ are defined in (12.25), (12.38). The order M will be fixed in (14.8). By (12.16), (12.17), (12.23) at $m = 2M$, the operator L_7 in (12.2) is conjugated, for all $\omega \in \mathtt{DC}(\gamma, \tau)$, to the even and reversible operator

$$L_8 := L_7^{(2M)} = W^{-1} L_7 W = \omega \cdot \partial_\varphi + \Lambda_{2M}(D) + P_{2M} \tag{12.67}$$

where $P_{2M} := P_7^{(2M)} \in OPS^{-M}$ and

$$\Lambda_{2M}(D) = \mathrm{i}\mathtt{m}_{\frac12} T_{\mathtt{h}}^{\frac12} |D|^{\frac12} + r_{2M}(D) , \quad r_{2M}(D) \in OPS^{-\frac12} . \tag{12.68}$$

Lemma 12.7. *Assume* (6.9) *with* $\mu_0 \geq \aleph_7^{(2M)}(M, 0)$. *Then, for any* $s \geq s_0$, *the following estimates hold:*

$$|r_{2M}(D)|_{-\frac12, s, 0}^{k_0,\gamma} \lesssim_M \varepsilon \gamma^{-(2M+1)} , \quad |\Delta_{12} r_{2M}(D)|_{-\frac12, s_1, 0} \lesssim_M \varepsilon \gamma^{-(2M+1)} \|\Delta_{12} i\|_{s_1 + \aleph_7^{(2M)}(M,0)} , \tag{12.69}$$

$$|P_{2M}|_{-M,s,0}^{k_0,\gamma} \lesssim_{M,s} \varepsilon \gamma^{-2(M+1)} \big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(2M)}(M,0)}^{k_0,\gamma}\big) , \tag{12.70}$$

$$|\Delta_{12} P_{2M}|_{-M,s_1,0} \lesssim_{M,s_1} \varepsilon \gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(2M)}(M,0)} , \tag{12.71}$$

$$|W^{\pm 1} - \mathrm{Id}|_{0,s,0}^{k_0,\gamma} \lesssim_{M,s} \varepsilon \gamma^{-2(M+1)} \big(1 + \|\mathfrak{I}_0\|_{s+\aleph_7^{(2M)}(M,0)}^{k_0,\gamma}\big) , \tag{12.72}$$

$$|\Delta_{12} W^{\pm 1}|_{0,s_1,0} \lesssim_{M,s_1} \varepsilon \gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph_7^{(2M)}(M,0)} . \tag{12.73}$$

Proof. Estimates (12.69), (12.70), (12.71) follow by (12.18), (12.19), (12.20) applied for $m = 2M$. Estimates (12.72)-(12.73) for the map W defined in (12.66), and its inverse W^{-1}, follow by (12.13), (12.32), (12.33), (12.58), (12.59), applying the composition estimate (2.44) (with $m = m' = \alpha = 0$). □

Since $\Lambda_{2M}(D)$ is even and reversible, we have that

$$\Lambda_{2M}(\xi), r_{2M}(\xi) \in \mathrm{i}\mathbb{R} \qquad \text{and} \qquad \Lambda_{2M}(\xi) = \Lambda_{2M}(-\xi) , \ r_{2M}(\xi) = r_{2M}(-\xi) . \tag{12.74}$$

In conclusion, we write the even and reversible operator L_8 in (12.67) as

$$L_8 = \omega \cdot \partial_\varphi + \mathrm{i} D_8 + P_{2M} \tag{12.75}$$

where D_8 is the diagonal operator

$$D_8 := -\mathrm{i}\Lambda_{2M}(D) := \mathrm{diag}_{j \in \mathbb{Z}}(\mu_j) , \qquad \mu_j := \mathtt{m}_{\frac12} |j|^{\frac12} \tanh(\mathtt{h}|j|)^{\frac12} + r_j , \quad r_j := -\mathrm{i}\, r_{2M}(j) , \tag{12.76}$$

$$\mu_j, r_j \in \mathbb{R} , \qquad \mu_j = \mu_{-j} , \ r_j = r_{-j} , \quad \forall j \in \mathbb{Z} , \tag{12.77}$$

with $r_j \in \mathbb{R}$ satisfying, by (12.69),

$$\sup_{j \in \mathbb{Z}} |j|^{\frac12} |r_j|^{k_0,\gamma} \lesssim_M \varepsilon \gamma^{-(2M+1)} , \quad \sup_{j \in \mathbb{Z}} |j|^{\frac12} |\Delta_{12} r_j| \lesssim_M \varepsilon \gamma^{-(2M+1)} \|\Delta_{12} i\|_{s_1 + \aleph_7^{(2M)}(M,0)} \tag{12.78}$$

and $P_{2M} \in OPS^{-M}$ satisfies (12.70)-(12.71).

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 2.24 and 2.29).

Remark 12.8. In view of Lemma 2.28, the tame constants of P_{2M} can be deduced by estimates (12.70)-(12.71) 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 (12.54)), and α keeps track of the regularity of symbols with respect to ξ. □

12.3 Conjugation of $\mathcal{L}_7$

In the previous subsections 12.1-12.2 we have conjugated the operator L_7 defined in (12.2) to L_8 in (12.67), whose symbol is constant in (φ, x), up to smoothing remainders of order $-M$. Now we conjugate the whole operator $\mathcal{L}_7$ in (12.1) by the real, even and reversibility preserving map

$$\mathcal{W} := \begin{pmatrix} W & 0 \\ 0 & \overline{W} \end{pmatrix} \tag{12.79}$$

where W is defined in (12.66). By (12.67), (12.75) we obtain, for all $\omega \in \mathtt{DC}(\gamma, \tau)$, the real, even and reversible operator

$$\mathcal{L}_8 := \mathcal{W}^{-1} \mathcal{L}_7 \mathcal{W} = \omega \cdot \partial_\varphi + \mathrm{i}\mathcal{D}_8 + \mathrm{i}\Pi_0 + \mathcal{T}_8 \tag{12.80}$$

where $\mathcal{D}_8$ is the diagonal operator

$$\mathcal{D}_8 := \begin{pmatrix} D_8 & 0 \\ 0 & -D_8 \end{pmatrix}, \tag{12.81}$$

with D_8 defined in (12.76), and the remainder $\mathcal{T}_8$ is

$$\mathcal{T}_8 := \mathrm{i}\mathcal{W}^{-1}\Pi_0 \mathcal{W} - \mathrm{i}\Pi_0 + \mathcal{W}^{-1}\mathcal{T}_7 \mathcal{W} + \mathcal{P}_{2M}, \quad \mathcal{P}_{2M} := \begin{pmatrix} P_{2M} & 0 \\ 0 & \overline{P_{2M}} \end{pmatrix} \tag{12.82}$$

with P_{2M} defined in (12.67). 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 (12.80) 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 12.9. *For any $M > 0$, there exists a constant $\aleph_8(M) > 0$ (depending also on τ, ν, k_0) such that, if* (6.9) *holds with $\mu_0 \geq \aleph_8(M)$, then for any $s \geq s_0$*

$$|\mathcal{W}^{\pm 1} - \mathrm{Id}|_{0,s,0}^{k_0,\gamma}, \ |\mathcal{W}^{*} - \mathrm{Id}|_{0,s,0}^{k_0,\gamma} \lesssim_{M,s} \varepsilon \gamma^{-2(M+1)} \big(1 + \|\mathfrak{I}_0\|_{s+\aleph_8(M)}^{k_0,\gamma}\big), \tag{12.83}$$

$$|\Delta_{12}\mathcal{W}^{\pm 1}|_{0,s_1,0}, \ |\Delta_{12}\mathcal{W}^{*}|_{0,s_1,0} \lesssim_{M,s_1} \varepsilon \gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph_8(M)}. \tag{12.84}$$

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 (6.9) *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*

$$\mathfrak{M}_{\langle D \rangle^{m_1} (\partial_\varphi^\beta \mathcal{T}_8) \langle D \rangle^{m_2}}(s) \lesssim_{M,S} \varepsilon \gamma^{-2(M+1)} \big(1 + \|\mathfrak{I}_0\|_{s+\aleph_8'(M,\beta_0)}\big), \qquad \forall s_0 \leq s \leq S \tag{12.85}$$

$$\|\langle D \rangle^{m_1} \Delta_{12}(\partial_\varphi^\beta \mathcal{T}_8) \langle D \rangle^{m_2}\|_{\mathcal{L}(H^{s_1})} \lesssim_{M,S} \varepsilon \gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph_8'(M,\beta_0)}. \tag{12.86}$$

Proof. Estimates (12.83), (12.84) follow by definition (12.79), by estimates (12.72), (12.73) and using also Lemma 2.12 to estimate the adjoint operator. Let us prove (12.85) (the proof of (12.86) 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 11.4, 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

$$\begin{aligned} &\langle D \rangle^{m_1} (\partial_\varphi^{\beta_1}\mathcal{W}^{-1})(\partial_\varphi^{\beta_2}\mathcal{T}_7)(\partial_\varphi^{\beta_3}\mathcal{W}) \langle D \rangle^{m_2} \\ &= \Big(\langle D \rangle^{m_1} \partial_\varphi^{\beta_1}\mathcal{W} \langle D \rangle^{-m_1}\Big) \circ \Big(\langle D \rangle^{m_1} \partial_\varphi^{\beta_2}\mathcal{T}_7 \langle D \rangle^{m_2}\Big) \circ \Big(\langle D \rangle^{-m_2} \partial_\varphi^{\beta_3}\mathcal{W} \langle D \rangle^{m_2}\Big). \end{aligned} \tag{12.87}$$

For any $m \geq 0$, $\beta \in \mathbb{N}^\nu$, $|\beta| \leq \beta_0$, by (2.68), (2.40), (2.46), (2.44), 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 (12.83). The estimate of (12.87) then follows by (11.39) and Lemma 2.26. The tame estimate of $\langle D \rangle^{m_1} \partial_\varphi^{\beta}\mathcal{P}_{2M} \langle D \rangle^{m_2}$ follows by (2.68), (12.70), (12.71). 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 2.36 (applied with $A = \mathcal{W}^{-1}$ and $B = \mathcal{W}$) and (2.68), (12.83), (12.84). $\square$

13 Conclusion: reduction of $\mathcal{L}_\omega$ up to smoothing operators

By Sections 6-12, 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 (6.8) is conjugated to the real, even and reversible operator $\mathcal{L}_8$ defined in (12.80), namely

$$\mathcal{P}^{-1}\mathcal{L}\mathcal{P} = \mathcal{L}_8 = \omega \cdot \partial_\varphi + \mathrm{i}\mathcal{D}_8 + \mathrm{i}\Pi_0 + \mathcal{T}_8\,, \tag{13.1}$$

where $\mathcal{P}$ is the real, even and reversibility preserving map

$$\mathcal{P} := \mathcal{Z}\mathcal{B}\mathcal{A}\mathcal{M}_2\mathcal{M}_3\mathcal{C}\mathbf{\Phi}_M\mathbf{\Phi}\mathcal{W}\,. \tag{13.2}$$

Moreover, as already noticed below (12.82), 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 (5.26), which acts on the normal subspace $H^\perp_{\mathbb{S}^+}$, whose relation with $\mathcal{L}$ is stated in (6.5). The operator $\mathcal{L}_\omega$ is even and reversible as stated in Lemma 6.1.

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 13.1. (Restriction of the conjugation map to $H^\perp_{\mathbb{S}_0}$) *Let $M > 0$. There exists a constant $\sigma_M > 0$ (depending also on k_0, τ, ν) such that, assuming (6.9) 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*

$$\mathcal{P}_\perp := \Pi^\perp_{\mathbb{S}_0}\mathcal{P}\Pi^\perp_{\mathbb{S}_0} \tag{13.3}$$

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

$$\|\mathcal{P}^{\pm 1}_\perp h\|^{k_0,\gamma}_s \lesssim_{M,s} \|h\|^{k_0,\gamma}_{s+\sigma_M} + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\sigma_M}\|h\|^{k_0,\gamma}_{s_0+\sigma_M}\,, \tag{13.4}$$

$$\|(\Delta_{12}\mathcal{P}^{\pm 1}_\perp)h\|_{s_1} \lesssim_{M,s_1} \varepsilon\gamma^{-2(M+1)}\|\Delta_{12} i\|_{s_1+\sigma_M}\|h\|_{s_1+1}\,. \tag{13.5}$$

The operator $\mathcal{P}_\perp$ is real, even and reversibility preserving. The operators $\mathcal{P}, \mathcal{P}^{-1}$ also satisfy (13.4), (13.5).

Proof. Applying (2.69) and (6.17), (7.28), (8.10), (9.7), (9.15), (2.60), (10.36), (11.35), (12.83) we get

$$\|Ah\|^{k_0,\gamma}_s \lesssim_s \|h\|^{k_0,\gamma}_{s+\mu_M} + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu_M}\|h\|^{k_0,\gamma}_{s_0+\mu_M}\,, \quad A \in \{\mathcal{Z}^{\pm 1}, \mathcal{B}^{\pm 1}, \mathcal{A}^{\pm 1}, \mathcal{M}_2^{\pm 1}, \mathcal{M}_3^{\pm 1}, \mathcal{C}^{\pm 1}, \mathbf{\Phi}_M^{\pm 1}, \mathbf{\Phi}^{\pm 1}, \mathcal{W}^{\pm 1}\}\,,$$

for some $\mu_M > 0$. Then by the definition (13.2) 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 (13.4). 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 (13.5) follows similarly by using (6.20), (7.30), (8.10), (9.19), (10.37), (11.36), (12.84). □

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 (5.26) is conjugated to

$$\mathcal{L}_\perp := \mathcal{P}^{-1}_\perp\mathcal{L}_\omega\mathcal{P}_\perp = \Pi^\perp_{\mathbb{S}_0}\mathcal{L}_8\Pi^\perp_{\mathbb{S}_0} + R_M \tag{13.6}$$

where

$$R_M := \mathcal{P}^{-1}_\perp\Pi^\perp_{\mathbb{S}_0}\big(\mathcal{P}\Pi_{\mathbb{S}_0}\mathcal{L}_8\Pi^\perp_{\mathbb{S}_0} - \mathcal{L}\Pi_{\mathbb{S}_0}\mathcal{P}\Pi^\perp_{\mathbb{S}_0} + \varepsilon R\mathcal{P}_\perp\big) \tag{13.7}$$

$$= \mathcal{P}^{-1}_\perp\Pi^\perp_{\mathbb{S}_0}\mathcal{P}\Pi_{\mathbb{S}_0}\mathcal{T}_8\Pi^\perp_{\mathbb{S}_0} + \mathcal{P}^{-1}_\perp\Pi^\perp_{\mathbb{S}_0} J\partial_u\nabla_u H(T_\delta(\varphi))\Pi_{\mathbb{S}_0}\mathcal{P}\Pi^\perp_{\mathbb{S}_0} + \varepsilon\mathcal{P}^{-1}_\perp\Pi^\perp_{\mathbb{S}_0} R\mathcal{P}_\perp \tag{13.8}$$

is a finite dimensional operator. To prove (13.6)-(13.7) we first use (6.5) and (13.3) 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 (13.1) 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 (13.8), we use (13.1), (6.5), 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 (12.81) and (12.76).

Lemma 13.2. *The operator R_M in* (13.7) *has the finite dimensional form* (6.3). *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* (6.9) *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 $\langle D\rangle^{m_1} \partial_\varphi^\beta R_M \langle D\rangle^{m_2}$, $\langle D\rangle^{m_1} \partial_\varphi^\beta \Delta_{12} R_M \langle D\rangle^{m_2}$ are $\mathcal{D}^{k_0}$-tame with tame constants*

$$\mathfrak{M}_{\langle D\rangle^{m_1} \partial_\varphi^\beta R_M \langle D\rangle^{m_2}}(s) \lesssim_{M,S} \varepsilon \gamma^{-2(M+1)} \big(1 + \|\mathfrak{I}_0\|_{s+\aleph_9(M,\beta_0)}^{k_0,\gamma}\big)\,, \qquad \forall s_0 \leq s \leq S \tag{13.9}$$

$$\|\langle D\rangle^{m_1} \Delta_{12} \partial_\varphi^\beta R_M \langle D\rangle^{m_2}\|_{\mathcal{L}(H^{s_1})} \lesssim_{M,S} \varepsilon \gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph_9(M,\beta_0)}\,. \tag{13.10}$$

Proof. To prove that the operator R_M has the finite dimensional form (6.3), notice that in the first two terms in (13.8) there is the finite dimensional projector $\Pi_{\mathbb{S}_0}$, that the operator R in the third term in (13.8) already has the finite dimensional form (6.3), 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 2.36 (and Lemma 6.30 in [21] and Lemma 8.3 in [8]). To estimate R_M, use (13.4), (13.5) for $\mathcal{P}$, (12.85), (12.86) for $\mathcal{T}_8$, (6.5), (6.8), (6.18), (6.19), (A.3) for $J\partial_u \nabla_u H(T_\delta(\varphi))$, (6.3), (6.4) 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 $\Pi_{\mathbb{S}_0}^\perp \left(\begin{smallmatrix} 0 & -D\tanh(\mathtt{h}D) \\ 1 & 0 \end{smallmatrix}\right) \Pi_{\mathbb{S}_0}$ is zero. $\square$

By (13.6) and (12.80) we get

$$\mathcal{L}_\perp = \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_\perp + \mathcal{R}_\perp \tag{13.11}$$

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),

$$\mathcal{D}_\perp := \begin{pmatrix} D_\perp & 0 \\ 0 & -D_\perp \end{pmatrix}, \qquad D_\perp := \Pi_{\mathbb{S}_0}^\perp D_8 \Pi_{\mathbb{S}_0}^\perp\,, \tag{13.12}$$

and $\mathcal{R}_\perp$ is the operator

$$\mathcal{R}_\perp := \Pi_{\mathbb{S}_0}^\perp \mathcal{T}_8 \Pi_{\mathbb{S}_0}^\perp + R_M\,, \quad \mathcal{R}_\perp = \begin{pmatrix} \mathcal{R}_{\perp,1} & \mathcal{R}_{\perp,2} \\ \overline{\mathcal{R}_{\perp,2}} & \overline{\mathcal{R}_{\perp,1}} \end{pmatrix}. \tag{13.13}$$

The operator $\mathcal{R}_\perp$ in (13.13) is defined for all $\lambda = (\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, because $\mathcal{T}_8$ in (12.82) and the operator in the right hand side of (13.8) are defined on the whole parameter space. As a consequence, the right hand side of (13.11) 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 13.3. (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* (6.5) *is conjugated by the map $\mathcal{P}_\perp$ defined in* (13.3) *to the real, even and reversible operator $\mathcal{L}_\perp$ in* (13.6). *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* (13.11) *has the form*

$$\mathcal{L}_\perp = \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_\perp + \mathcal{R}_\perp \tag{13.14}$$

where $\mathcal{D}_\perp$ is the diagonal operator

$$\mathcal{D}_\perp := \begin{pmatrix} D_\perp & 0 \\ 0 & -D_\perp \end{pmatrix}, \qquad D_\perp = \mathrm{diag}_{j \in \mathbb{S}_0^c}\, \mu_j\,, \qquad \mu_{-j} = \mu_j\,, \tag{13.15}$$

with eigenvalues μ_j, defined in (12.76), *given by*

$$\mu_j = \mathtt{m}_{\frac{1}{2}} |j|^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h}|j|) + r_j \in \mathbb{R}\,, \qquad r_{-j} = r_j\,, \tag{13.16}$$

where $\mathtt{m}_{\frac{1}{2}}, r_j \in \mathbb{R}$ satisfy (11.24), (12.78). *The operator $\mathcal{R}_\perp$ defined in* (13.13) *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 (6.9) *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*

$$\mathfrak{M}_{\langle D\rangle^{m_1} \partial_\varphi^\beta \mathcal{R}_\perp \langle D\rangle^{m_2}}(s) \lesssim_{M,S} \varepsilon \gamma^{-2(M+1)} \big(1 + \|\mathfrak{I}_0\|_{s+\aleph(M,\beta_0)}^{k_0,\gamma}\big)\,, \qquad \forall s_0 \leq s \leq S \tag{13.17}$$

$$\|\langle D\rangle^{m_1} \Delta_{12} \partial_\varphi^\beta \mathcal{R}_\perp \langle D\rangle^{m_2}\|_{\mathcal{L}(H^{s_1})} \lesssim_{M,S} \varepsilon \gamma^{-2(M+1)} \|\Delta_{12} i\|_{s_1+\aleph(M,\beta_0)}\,. \tag{13.18}$$

Proof. Estimates (13.17)-(13.18) for the term $\Pi_{\mathbb{S}_0}^\perp \mathcal{T}_8 \Pi_{\mathbb{S}_0}^\perp$ in (13.13) follow directly by (12.85), (12.86). Estimates (13.17)-(13.18) for R_M are (13.9)-(13.10). $\square$

14 Almost-diagonalization and invertibility of $\mathcal{L}_\omega$

In Proposition 13.3 we obtained the operator $\mathcal{L}_\perp = \mathcal{L}_\perp(\varphi)$ in (13.14) 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 (13.14) with the operator $\mathcal{L}_\perp^{sym}$ defined in (14.1) below, which coincides with $\mathcal{L}_\perp$ on the subspace of functions even in x, see Lemma 14.1. This trick enables to reduce an even operator using its matrix representation in the exponential basis $(e^{\mathrm{i}jx})_{j\in\mathbb{Z}}$ and exploiting the fact that on the subspace of functions even(x) its eigenvalues are simple. We define the linear operator $\mathcal{L}_\perp^{sym}$, acting on $H^\perp_{\mathbb{S}_0}$, as

$$\mathcal{L}_\perp^{sym} := \omega\cdot\partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_\perp + \mathcal{R}_\perp^{sym}, \quad \mathcal{R}_\perp^{sym} := \begin{pmatrix} \mathcal{R}^{sym}_{\perp,1} & \mathcal{R}^{sym}_{\perp,2} \\ \overline{\mathcal{R}^{sym}_{\perp,2}} & \overline{\mathcal{R}^{sym}_{\perp,1}} \end{pmatrix}, \tag{14.1}$$

where $\mathcal{R}^{sym}_{\perp,i}$, $i=1,2$, are defined by their matrix entries

$$(\mathcal{R}^{sym}_{\perp,i})_j^{j'}(\ell) := \begin{cases} (\mathcal{R}_{\perp,i})_j^{j'}(\ell) + (\mathcal{R}_{\perp,i})_j^{-j'}(\ell) & \text{if } jj'>0, \\ 0 & \text{if } jj'<0, \end{cases} \qquad j,j'\in\mathbb{S}_0^c\,, \qquad i=1,2, \tag{14.2}$$

and $\mathcal{R}_{\perp,i}$, $i=1,2$ are introduced in (13.13). Note that, in particular, $(\mathcal{R}^{sym}_{\perp,i})_j^{j'} = 0$, $i=1,2$ on the anti-diagonal $j'=-j$. Using definition (14.2), one has the following lemma.

Lemma 14.1. *The operator $\mathcal{R}_\perp^{sym}$ coincides with $\mathcal{R}_\perp$ on the subspace of functions even(x) in $H^\perp_{\mathbb{S}_0}\times H^\perp_{\mathbb{S}_0}$, namely*

$$\mathcal{R}_\perp h = \mathcal{R}_\perp^{sym} h\,, \quad \forall h\in H^\perp_{\mathbb{S}_0}\times H^\perp_{\mathbb{S}_0}\,, \quad h = h(\varphi,x) = even(x)\,. \tag{14.3}$$

$\mathcal{R}_\perp^{sym}$ is real, even and reversible, and it satisfies the same bounds (13.17), (13.18) *as $\mathcal{R}_\perp$.*

As a starting point of the recursive scheme, we consider the real, even, reversible linear operator $\mathcal{L}_\perp^{sym}$ in (14.1), acting on $H^\perp_{\mathbb{S}_0}$, defined for all $(\omega,\mathtt{h})\in\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$, which we rename

$$\mathcal{L}_0 := \mathcal{L}_\perp^{sym} := \omega\cdot\partial_\varphi\mathbb{I}_\perp + \mathrm{i}\mathcal{D}_0 + \mathcal{R}_0\,, \qquad \mathcal{D}_0 := \mathcal{D}_\perp, \quad \mathcal{R}_0 := \mathcal{R}_\perp^{sym}\,, \tag{14.4}$$

with

$$\mathcal{D}_0 := \begin{pmatrix} D_0 & 0 \\ 0 & -D_0 \end{pmatrix}, \qquad D_0 := \mathrm{diag}_{j\in\mathbb{S}_0^c}\,\mu_j^0\,, \qquad \mu_j^0 := \mathtt{m}_{\frac12}|j|^{\frac12}\tanh^{\frac12}(\mathtt{h}|j|) + r_j\,, \tag{14.5}$$

where $\mathtt{m}_{\frac12} := \mathtt{m}_{\frac12}(\omega,\mathtt{h})\in\mathbb{R}$ satisfies (11.24), $r_j := r_j(\omega,\mathtt{h})\in\mathbb{R}$, $r_j = r_{-j}$ satisfy (12.78), and

$$\mathcal{R}_0 := \begin{pmatrix} R_1^{(0)} & R_2^{(0)} \\ \overline{R_2^{(0)}} & \overline{R_1^{(0)}} \end{pmatrix}, \quad R_i^{(0)} : H^\perp_{\mathbb{S}_0}\to H^\perp_{\mathbb{S}_0}\,, \qquad i=1,2\,. \tag{14.6}$$

Notation. In this section we use the following notation: given an operator R, we denote by $\partial_{\varphi_i}^s\langle D\rangle^{\mathtt{m}}R\langle D\rangle^{\mathtt{m}}$ the operator $\langle D\rangle^{\mathtt{m}}\circ\big(\partial^s_{\varphi_i}R(\varphi)\big)\circ\langle D\rangle^{\mathtt{m}}$. Similarly $\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}R\langle D\rangle^{\mathtt{m}}$ denotes $\langle D\rangle^{\mathtt{m}}\circ\big(\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}R\big)\circ\langle D\rangle^{\mathtt{m}}$ where $\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}$ is introduced in Definition 2.7.

The operator $\mathcal{R}_0$ in (14.6) satisfies the tame estimates of Lemma 14.2 below. Define the constants

$$\begin{aligned} &\mathtt{b} := [\mathtt{a}] + 2\in\mathbb{N}\,, \quad \mathtt{a} := \max\{3\tau_1, \chi(\tau+1)(4\mathtt{d}+1)+1\}\,, \quad \chi := 3/2\,, \\ &\tau_1 := \tau(k_0+1)+k_0+\mathtt{m}\,, \quad \mathtt{m} := \mathtt{d}(k_0+1)+\frac{k_0}{2}\,, \end{aligned} \tag{14.7}$$

where $\mathtt{d} > \frac34 k_0^*$, by (4.22). The condition $\mathtt{a}\geq\chi(\tau+1)(4\mathtt{d}+1)+1$ in (14.7) will be used in Section 15 in order to verify inequality (15.5). Proposition 13.3 implies that $\mathcal{R}_0$ satisfies the tame estimates of Lemma 14.2 by fixing the constant M large enough (which means that one has to perform a sufficiently large number of regularizing steps in Sections 10 and 12), namely

$$M := \Big[2\mathtt{m} + 2\mathtt{b} + 1 + \frac{\mathtt{b}+s_0+k_0}{2}\Big] + 1\in\mathbb{N} \tag{14.8}$$

where $[\,\cdot\,]$ denotes the integer part, and $\mathtt{m}$ and $\mathtt{b}$ are defined in (14.7). We also set

$$\mu(\mathtt{b}) := \aleph(M, s_0 + \mathtt{b}) \,, \tag{14.9}$$

where the constant $\aleph(M, s_0 + \mathtt{b})$ is given in Proposition 13.3.

Lemma 14.2. (Tame estimates of $\mathcal{R}_0 := \mathcal{R}_\perp^{sym}$) *Assume* (6.9) *with $\mu_0 \geq \mu(\mathtt{b})$. Then $\mathcal{R}_0$ in* (14.4) *satisfies the following property: the operators*

$$\langle D\rangle^{\mathtt{m}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+1}, \qquad \partial_{\varphi_i}^{s_0}\langle D\rangle^{\mathtt{m}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+1} \,, \qquad \forall i = 1, \ldots, \nu \,, \tag{14.10}$$

$$\langle D\rangle^{\mathtt{m}+\mathtt{b}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+\mathtt{b}+1}, \quad \partial_{\varphi_i}^{s_0+\mathtt{b}}\langle D\rangle^{\mathtt{m}+\mathtt{b}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+\mathtt{b}+1} \,, \tag{14.11}$$

where $\mathtt{m}, \mathtt{b}$ are defined in (14.7)*, are $\mathcal{D}^{k_0}$-tame with tame constants*

$$\mathbb{M}_0(s) := \max_{i=1,\ldots,\nu} \big\{\mathfrak{M}_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+1}}(s), \mathfrak{M}_{\partial_{\varphi_i}^{s_0}\langle D\rangle^{\mathtt{m}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+1}}(s)\big\} \tag{14.12}$$

$$\mathbb{M}_0(s, \mathtt{b}) := \max_{i=1,\ldots,\nu} \big\{\mathfrak{M}_{\langle D\rangle^{\mathtt{m}+\mathtt{b}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+\mathtt{b}+1}}(s), \mathfrak{M}_{\partial_{\varphi_i}^{s_0+\mathtt{b}}\langle D\rangle^{\mathtt{m}+\mathtt{b}}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+\mathtt{b}+1}}(s)\big\} \tag{14.13}$$

satisfying, for all $s_0 \leq s \leq S$,

$$\mathfrak{M}_0(s, \mathtt{b}) := \max\{\mathbb{M}_0(s), \mathbb{M}_0(s, \mathtt{b})\} \lesssim_S \varepsilon\gamma^{-2(M+1)}\big(1 + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})}^{k_0,\gamma}\big) \,. \tag{14.14}$$

In particular we have

$$\mathfrak{M}_0(s_0, \mathtt{b}) \leq C(S)\varepsilon\gamma^{-2(M+1)} \,. \tag{14.15}$$

Moreover, for all $i = 1, \ldots, \nu$, $\beta \in \mathbb{N}$, $\beta \leq s_0 + \mathtt{b}$, we have

$$\|\partial_{\varphi_i}^{\beta}\langle D\rangle^{\mathtt{m}}\Delta_{12}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+1}\|_{\mathcal{L}(H^{s_0})}, \|\partial_{\varphi_i}^{\beta}\langle D\rangle^{\mathtt{m}+\mathtt{b}}\Delta_{12}\mathcal{R}_0\langle D\rangle^{\mathtt{m}+\mathtt{b}+1}\|_{\mathcal{L}(H^{s_0})} \lesssim_S \varepsilon\gamma^{-2(M+1)}\|\Delta_{12} i\|_{s_0+\mu(\mathtt{b})} \,. \tag{14.16}$$

Proof. Estimate (14.14) follows by Lemma 14.1, by (13.17) 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 (14.7), (14.8), (14.9). Estimates (14.16) follow similarly, applying (13.18) with the same choices of m_1, m_2 and with $s_1 = s_0$. $\square$

We perform the almost-reducibility of $\mathcal{L}_0$ along the scale

$$N_{-1} := 1 \,, \quad N_{\mathtt{n}} := N_0^{\chi^{\mathtt{n}}} \quad \forall \mathtt{n} \geq 0 \,, \quad \chi = 3/2 \,, \tag{14.17}$$

requiring inductively at each step the second order Melnikov non-resonance conditions in (14.26). Note that the non-diagonal remainder $\mathcal{R}_{\mathtt{n}}$ in (14.19) is small according to the first inequality in (14.25).

Theorem 14.3. (Almost-reducibility of $\mathcal{L}_0$: KAM iteration) *There exists $\tau_2 := \tau_2(\tau, \nu) > \tau_1 + \mathtt{a}$ (where $\tau_1, \mathtt{a}$ are defined in* (14.7)*) 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*

$$\varepsilon\gamma^{-2(M+1)} \leq \delta_0, \quad N_0^{\tau_2}\mathfrak{M}_0(s_0, \mathtt{b})\gamma^{-1} \leq 1 \tag{14.18}$$

(see (14.15)*), then, for all $n \in \mathbb{N}$, $\mathtt{n} = 0, 1, \ldots, n$:*

(S1)$_{\mathtt{n}}$ *There exists a real, even and reversible operator*

$$\mathcal{L}_{\mathtt{n}} := \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_{\mathtt{n}} + \mathcal{R}_{\mathtt{n}} \,, \quad \mathcal{D}_{\mathtt{n}} := \begin{pmatrix} D_{\mathtt{n}} & 0 \\ 0 & -D_{\mathtt{n}} \end{pmatrix} \,, \quad D_{\mathtt{n}} := \mathrm{diag}_{j\in\mathbb{S}_0^c}\mu_j^{\mathtt{n}} \,, \tag{14.19}$$

defined for all $(\omega, \mathtt{h})$ in $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ where $\mu_j^{\mathtt{n}}$ are k_0 times differentiable functions of the form

$$\mu_j^{\mathtt{n}}(\omega, \mathtt{h}) := \mu_j^0(\omega, \mathtt{h}) + r_j^{\mathtt{n}}(\omega, \mathtt{h}) \in \mathbb{R} \tag{14.20}$$

where μ_j^0 are defined in (14.5)*, satisfying*

$$\mu_j^{\mathtt{n}} = \mu_{-j}^{\mathtt{n}} \,, \quad \textit{i.e. } r_j^{\mathtt{n}} = r_{-j}^{\mathtt{n}} \,, \quad |r_j^{\mathtt{n}}|^{k_0,\gamma} \leq C(S, \mathtt{b})\varepsilon\gamma^{-2(M+1)}|j|^{-2\mathtt{m}} \,, \quad \forall j \in \mathbb{S}_0^c \tag{14.21}$$

and, for $\mathtt{n} \geq 1$,

$$|\mu_j^{\mathtt{n}} - \mu_j^{\mathtt{n}-1}|^{k_0,\gamma} \leq C|j|^{-2\mathtt{m}} \mathfrak{M}^{\sharp}_{\langle D\rangle^{\mathtt{m}} \mathcal{R}_{\mathtt{n}-1} \langle D\rangle^{\mathtt{m}}}(s_0) \leq C(S,\mathtt{b})\varepsilon\gamma^{-2(M+1)}|j|^{-2\mathtt{m}} N_{\mathtt{n}-2}^{-\mathtt{a}} . \tag{14.22}$$

The remainder

$$\mathcal{R}_{\mathtt{n}} := \begin{pmatrix} R_1^{(\mathtt{n})} & R_2^{(\mathtt{n})} \\ \overline{R}_2^{(\mathtt{n})} & \overline{R}_1^{(\mathtt{n})} \end{pmatrix} \tag{14.23}$$

satisfies

$$(R_1^{(\mathtt{n})})_j^{j'}(\ell) = (R_2^{(\mathtt{n})})_j^{j'}(\ell) = 0 \quad \forall (\ell, j, j'), \ jj' < 0, \tag{14.24}$$

and it is $\mathcal{D}^{k_0}$*-modulo-tame: more precisely, the operators* $\langle D\rangle^{\mathtt{m}} \mathcal{R}_{\mathtt{n}} \langle D\rangle^{\mathtt{m}}$ *and* $\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^{\mathtt{m}} \mathcal{R}_{\mathtt{n}} \langle D\rangle^{\mathtt{m}}$ *are* $\mathcal{D}^{k_0}$*-modulo-tame and there exists a constant* $C_* := C_*(s_0, \mathtt{b}) > 0$ *such that, for any* $s \in [s_0, S]$,

$$\mathfrak{M}^{\sharp}_{\langle D\rangle^{\mathtt{m}} \mathcal{R}_{\mathtt{n}} \langle D\rangle^{\mathtt{m}}}(s) \leq \frac{C_* \mathfrak{M}_0(s,\mathtt{b})}{N_{\mathtt{n}-1}^{\mathtt{a}}} , \quad \mathfrak{M}^{\sharp}_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^{\mathtt{m}} \mathcal{R}_{\mathtt{n}} \langle D\rangle^{\mathtt{m}}}(s) \leq C_* \mathfrak{M}_0(s,\mathtt{b}) N_{\mathtt{n}-1} . \tag{14.25}$$

Define the sets $\Lambda_{\mathtt{n}}^{\gamma}$ *by* $\Lambda_0^{\gamma} := \mathtt{DC}(2\gamma, \tau) \times [\mathtt{h}_1, \mathtt{h}_2]$, *and, for all* $\mathtt{n} \geq 1$,

$$\begin{aligned} \Lambda_{\mathtt{n}}^{\gamma} := \Lambda_{\mathtt{n}}^{\gamma}(i) := \Big\{ & \lambda = (\omega, \mathtt{h}) \in \Lambda_{\mathtt{n}-1}^{\gamma} \ : \\ & |\omega \cdot \ell + \mu_j^{\mathtt{n}-1} - \mu_{j'}^{\mathtt{n}-1}| \geq \gamma j^{-\mathtt{d}} j'^{-\mathtt{d}} \langle \ell \rangle^{-\tau} \ \ \forall |\ell|, |j-j'| \leq N_{\mathtt{n}-1}, \ \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+, \ \ (\ell, j, j') \neq (0, j, j), \\ & |\omega \cdot \ell + \mu_j^{\mathtt{n}-1} + \mu_{j'}^{\mathtt{n}-1}| \geq \gamma (\sqrt{j} + \sqrt{j'}) \langle \ell \rangle^{-\tau} \ \ \forall |\ell|, |j-j'| \leq N_{\mathtt{n}-1} , \ \ j, j' \in \mathbb{N}^+ \setminus \mathbb{S}^+ \Big\} . \end{aligned} \tag{14.26}$$

For $\mathtt{n} \geq 1$, *there exists a real, even and reversibility preserving map, defined for all* $(\omega, \mathtt{h})$ *in* $\mathbb{R}^{\nu} \times [\mathtt{h}_1, \mathtt{h}_2]$, *of the form*

$$\Phi_{\mathtt{n}-1} := \mathbb{I}_{\perp} + \Psi_{\mathtt{n}-1} , \quad \Psi_{\mathtt{n}-1} := \begin{pmatrix} \Psi_{\mathtt{n}-1,1} & \Psi_{\mathtt{n}-1,2} \\ \overline{\Psi}_{\mathtt{n}-1,2} & \overline{\Psi}_{\mathtt{n}-1,1} \end{pmatrix} \tag{14.27}$$

such that for all $\lambda = (\omega, \mathtt{h}) \in \Lambda_{\mathtt{n}}^{\gamma}$ *the following conjugation formula holds:*

$$\mathcal{L}_{\mathtt{n}} = \Phi_{\mathtt{n}-1}^{-1} \mathcal{L}_{\mathtt{n}-1} \Phi_{\mathtt{n}-1} . \tag{14.28}$$

The operators $\langle D\rangle^{\pm\mathtt{m}} \Psi_{\mathtt{n}-1} \langle D\rangle^{\mp\mathtt{m}}$ *and* $\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^{\pm\mathtt{m}} \Psi_{\mathtt{n}-1} \langle D\rangle^{\mp\mathtt{m}}$ *are* $\mathcal{D}^{k_0}$*-modulo-tame on* $\mathbb{R}^{\nu} \times [\mathtt{h}_1, \mathtt{h}_2]$ *with modulo-tame constants satisfying, for all* $s \in [s_0, S]$, *(*$\tau_1, \mathtt{a}$ *are defined in* (14.7)*)*

$$\mathfrak{M}^{\sharp}_{\langle D\rangle^{\pm\mathtt{m}} \Psi_{\mathtt{n}-1} \langle D\rangle^{\mp\mathtt{m}}}(s) \leq C(s_0,\mathtt{b})\gamma^{-1} N_{\mathtt{n}-1}^{\tau_1} N_{\mathtt{n}-2}^{-\mathtt{a}} \mathfrak{M}_0(s,\mathtt{b}) , \tag{14.29}$$

$$\mathfrak{M}^{\sharp}_{\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^{\pm\mathtt{m}} \Psi_{\mathtt{n}-1} \langle D\rangle^{\mp\mathtt{m}}}(s) \leq C(s_0,\mathtt{b})\gamma^{-1} N_{\mathtt{n}-1}^{\tau_1} N_{\mathtt{n}-2} \mathfrak{M}_0(s,\mathtt{b}) , \tag{14.30}$$

$$\mathfrak{M}^{\sharp}_{\Psi_{\mathtt{n}-1}}(s) \leq C(s_0,\mathtt{b})\gamma^{-1} N_{\mathtt{n}-1}^{\tau_1} N_{\mathtt{n}-2}^{-\mathtt{a}} \mathfrak{M}_0(s,\mathtt{b}) . \tag{14.31}$$

(S2)$_{\mathtt{n}}$ *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* (14.15). *Then for all* $(\omega, \mathtt{h}) \in \Lambda_{\mathtt{n}}^{\gamma_1}(i_1) \cap \Lambda_{\mathtt{n}}^{\gamma_2}(i_2)$ *with* $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$, *the following estimates hold*

$$|||\langle D\rangle^{\mathtt{m}} \Delta_{12} \mathcal{R}_{\mathtt{n}} \langle D\rangle^{\mathtt{m}}|||_{\mathcal{L}(H^{s_0})} \lesssim_{S,\mathtt{b}} \varepsilon\gamma^{-2(M+1)} N_{\mathtt{n}-1}^{-\mathtt{a}} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})}, \tag{14.32}$$

$$|||\langle \partial_{\varphi,x}\rangle^{\mathtt{b}} \langle D\rangle^{\mathtt{m}} \Delta_{12} \mathcal{R}_{\mathtt{n}} \langle D\rangle^{\mathtt{m}}|||_{\mathcal{L}(H^{s_0})} \lesssim_{S,\mathtt{b}} \varepsilon\gamma^{-2(M+1)} N_{\mathtt{n}-1} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} . \tag{14.33}$$

Moreover for $\mathtt{n} \geq 1$, *for all* $j \in \mathbb{S}_0^c$,

$$\big|\Delta_{12}(r_j^{\mathtt{n}} - r_j^{\mathtt{n}-1})\big| \lesssim_{S,\mathtt{b}} \varepsilon\gamma^{-2(M+1)} |j|^{-2\mathtt{m}} N_{\mathtt{n}-2}^{-\mathtt{a}} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} , \tag{14.34}$$

$$|\Delta_{12} r_j^{\mathtt{n}}| \lesssim_{S,\mathtt{b}} \varepsilon\gamma^{-2(M+1)} |j|^{-2\mathtt{m}} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} . \tag{14.35}$$

(S3)$_{\mathtt{n}}$ *Let* i_1, i_2 *be like in* **(S2)**$_{\mathtt{n}}$ *and* $0 < \rho \leq \gamma/2$. *Then*

$$C(S) N_{\mathtt{n}-1}^{(\tau+1)(4\mathtt{d}+1)} \gamma^{-4\mathtt{d}} \|i_2 - i_1\|_{s_0+\mu(\mathtt{b})} \leq \rho \quad \Longrightarrow \quad \Lambda_{\mathtt{n}}^{\gamma}(i_1) \subseteq \Lambda_{\mathtt{n}}^{\gamma-\rho}(i_2) . \tag{14.36}$$

We make some comments:

1. Note that in (14.34)-(14.35) 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 $|\omega\cdot\ell + \mu_j^{\mathtt{n}-1} - \mu_{j'}^{\mathtt{n}-1}| \geq \gamma |j|^{-\mathtt{d}}|j'|^{-\mathtt{d}}\langle\ell\rangle^{-\tau}$ lose regularity both in φ and in x, for the convergence of the reducibility scheme we use the smoothing operators Π_N, defined in (2.25), which regularize in both φ and x. As a consequence, the natural smallness condition to impose at the zero step of the recursion is (14.25) at $\mathtt{n}=0$ that we verify in the step $(\mathbf{S1})_0$ thanks to Lemma 2.35 and (14.14).
3. An important point of Theorem 14.3 is to require bound (14.18) for $\mathfrak{M}_0(s_0,\mathtt{b})$ only in low norm, which is verified in Lemma 14.2. On the other hand, Theorem 14.3 provides the smallness (14.25) of the tame constants $\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_{\mathtt{n}}\langle D\rangle^{\mathtt{m}}}(s)$ and proves that $\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_{\mathtt{n}}\langle D\rangle^{\mathtt{m}}}(s,\mathtt{b})$, $\mathtt{n}\geq 0$, do not diverge too much.

Theorem 14.3 implies that the invertible operator

$$\mathcal{U}_n := \Phi_0 \circ \ldots \circ \Phi_{n-1}\,, \quad n \geq 1, \tag{14.37}$$

has almost-diagonalized $\mathcal{L}_0$, i.e. (14.42) below holds. As a corollary, we deduce the following theorem.

Theorem 14.4. (Almost-reducibility of $\mathcal{L}_0$) *Assume* (6.9) *with $\mu_0 \geq \mu(\mathtt{b})$. Let $\mathcal{R}_0 = \mathcal{R}_\perp^{sym}$, $\mathcal{L}_0 = \mathcal{L}_\perp^{sym}$ in* (14.1)-(14.2)*. 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*

$$N_0^{\tau_2}\varepsilon\gamma^{-(2M+3)} \leq \delta_0 \tag{14.38}$$

holds, where the constant $\tau_2 := \tau_2(\tau,\nu)$ is defined in Theorem 14.3 and M is defined in (14.8)*, then, for all $n\in\mathbb{N}$, for all $\lambda = (\omega,\mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1,\mathtt{h}_2]$, the operator $\mathcal{U}_n$ in* (14.37) *and its inverse $\mathcal{U}_n^{-1}$ are real, even, reversibility preserving, and $\mathcal{D}^{k_0}$-modulo-tame, with*

$$\mathfrak{M}^\sharp_{\mathcal{U}_n^{\pm 1}-\mathbb{I}_\perp}(s) \lesssim_S \varepsilon\gamma^{-(2M+3)}N_0^{\tau_1}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu(\mathtt{b})}\big) \quad \forall s_0 \leq s \leq S\,, \tag{14.39}$$

where τ_1 is defined in (14.7)*.*

The operator $\mathcal{L}_n = \omega\cdot\partial_\varphi\mathbb{I}_\perp + \mathrm{i}\mathcal{D}_n + \mathcal{R}_n$ defined in (14.19) *(with $\mathtt{n}=n$) 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*

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s) \lesssim_S \varepsilon\gamma^{-2(M+1)}N_{n-1}^{-\mathtt{a}}\big(1 + \|\mathfrak{I}_0\|^{k_0,\gamma}_{s+\mu(\mathtt{b})}\big) \quad \forall s_0 \leq s \leq S\,. \tag{14.40}$$

Moreover, for all $\lambda = (\omega,\mathtt{h})$ in the set

$$\mathtt{\Lambda}_n^\gamma = \bigcap_{\mathtt{n}=0}^{n} \mathtt{\Lambda}_{\mathtt{n}}^\gamma \tag{14.41}$$

defined in (14.26)*, the following conjugation formula holds:*

$$\mathcal{L}_n = \mathcal{U}_n^{-1}\mathcal{L}_0\mathcal{U}_n\,. \tag{14.42}$$

Proof. Assumption (14.18) of Theorem 14.3 holds by (14.14), (6.9) with $\mu_0 \geq \mu(\mathtt{b})$, and (14.38). Estimate (14.40) follows by (14.25) (for $\mathtt{n}=n$) and (14.14). It remains to prove (14.39). The estimates of $\mathfrak{M}^\sharp_{\Phi_{\mathtt{n}}^{\pm1}-\mathbb{I}_\perp}(s)$, $\mathtt{n}=0,\ldots,n-1$, are obtained by using (14.31), (14.18) and Lemma 2.32. Then the estimate of $\mathcal{U}_n^{\pm1}-\mathbb{I}_\perp$ follows as in the proof of Theorem 7.5 in [21], using Lemma 2.31. □

14.1 Proof of Theorem 14.3

Initialization.

PROOF OF $(\mathbf{S1})_0$. The real, even and reversible operator $\mathcal{L}_0$ defined in (14.4)-(14.6) has the form (14.19)-(14.20) for $\mathtt{n}=0$ with $r_j^0(\omega,\mathtt{h}) = 0$, and (14.21) holds trivially. Moreover (14.24) is satisfied for $\mathtt{n}=0$ by the definition of $\mathcal{R}_0 := \mathcal{R}_\perp^{sym}$ in (14.2). The estimate (14.25) for $\mathtt{n}=0$ follows by applying Lemma 2.35 to $A \in \{R_1^{(0)}, R_2^{(0)}\}$ and by recalling definition of $\mathfrak{M}_0(s,\mathtt{b})$ in (14.14).

PROOF OF $(\mathbf{S2})_0$. The proof of (14.32), (14.33) for $\mathtt{n}=0$ follows similarly using Lemma 2.35 and (14.16).

PROOF OF $(\mathbf{S3})_0$. It is trivial because, by definition, $\mathtt{\Lambda}_0^\gamma = \mathtt{DC}(2\gamma,\tau)\times[\mathtt{h}_1,\mathtt{h}_2] \subseteq \mathtt{DC}(2\gamma-2\rho,\tau)\times[\mathtt{h}_1,\mathtt{h}_2] = \mathtt{\Lambda}_0^{\gamma-\rho}$.

14.1.1 Reducibility step

In this section we describe the inductive step and show how to define $\mathcal{L}_{\mathtt{n}+1}$ (and $\Psi_{\mathtt{n}}$, $\Phi_{\mathtt{n}}$, etc). To simplify the notation we drop the index $\mathtt{n}$ and write $+$ instead of $\mathtt{n}+1$, so that we write $\mathcal{L} := \mathcal{L}_{\mathtt{n}}$, $\mathcal{D} := \mathcal{D}_{\mathtt{n}}$, $D := D_{\mathtt{n}}$, $\mu_j = \mu_j^{\mathtt{n}}$, $\mathcal{R} := \mathcal{R}_{\mathtt{n}}$, $R_1 := R_1^{(\mathtt{n})}$, $R_2 := R_2^{(\mathtt{n})}$, and $\mathcal{L}_+ := \mathcal{L}_{\mathtt{n}+1}$, $\mathcal{D}_+ := \mathcal{D}_{\mathtt{n}+1}$, and so on.

We conjugate the operator $\mathcal{L}$ in (14.19) by a transformation of the form (see (14.27))

$$\Phi := \mathbb{I}_\perp + \Psi\,, \qquad \Psi := \begin{pmatrix} \Psi_1 & \Psi_2 \\ \overline{\Psi}_2 & \overline{\Psi}_1 \end{pmatrix}. \tag{14.43}$$

We have

$$\mathcal{L}\Phi = \Phi(\omega\cdot\partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}) + (\omega\cdot\partial_\varphi\Psi + \mathrm{i}[\mathcal{D},\Psi] + \Pi_N\mathcal{R}) + \Pi_N^\perp\mathcal{R} + \mathcal{R}\Psi \tag{14.44}$$

where the projector Π_N is defined in (2.25), $\Pi_N^\perp := \mathbb{I}_2 - \Pi_N$, and $\omega\cdot\partial_\varphi\Psi$ is the commutator $[\omega\cdot\partial_\varphi,\Psi]$. We want to solve the homological equation

$$\omega\cdot\partial_\varphi\Psi + \mathrm{i}[\mathcal{D},\Psi] + \Pi_N\mathcal{R} = [\mathcal{R}] \tag{14.45}$$

where

$$[\mathcal{R}] := \begin{pmatrix} [R_1] & 0 \\ 0 & [\overline{R}_1] \end{pmatrix}\,, \quad [R_1] := \mathrm{diag}_{j\in\mathbb{S}_0^c}(R_1)_j^j(0)\,. \tag{14.46}$$

By (14.19), (14.23), (14.43), equation (14.45) is equivalent to the two scalar homological equations

$$\omega\cdot\partial_\varphi\Psi_1 + \mathrm{i}[D,\Psi_1] + \Pi_N R_1 = [R_1]\,, \qquad \omega\cdot\partial_\varphi\Psi_2 + \mathrm{i}(D\Psi_2 + \Psi_2 D) + \Pi_N R_2 = 0 \tag{14.47}$$

(note that $[R_1] = [\Pi_N R_1]$). We choose the solution of (14.47) given by

$$(\Psi_1)_j^{j'}(\ell) := \begin{cases} -\dfrac{(R_1)_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + \mu_j - \mu_{j'})} & \forall(\ell,j,j') \neq (0,j,\pm j)\,,\ |\ell|, |j-j'| \leq N, \\ 0 & \text{otherwise;} \end{cases} \tag{14.48}$$

$$(\Psi_2)_j^{j'}(\ell) := \begin{cases} -\dfrac{(R_2)_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + \mu_j + \mu_{j'})} & \forall(\ell,j,j') \in \mathbb{Z}^\nu\times\mathbb{S}_0^c\times\mathbb{S}_0^c\,,\ |\ell|, |j-j'| \leq N, \\ 0 & \text{otherwise.} \end{cases} \tag{14.49}$$

Note that, since $\mu_j = \mu_{-j}$ for all $j\in\mathbb{S}_0^c$ (see (14.21)), the denominators in (14.48), (14.49) are different from zero for $(\omega,\mathtt{h}) \in \Lambda_{\mathtt{n}+1}^\gamma$ (see (14.26) with $\mathtt{n} \rightsquigarrow \mathtt{n}+1$) and the maps Ψ_1, Ψ_2 are well defined on $\Lambda_{\mathtt{n}+1}^\gamma$. Also note that the term $[R_1]$ in (14.46) (which is the term we are not able to remove by conjugation with Ψ_1 in (14.47)) contains only the diagonal entries $j'=j$ and not the anti-diagonal ones $j'=-j$, because $\mathcal{R}$ is zero on $j'=-j$ by (14.24). Thus, by construction,

$$(\Psi_1)_j^{j'}(\ell) = (\Psi_2)_j^{j'}(\ell) = 0 \quad \forall(\ell,j,j'),\ jj' < 0\,. \tag{14.50}$$

Lemma 14.5. (Homological equations) *The operators Ψ_1, Ψ_2 defined in* (14.48), (14.49) *(which, for all $\lambda\in\Lambda_{\mathtt{n}+1}^\gamma$, solve the homological equations* (14.47)*) admit an extension to the whole parameter space $\mathbb{R}^\nu\times[\mathtt{h}_1,\mathtt{h}_2]$. Such extended operators are $\mathcal{D}^{k_0}$-modulo-tame with modulo-tame constants satisfying*

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\pm\mathtt{m}}\Psi\langle D\rangle^{\mp\mathtt{m}}}(s) \lesssim_{k_0} N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}\langle D\rangle^{\mathtt{m}}}(s), \tag{14.51}$$

$$\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\Psi\langle D\rangle^{\mp\mathtt{m}}}(s) \lesssim_{k_0} N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}\langle D\rangle^{\mathtt{m}}}(s) \tag{14.52}$$

$$\mathfrak{M}^\sharp_{\Psi}(s) \lesssim_{k_0} N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\mathcal{R}}(s) \tag{14.53}$$

where $\tau_1, \mathtt{b}, \mathtt{m}$ are defined in (14.7).

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 $(\omega, \mathtt{h}) \in \Lambda^{\gamma_1}_{\mathtt{n}+1}(i_1) \cap \Lambda^{\gamma_2}_{\mathtt{n}+1}(i_2)$,

$$
\begin{aligned}
\| \, |\langle D\rangle^{\pm\mathtt{m}} \Delta_{12}\Psi \langle D\rangle^{\mp\mathtt{m}}| \, \|_{\mathcal{L}(H^{s_0})} &\lesssim N^{2\tau+2\mathtt{d}+\frac{1}{2}}\gamma^{-1}\big(\| \, |\langle D\rangle^{\mathtt{m}} \mathcal{R}(i_2)\langle D\rangle^{\mathtt{m}}| \, \|_{\mathcal{L}(H^{s_0})} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} \\
&\quad + \| \, |\langle D\rangle^{\mathtt{m}} \Delta_{12}\mathcal{R}\langle D\rangle^{\mathtt{m}}| \, \|_{\mathcal{L}(H^{s_0})}\big) , \qquad (14.54)\\
\| \, |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}} \Delta_{12}\Psi \langle D\rangle^{\mp\mathtt{m}}| \, \|_{\mathcal{L}(H^{s_0})} &\lesssim N^{2\tau+2\mathtt{d}+\frac{1}{2}}\gamma^{-1}\big(\| \, |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}} \mathcal{R}(i_2)\langle D\rangle^{\mathtt{m}}| \, \|_{\mathcal{L}(H^{s_0})} \|i_1 - i_2\|_{s_0+\mu(\mathtt{b})} \\
&\quad + \| \, |\langle \partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}} \Delta_{12}\mathcal{R}\langle D\rangle^{\mathtt{m}}| \, \|_{\mathcal{L}(H^{s_0})}\big) . \qquad (14.55)
\end{aligned}
$$

Moreover Ψ is real, even and reversibility preserving.

Proof. For all $\lambda \in \Lambda^{\gamma}_{\mathtt{n}+1}$, $(\ell, j, j') \neq (0, j, \pm j)$, $j, j' \in \mathbb{S}_0^c$ $|\ell|, |j - j'| \leq N$, we have the small divisor estimate

$$|\omega \cdot \ell + \mu_j - \mu_{j'}| = |\omega \cdot \ell + \mu_{|j|} - \mu_{|j'|}| \geq \gamma |j|^{-\mathtt{d}}|j'|^{-\mathtt{d}}\langle \ell\rangle^{-\tau}$$

by (14.26), because $||j| - |j'|| \leq |j - j'| \leq N$. As in Lemma B.4, we extend the restriction to $F = \Lambda^{\gamma}_{\mathtt{n}+1}$ of the function $(\omega \cdot \ell + \mu_j - \mu_{j'})^{-1}$ to the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ 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 (B.14). For $n \geq 1$, $x > 0$, the n-th derivative of the function $\tanh^{\frac{1}{2}}(x)$ is $P_n(\tanh(x)) \tanh^{\frac{1}{2}-n}(x)(1 - \tanh^2(x))$, where P_n is a polynomial of degree $\leq 2n - 2$. Hence $|\partial^n_{\mathtt{h}}\{\tanh^{\frac{1}{2}}(\mathtt{h}|j|)\}| \leq C$ 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 (14.20), (14.21), (14.5), (11.24), (12.78) (and recalling that μ_j here denotes $\mu_j^{\mathtt{n}}$), since $\varepsilon\gamma^{-2(M+1)} \leq \gamma$, we deduce that

$$\gamma^{|\alpha|}|\partial^\alpha_\lambda \mu_j| \lesssim \gamma |j|^{\frac{1}{2}} \quad \forall \alpha \in \mathbb{N}^{\nu+1}, \ 1 \leq |\alpha| \leq k_0 . \qquad (14.56)$$

Since $\gamma^{|\alpha|}|\partial^\alpha_\lambda(\omega \cdot \ell)| \leq \gamma|\ell|$ for all $|\alpha| \geq 1$, we conclude that

$$\gamma^{|\alpha|}|\partial^\alpha_\lambda(\omega \cdot \ell + \mu_j - \mu_{j'})| \lesssim \gamma(|\ell| + |j|^{\frac{1}{2}} + |j'|^{\frac{1}{2}}) \lesssim \gamma\langle \ell\rangle|j|^{\frac{1}{2}}|j'|^{\frac{1}{2}} , \quad \forall \ 1 \leq |\alpha| \leq k_0 . \qquad (14.57)$$

Thus (B.14) holds with $M = C\gamma\langle \ell\rangle|j|^{\frac{1}{2}}|j'|^{\frac{1}{2}}$ (which is $\geq \rho$) and (B.15) implies that

$$|g_{\ell,j,j'}|^{k_0,\gamma} \lesssim \gamma^{-1}\langle \ell\rangle^{\tau(k_0+1)+k_0}|j|^{\mathtt{m}}|j'|^{\mathtt{m}} \quad \text{with} \quad \mathtt{m} = (k_0 + 1)\mathtt{d} + \frac{k_0}{2} \qquad (14.58)$$

defined in (14.7). Formula (14.48) with $(\omega\cdot\ell+\mu_j-\mu_{j'})^{-1}$ replaced by $g_{\ell,j,j'}(\lambda)$ defines the extended operator Ψ_1 to $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$. Analogously, we construct an extension of the function $(\omega \cdot \ell + \mu_j + \mu_{j'})^{-1}$ to the whole $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, and we obtain an extension of the operator Ψ_2 in (14.49).

Proof of (14.51), (14.52), (14.53). We prove (14.52) for Ψ_1, then the estimate for Ψ_2 follows in the same way, as well as (14.51), (14.53). Furthermore, we analyze $\langle D\rangle^{\mathtt{m}}\partial^k_\lambda\Psi_1\langle D\rangle^{-\mathtt{m}}$, 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 $|k| \leq k_0$,

$$
\begin{aligned}
|\partial^k_\lambda(\Psi_1)^{j'}_j(\ell)| &\lesssim \sum_{k_1+k_2=k} |\partial^{k_1}_\lambda g_{\ell,j,j'}||\partial^{k_2}_\lambda(R_1)^{j'}_j(\ell)| \lesssim \sum_{k_1+k_2=k} \gamma^{-|k_1|}|g_{\ell,j,j'}|^{k_0,\gamma}|\partial^{k_2}_\lambda(R_1)^{j'}_j(\ell)| \\
&\overset{(14.58)}{\lesssim} \langle \ell\rangle^{\tau(k_0+1)+k_0}|j|^{\mathtt{m}}|j'|^{\mathtt{m}}\gamma^{-1-|k|}\sum_{|k_2|\leq|k|}\gamma^{|k_2|}|\partial^{k_2}_\lambda(R_1)^{j'}_j(\ell)| . \qquad (14.59)
\end{aligned}
$$

For $|j - j'| \leq N$, $j, j' \neq 0$, one has

$$|j|^{2\mathtt{m}} \lesssim |j|^{\mathtt{m}}(|j'|^{\mathtt{m}} + |j - j'|^{\mathtt{m}}) \lesssim |j|^{\mathtt{m}}(|j'|^{\mathtt{m}} + N^{\mathtt{m}}) \lesssim |j|^{\mathtt{m}}|j'|^{\mathtt{m}}N^{\mathtt{m}}. \qquad (14.60)$$

Hence, by (14.59) and (14.60), for all $|k| \leq k_0$, $j, j' \in \mathbb{S}_0^c$, $\ell \in \mathbb{Z}^\nu$, $|\ell| \leq N$, $|j - j'| \leq N$, one has

$$|j|^{\mathtt{m}}|\partial^k_\lambda(\Psi_1)^{j'}_j(\ell)||j'|^{-\mathtt{m}} \lesssim N^{\tau_1}\gamma^{-1-|k|}\sum_{|k_2|\leq|k|}\gamma^{|k_2|}|j|^{\mathtt{m}}|\partial^{k_2}_\lambda(R_1)^{j'}_j(\ell)||j'|^{\mathtt{m}} \qquad (14.61)$$

where $\tau_1 = \tau(k_0+1) + k_0 + \mathtt{m}$ is defined in (14.7). Therefore, for all $0 \le |k| \le k_0$, we get

$$
\begin{aligned}
& \| \, |\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} \partial_\lambda^k \Psi_1 \langle D \rangle^{-\mathtt{m}} | \, h \|_s^2 \\
& \le \sum_{\ell,j} \langle \ell, j \rangle^{2s} \Big(\sum_{|\ell'-\ell|,|j'-j| \le N} \langle \ell - \ell', j - j' \rangle^{\mathtt{b}} \langle j \rangle^{\mathtt{m}} |\partial_\lambda^k (\Psi_1)_j^{j'}(\ell - \ell')| \langle j' \rangle^{-\mathtt{m}} |h_{\ell',j'}| \Big)^2 \\
& \overset{(14.61)}{\lesssim_{k_0}} N^{2\tau_1} \gamma^{-2(1+|k|)} \sum_{|k_2| \le |k|} \gamma^{2|k_2|} \sum_{\ell,j} \langle \ell, j \rangle^{2s} \Big(\sum_{\ell',j'} |\langle \ell - \ell', j - j' \rangle^{\mathtt{b}} \langle j \rangle^{\mathtt{m}} \partial_\lambda^{k_2} (R_1)_j^{j'}(\ell - \ell') \langle j' \rangle^{\mathtt{m}} | |h_{\ell',j'}| \Big)^2 \\
& \lesssim_{k_0} N^{2\tau_1} \gamma^{-2(1+|k|)} \sum_{|k_2| \le |k|} \gamma^{2|k_2|} \| \, |\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} \partial_\lambda^{k_2} (R_1) \langle D \rangle^{\mathtt{m}} | \, [\, |h| \,] \, \|_s^2 \\
& \overset{(2.70),(2.28)}{\lesssim_{k_0}} N^{2\tau_1} \gamma^{-2(1+|k|)} \Big(\mathfrak{M}^\sharp_{\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} R_1 \langle D \rangle^{\mathtt{m}}}(s) \|h\|_{s_0} + \mathfrak{M}^\sharp_{\langle \partial_{\varphi,x} \rangle^{\mathtt{b}} \langle D \rangle^{\mathtt{m}} R_1 \langle D \rangle^{\mathtt{m}}}(s_0) \|h\|_s \Big)^2
\end{aligned}
\tag{14.62}
$$

and, recalling Definition 2.29, inequality (14.52) follows. The proof of (14.54)-(14.55) follow similarly. $\square$

If Ψ, with Ψ_1, Ψ_2 defined in (14.48)-(14.49), satisfies the smallness condition

$$4C(\mathtt{b})C(k_0)\mathfrak{M}^\sharp_\Psi(s_0) \le 1/2 \,, \tag{14.63}$$

then, by Lemma 2.32, Φ is invertible, and (14.44), (14.45) imply that, for all $\lambda \in \Lambda^\gamma_{\mathtt{n}+1}$,

$$\mathcal{L}_+ = \Phi^{-1}\mathcal{L}\Phi = \omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i}\mathcal{D}_+ + \mathcal{R}_+ \tag{14.64}$$

which proves (14.28) and (14.19) at the step $\mathtt{n}+1$, with

$$\mathrm{i}\mathcal{D}_+ := \mathrm{i}\mathcal{D} + [\mathcal{R}] \,, \quad \mathcal{R}_+ := \Phi^{-1}\big(\Pi_N^\perp \mathcal{R} + \mathcal{R}\Psi - \Psi[\mathcal{R}]\big) \,. \tag{14.65}$$

We note that $\mathcal{R}_+$ satisfies

$$\mathcal{R}_+ = \begin{pmatrix} (R_+)_1 & (R_+)_2 \\ (\overline{R}_+)_2 & (\overline{R}_+)_1 \end{pmatrix}, \quad [(R_+)_1]_j^{j'}(\ell) = [(R_+)_2]_j^{j'}(\ell) = 0 \quad \forall (\ell, j, j'), \ jj' < 0 \,, \tag{14.66}$$

similarly as $\mathcal{R}_{\mathtt{n}}$ in (14.24), 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 (14.24), (14.50), (14.46)), and therefore, by Neumann series, also Φ^{-1} does.

The right hand sides of (14.64)-(14.65) define an extension of $\mathcal{L}_+$ to the whole parameter space $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, since $\mathcal{R}$ and Ψ are defined on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$.

The new operator $\mathcal{L}_+$ in (14.64) has the same form as $\mathcal{L}$ in (14.19), with the non-diagonal remainder $\mathcal{R}_+$ defined in (14.65) 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 (14.65) is diagonal:

Lemma 14.6. (New diagonal part). *For all $(\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$ we have*

$$\mathrm{i}\mathcal{D}_+ = \mathrm{i}\mathcal{D} + [\mathcal{R}] = \mathrm{i}\begin{pmatrix} D_+ & 0 \\ 0 & -D_+ \end{pmatrix}, \quad D_+ := \mathrm{diag}_{j \in \mathbb{S}_0^c} \mu_j^+ \,, \quad \mu_j^+ := \mu_j + \mathtt{r}_j \in \mathbb{R} \,, \tag{14.67}$$

with $\mathtt{r}_j = \mathtt{r}_{-j}$, $\mu_j^+ = \mu_{-j}^+$ for all $j \in \mathbb{S}_0^c$, and, on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$,

$$|\mathtt{r}_j|^{k_0,\gamma} = |\mu_j^+ - \mu_j|^{k_0,\gamma} \lesssim |j|^{-2\mathtt{m}} \mathfrak{M}^\sharp_{\langle D \rangle^{\mathtt{m}} \mathcal{R} \langle D \rangle^{\mathtt{m}}}(s_0). \tag{14.68}$$

Moreover, given tori $i_1(\omega, \mathtt{h})$, $i_2(\omega, \mathtt{h})$, the difference

$$|\mathtt{r}_j(i_1) - \mathtt{r}_j(i_2)| \lesssim |j|^{-2\mathtt{m}} \| |\langle D \rangle^{\mathtt{m}} \Delta_{12} \mathcal{R} \langle D \rangle^{\mathtt{m}} | \|_{\mathcal{L}(H^{s_0})} \,. \tag{14.69}$$

Proof. Identity (14.67) follows by (14.19) and (14.46) with $\mathtt{r}_j := -\mathrm{i}(R_1)_j^j(0)$. Since R_1 satisfies (14.24) and it is even, we deduce, by (2.58), that $\mathtt{r}_{-j} = \mathtt{r}_j$. Since $\mathcal{R}$ is reversible, (2.63) 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 2.29, 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 (2.67))

$$|\partial_\lambda^k (R_1)_j^j(0)| \lesssim |j|^{-2\mathtt{m}}\gamma^{-|k|}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}} R_1 \langle D\rangle^{\mathtt{m}}}(s_0) \lesssim |j|^{-2\mathtt{m}}\gamma^{-|k|}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}} \mathcal{R} \langle D\rangle^{\mathtt{m}}}(s_0)$$

which implies (14.68). Estimate (14.69) follows by $|\Delta_{12}(R_1)_j^j(0)| \lesssim |j|^{-2\mathtt{m}}|||\langle D\rangle^{\mathtt{m}}\Delta_{12}\mathcal{R}\langle D\rangle^{\mathtt{m}}|||_{\mathcal{L}(H^{s_0})}$. $\square$

14.1.2 Reducibility iteration

Let $n \geq 0$ and suppose that $(\mathbf{S1})_{\mathtt{n}}$-$(\mathbf{S3})_{\mathtt{n}}$ are true for all $\mathtt{n} = 0, \ldots, n$. We prove $(\mathbf{S1})_{n+1}$-$(\mathbf{S3})_{n+1}$. For simplicity of notation we omit to write the dependence on k_0 which is considered as a fixed constant.

Proof of $(\mathbf{S1})_{n+1}$. By (14.51)-(14.53), (14.25), 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 14.5 satisfies estimates (14.29)-(14.31) with $\mathtt{n} = n+1$. In particular at $s = s_0$ we have

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\pm\mathtt{m}}\Psi_n\langle D\rangle^{\mp\mathtt{m}}}(s_0)\,,\ \mathfrak{M}^\sharp_{\Psi_n}(s_0) \leq C(s_0,\mathtt{b})N_n^{\tau_1}N_{n-1}^{-\mathtt{a}}\gamma^{-1}\mathfrak{M}_0(s_0,\mathtt{b})\,. \tag{14.70}$$

Therefore, by (14.70), (14.7), (14.18), choosing $\tau_2 > \tau_1$, the smallness condition (14.63) 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

$$\Phi_n^{-1} = \mathbb{I}_\perp + \check{\Psi}_n\,, \qquad \check{\Psi}_n := \begin{pmatrix} \check{\Psi}_{n,1} & \check{\Psi}_{n,2} \\ \overline{\check{\Psi}}_{n,2} & \overline{\check{\Psi}}_{n,1} \end{pmatrix}. \tag{14.71}$$

Moreover also the smallness condition (2.75) (of Corollary 2.33) with $A = \Psi_n$, holds, and Lemma 2.32, Corollary 2.33 and Lemma 14.5 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

$$\mathfrak{M}^\sharp_{\check{\Psi}_n}(s),\ \mathfrak{M}^\sharp_{\langle D\rangle^{\pm\mathtt{m}}\check{\Psi}_n\langle D\rangle^{\mp\mathtt{m}}}(s) \lesssim_{s_0,\mathtt{b}} N_n^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s) \tag{14.72}$$

$$\overset{(14.25)_{|n}}{\lesssim_{s_0,\mathtt{b}}} N_n^{\tau_1}N_{n-1}^{-\mathtt{a}}\gamma^{-1}\mathfrak{M}_0(s,\mathtt{b})\,, \tag{14.73}$$

and

$$\begin{aligned}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\pm\mathtt{m}}\check{\Psi}_n\langle D\rangle^{\mp\mathtt{m}}}(s) &\lesssim_{s_0,\mathtt{b}} N_n^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s) \\ &\quad + N_n^{2\tau_1}\gamma^{-2}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s_0)\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s) \qquad (14.74)\\ &\overset{(14.25)_{|n},(14.7),(14.18)}{\lesssim_{s_0,\mathtt{b}}} N_n^{\tau_1}N_{n-1}\gamma^{-1}\mathfrak{M}_0(s,\mathtt{b})\,. \qquad (14.75)\end{aligned}$$

Conjugating $\mathcal{L}_n$ by Φ_n, we obtain, by (14.64)-(14.65), for all $\lambda \in \Lambda_{n+1}^\gamma$,

$$\mathcal{L}_{n+1} = \Phi_n^{-1}\mathcal{L}_n\Phi_n = \omega\cdot\partial_\varphi\mathbb{I}_\perp + \mathrm{i}\mathcal{D}_{n+1} + \mathcal{R}_{n+1}\,, \tag{14.76}$$

namely (14.28) at $\mathtt{n} = n+1$, where

$$\mathrm{i}\mathcal{D}_{n+1} := \mathrm{i}\mathcal{D}_n + [\mathcal{R}_n]\,, \qquad \mathcal{R}_{n+1} := \Phi_n^{-1}\big(\Pi_{N_n}^\perp\mathcal{R}_n + \mathcal{R}_n\Psi_n - \Psi_n[\mathcal{R}_n]\big)\,. \tag{14.77}$$

The operator $\mathcal{L}_{n+1}$ is real, even and reversible because Φ_n is real, even and reversibility preserving (Lemma 14.5) and $\mathcal{L}_n$ is real, even and reversible. Note that the operators $\mathcal{D}_{n+1}, \mathcal{R}_{n+1}$ are defined on $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$, and the identity (14.76) holds on Λ_{n+1}^γ.

By Lemma 14.6 the operator $\mathcal{D}_{n+1}$ is diagonal and, by (14.15), (14.25), (14.14), its eigenvalues $\mu_j^{n+1} : \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] \to \mathbb{R}$ 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 (14.22) with $\mathtt{n} = n+1$. Thus also (14.21) at $\mathtt{n} = n+1$ holds, by a telescoping sum. In addition, by (14.66) the operator $\mathcal{R}_{n+1}$ satisfies (14.24) with $\mathtt{n} = n+1$. In order to prove that (14.25) holds with $\mathtt{n} = n+1$, we first provide the following inductive estimates on the new remainder $\mathcal{R}_{n+1}$.

Lemma 14.7. *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*

$$\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) \lesssim_{s_0,\mathtt{b}} N_n^{-\mathtt{b}}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s) + \frac{N_n^{\tau_1}}{\gamma}\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s)\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s_0)\,, \tag{14.78}$$

$$\begin{aligned}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) &\lesssim_{s_0,\mathtt{b}} \mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s)\\ &\quad + N_n^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s_0)\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}}}(s)\,.\end{aligned} \tag{14.79}$$

Proof. By (14.77) and (14.71), we write

$$\begin{aligned}\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}} &= \langle D\rangle^{\mathtt{m}}\Pi^\perp_{N_n}\mathcal{R}_n\langle D\rangle^{\mathtt{m}} + (\langle D\rangle^{\mathtt{m}}\check{\Psi}_n\langle D\rangle^{-\mathtt{m}})(\langle D\rangle^{\mathtt{m}}\Pi^\perp_{N_n}\mathcal{R}_n\langle D\rangle^{\mathtt{m}})\\ &\quad + \Big(\mathbb{I}_\perp + \langle D\rangle^{\mathtt{m}}\check{\Psi}_n\langle D\rangle^{-\mathtt{m}}\Big)\Big((\langle D\rangle^{\mathtt{m}}\mathcal{R}_n\langle D\rangle^{\mathtt{m}})(\langle D\rangle^{-\mathtt{m}}\Psi_n\langle D\rangle^{\mathtt{m}})\Big)\\ &\quad - \Big(\mathbb{I}_\perp + \langle D\rangle^{\mathtt{m}}\check{\Psi}_n\langle D\rangle^{-\mathtt{m}}\Big)\Big((\langle D\rangle^{\mathtt{m}}\Psi_n\langle D\rangle^{-\mathtt{m}})(\langle D\rangle^{\mathtt{m}}[\mathcal{R}_n]\langle D\rangle^{\mathtt{m}})\Big)\,.\end{aligned} \tag{14.80}$$

The proof of (14.78) follows by estimating separately all the terms in (14.80), applying Lemmata 2.34, 2.31, and (14.51), (14.72), $(14.25)_{|n}$, (14.7), (14.18). The proof of (14.79) follows by formula (14.80), Lemmata 2.31, 2.34 and estimates (14.51), (14.52), (14.72), $(14.25)_{|n}$, (14.7), (14.18). □

In the next lemma we prove that (14.25) holds at $\mathtt{n} = n+1$, concluding the proof of $(\mathbf{S1})_{n+1}$.

Lemma 14.8. *For $N_0 = N_0(S,\mathtt{b}) > 0$ large enough we have*

$$\begin{aligned}&\mathfrak{M}^\sharp_{\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) \leq C_*(s_0,\mathtt{b})N_n^{-\mathtt{a}}\mathfrak{M}_0(s,\mathtt{b})\\ &\mathfrak{M}^\sharp_{\langle\partial_{\varphi,x}\rangle^{\mathtt{b}}\langle D\rangle^{\mathtt{m}}\mathcal{R}_{n+1}\langle D\rangle^{\mathtt{m}}}(s) \leq C_*(s_0,\mathtt{b})N_n\mathfrak{M}_0(s,\mathtt{b})\,.\end{aligned}$$

Proof. By (14.78) and (14.25) 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 (14.7), (14.18), taking $N_0(S,\mathtt{b}) > 0$ large enough and $\tau_2 > \tau_1 + \mathtt{a}$. Then by (14.79), (14.25) 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 (14.7), (14.18) and taking $N_0(S,\mathtt{b}) > 0$ large enough. □

Proof of $(\mathbf{S2})_{n+1}$. The proof of the estimates (14.32), (14.33) for $\mathtt{n} = n+1$ for the term $\Delta_{12}\mathcal{R}_{n+1}$ (where $\mathcal{R}_{n+1}$ is defined in (14.77)) follow as above. The proof of (14.34) for $\mathtt{n} = n+1$ follows estimating $\Delta_{12}(r_j^{n+1} - r_j^n) = \Delta_{12}\mathtt{r}_j^n$ by (14.69) of Lemma 14.6 and by (14.32) for $\mathtt{n} = n$. Estimate (14.35) for $\mathtt{n} = n+1$ follows by a telescoping argument using (14.34) and (14.32).

Proof of $(\mathbf{S3})_{n+1}$. First we note that the non-resonance conditions imposed in (14.26) 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 $\mathtt{n} = 0,\ldots,n$, for all $|\ell|, |j-j'| \leq N_{\mathtt{n}}$, $j,j' \in \mathbb{N}^+ \setminus \mathbb{S}^+$, if
$$\min\{j,j'\} \geq C_0 N_{\mathtt{n}}^{2(\tau+1)}\gamma^{-2}, \tag{14.81}$$
then $|\omega\cdot\ell + \mu_j^{\mathtt{n}} - \mu_{j'}^{\mathtt{n}}| \geq \gamma\langle\ell\rangle^{-\tau}$.

PROOF OF THE CLAIM. By (14.20), (14.21) and recalling also (12.78), one has

$$\mu_j^{\mathtt{n}} = \mathtt{m}_{\frac{1}{2}} j^{\frac{1}{2}} \tanh^{\frac{1}{2}}(\mathtt{h}j) + \mathfrak{r}_j^{\mathtt{n}}, \quad \mathfrak{r}_j^{\mathtt{n}} := r_j + r_j^{\mathtt{n}}, \quad \sup_{j \in \mathbb{S}^c} j^{\frac{1}{2}} |\mathfrak{r}_j^{\mathtt{n}}|^{k_0,\gamma} \lesssim_S \varepsilon \gamma^{-2(M+1)} . \tag{14.82}$$

For all $j, j' \in \mathbb{N} \setminus \{0\}$, one has

$$|\sqrt{j \tanh(\mathtt{h}j)} - \sqrt{j' \tanh(\mathtt{h}j')}| \leq \frac{C(\mathtt{h})}{\min\{\sqrt{j}, \sqrt{j'}\}} |j - j'|. \tag{14.83}$$

Then, using (14.83) and that $\omega \in \mathtt{DC}(2\gamma, \tau)$, we have, for $|j - j'| \leq N_{\mathtt{n}}$, $|\ell| \leq N_{\mathtt{n}}$,

$$\begin{aligned} |\omega \cdot \ell + \mu_j^{\mathtt{n}} - \mu_{j'}^{\mathtt{n}}| &\geq |\omega \cdot \ell| - |\mathtt{m}_{\frac{1}{2}}| \frac{C(\mathtt{h})}{\min\{\sqrt{j}, \sqrt{j'}\}} |j - j'| - |\mathfrak{r}_j^{\mathtt{n}}| - |\mathfrak{r}_{j'}^{\mathtt{n}}| \\ &\overset{(11.24),(14.82)}{\geq} \frac{2\gamma}{\langle \ell \rangle^\tau} - \frac{2C(\mathtt{h}) N_{\mathtt{n}}}{\min\{\sqrt{j}, \sqrt{j'}\}} - \frac{C(S)\varepsilon\gamma^{-2(M+1)}}{\min\{\sqrt{j}, \sqrt{j'}\}} \overset{(14.81)}{\geq} \frac{\gamma}{\langle \ell \rangle^\tau} , \end{aligned}$$

where the last inequality holds for C_0 large enough. This proves the claim.

Now we prove $(\mathbf{S3})_{n+1}$, namely that

$$C(S) N_n^{(\tau+1)(4\mathtt{d}+1)} \gamma^{-4\mathtt{d}} \|i_2 - i_1\|_{s_0+\mu(\mathtt{b})} \leq \rho \quad \Longrightarrow \quad \Lambda_{n+1}^\gamma(i_1) \subseteq \Lambda_{n+1}^{\gamma-\rho}(i_2) . \tag{14.84}$$

Let $\lambda \in \Lambda_{n+1}^\gamma(i_1)$. Definition (14.26) and (14.36) with $\mathtt{n} = n$ (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 (14.35) 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 (14.81) with $\mathtt{n} = n$ 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 (14.26) with $\mathtt{n} = n + 1$ also hold for j, j' violating $(14.81)|_{\mathtt{n}=n}$, namely that

$$|\omega \cdot \ell + \mu_j^n(\lambda, i_2(\lambda)) - \mu_{j'}^n(\lambda, i_2(\lambda))| \geq \frac{\gamma - \rho}{\langle \ell \rangle^\tau j^{\mathtt{d}} j'^{\mathtt{d}}} , \ \forall |\ell|, |j - j'| \leq N_n , \ \min\{j, j'\} \leq C_0 N_n^{2(\tau+1)} \gamma^{-2} . \tag{14.85}$$

The conditions on j, j' in (14.85) imply that

$$\max\{j, j'\} = \min\{j, j'\} + |j - j'| \leq C_0 N_n^{2(\tau+1)} \gamma^{-2} + N_n \leq 2C_0 N_n^{2(\tau+1)} \gamma^{-2} . \tag{14.86}$$

Now by (14.20), (14.21), (14.83), recalling (11.24), (12.78), (14.35) and the bound $\varepsilon \gamma^{-2(M+1)} \leq 1$, we get

$$\begin{aligned} |(\mu_j^n - \mu_{j'}^n)(\lambda, i_2(\lambda)) - (\mu_j^n - \mu_{j'}^n)(\lambda, i_1(\lambda))| &\leq |(\mu_j^0 - \mu_{j'}^0)(\lambda, i_2(\lambda)) - (\mu_j^0 - \mu_{j'}^0)(\lambda, i_1(\lambda))| \\ &\quad + |r_j^n(\lambda, i_2(\lambda)) - r_j^n(\lambda, i_1(\lambda))| + |r_{j'}^n(\lambda, i_2(\lambda)) - r_{j'}^n(\lambda, i_1(\lambda))| \\ &\leq \frac{C(S) N_n}{\min\{\sqrt{j}, \sqrt{j'}\}} \|i_2 - i_1\|_{s_0+\mu(\mathtt{b})} . \end{aligned} \tag{14.87}$$

Since $\lambda \in \Lambda_{n+1}^\gamma(i_1)$, by (14.87) we have, for all $|\ell| \leq N_n$, $|j - j'| \leq N_n$,

$$\begin{aligned} |\omega \cdot \ell + \mu_j^n(i_2) - \mu_{j'}^n(i_2)| &\geq |\omega \cdot \ell + \mu_j^n(i_1) - \mu_{j'}^n(i_1)| - |(\mu_j^n - \mu_{j'}^n)(i_2) - (\mu_j^n - \mu_{j'}^n)(i_1)| \\ &\geq \frac{\gamma}{\langle \ell \rangle^\tau j^{\mathtt{d}} j'^{\mathtt{d}}} - \frac{C(S) N_n}{\min\{\sqrt{j}, \sqrt{j'}\}} \|i_2 - i_1\|_{s_0+\mu(\mathtt{b})} \\ &\geq \frac{\gamma}{\langle \ell \rangle^\tau j^{\mathtt{d}} j'^{\mathtt{d}}} - C(S) N_n \|i_2 - i_1\|_{s_0+\mu(\mathtt{b})} \geq \frac{\gamma - \rho}{\langle \ell \rangle^\tau j^{\mathtt{d}} j'^{\mathtt{d}}} \end{aligned}$$

provided $C(S) N_n \langle \ell \rangle^\tau j^{\mathtt{d}} j'^{\mathtt{d}} \|i_2 - i_1\|_{s_0+\mu(\mathtt{b})} \leq \rho$. Using that $|\ell| \leq N_n$ and (14.86), the above inequality is implied by the inequality assumed in (14.84). The proof for the second Melnikov conditions for $\omega \cdot \ell + \mu_j^n + \mu_{j'}^n$ can be carried out similarly (in fact, it is simpler). This completes the proof of (14.36) with $\mathtt{n} = n + 1$. $\square$

14.2 Almost-invertibility of $\mathcal{L}_\omega$

By (13.6), $\mathcal{L}_\omega = \mathcal{P}_\perp \mathcal{L}_\perp \mathcal{P}_\perp^{-1}$, where $\mathcal{P}_\perp$ is defined in (13.2), (13.3). By (14.42), for any $\lambda \in \mathtt{\Lambda}_n^\gamma$, we have that $\mathcal{L}_0 = \mathcal{U}_n \mathcal{L}_n \mathcal{U}_n^{-1}$, where $\mathcal{U}_n$ is defined in (14.37), $\mathcal{L}_0 = \mathcal{L}_\perp^{sym}$, and $\mathcal{L}_\perp^{sym} = \mathcal{L}_\perp$ on the subspace of functions even in x (see (14.3)). Thus

$$\mathcal{L}_\omega = \mathcal{V}_n \mathcal{L}_n \mathcal{V}_n^{-1}, \quad \mathcal{V}_n := \mathcal{P}_\perp \mathcal{U}_n. \tag{14.88}$$

By Lemmata 2.27, 2.30, by estimate (14.39), using the smallness condition (14.38) and $\tau_2 > \tau_1$ (see Theorem 14.3), the operators $\mathcal{U}_n^{\pm 1}$ satisfy, for all $s_0 \leq s \leq S$, $\|\mathcal{U}_n^{\pm 1} h\|_s^{k_0,\gamma} \lesssim_S \|h\|_s^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})}^{k_0,\gamma} \|h\|_{s_0}^{k_0,\gamma}$. Therefore, by definition (14.88) and recalling (13.4), (14.8), (14.9), the operators $\mathcal{V}_n^{\pm 1}$ satisfy, for all $s_0 \leq s \leq S$,

$$\|\mathcal{V}_n^{\pm 1} h\|_s^{k_0,\gamma} \lesssim_S \|h\|_{s+\sigma}^{k_0,\gamma} + \|\mathfrak{I}_0\|_{s+\mu(\mathtt{b})}^{k_0,\gamma} \|h\|_{s_0+\sigma}^{k_0,\gamma}, \tag{14.89}$$

for some $\sigma = \sigma(k_0, \tau, \nu) > 0$.

In order to verify the inversion assumption (5.29)-(5.33) we decompose the operator $\mathcal{L}_n$ in (14.42) as

$$\mathcal{L}_n = \mathfrak{L}_n^< + \mathcal{R}_n + \mathcal{R}_n^\perp \tag{14.90}$$

where

$$\mathfrak{L}_n^< := \Pi_{K_n} \big(\omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i} \mathcal{D}_n \big) \Pi_{K_n} + \Pi_{K_n}^\perp, \quad \mathcal{R}_n^\perp := \Pi_{K_n}^\perp \big(\omega \cdot \partial_\varphi \mathbb{I}_\perp + \mathrm{i} \mathcal{D}_n \big) \Pi_{K_n}^\perp - \Pi_{K_n}^\perp, \tag{14.91}$$

the diagonal operator $\mathcal{D}_n$ is defined in (14.19) (with $\mathtt{n} = n$), and $K_n := K_0^{\chi^n}$ is the scale of the nonlinear Nash-Moser iterative scheme.

Lemma 14.9. (First order Melnikov non-resonance conditions) *For all $\lambda = (\omega, \mathtt{h})$ in*

$$\mathtt{\Lambda}_{n+1}^{\gamma,I} := \mathtt{\Lambda}_{n+1}^{\gamma,I}(i) := \big\{ \lambda \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] : |\omega \cdot \ell + \mu_j^n| \geq 2\gamma j^{\frac12} \langle \ell \rangle^{-\tau}, \quad \forall |\ell| \leq K_n, \ j \in \mathbb{N}^+ \setminus \mathbb{S}^+ \big\}, \tag{14.92}$$

the operator $\mathfrak{L}_n^<$ in (14.91) *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*

$$\|(\mathfrak{L}_n^<)^{-1} g\|_s^{k_0,\gamma} \lesssim_{k_0} \gamma^{-1} \|g\|_{s+\mu}^{k_0,\gamma}, \tag{14.93}$$

where $\mu = k_0 + \tau(k_0+1)$ is the constant in (2.18) *with $k_0 = k+1$.*

Proof. By (14.56), similarly as in (14.57) one has $\gamma^{|\alpha|} |\partial_\lambda^\alpha (\omega \cdot \ell + \mu_j^n)| \lesssim \gamma \langle \ell \rangle |j|^{\frac12}$ for all $1 \leq |\alpha| \leq k_0$. Hence Lemma B.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 (14.93). $\square$

Standard smoothing properties imply that the operator $\mathcal{R}_n^\perp$ defined in (14.91) satisfies, for all $b > 0$,

$$\|\mathcal{R}_n^\perp h\|_{s_0}^{k_0,\gamma} \lesssim K_n^{-b} \|h\|_{s_0+b+1}^{k_0,\gamma}, \quad \|\mathcal{R}_n^\perp h\|_s^{k_0,\gamma} \lesssim \|h\|_{s+1}^{k_0,\gamma}. \tag{14.94}$$

By (14.88), (14.90), Theorem 14.4, Proposition 13.3, and estimates (14.93), (14.94), (14.89), we deduce the following theorem.

Theorem 14.10. (Almost-invertibility of $\mathcal{L}_\omega$) *Assume* (5.6). *Let $\mathtt{a}, \mathtt{b}$ as in* (14.7) *and M as in* (14.8). *Let $S > s_0$, and assume the smallness condition* (14.38). *Then for all*

$$(\omega, \mathtt{h}) \in \mathbf{\Lambda}_{n+1}^\gamma := \mathbf{\Lambda}_{n+1}^\gamma(i) := \mathtt{\Lambda}_{n+1}^\gamma \cap \mathtt{\Lambda}_{n+1}^{\gamma,I} \tag{14.95}$$

(see (14.41), (14.92)*) the operator $\mathcal{L}_\omega$ defined in* (5.26) *(see also* (6.5)*) can be decomposed as (cf.* (5.29)*)*

$$\mathcal{L}_\omega = \mathcal{L}_\omega^< + \mathcal{R}_\omega + \mathcal{R}_\omega^\perp, \quad \mathcal{L}_\omega^< := \mathcal{V}_n \mathfrak{L}_n^< \mathcal{V}_n^{-1}, \quad \mathcal{R}_\omega := \mathcal{V}_n \mathcal{R}_n \mathcal{V}_n^{-1}, \quad \mathcal{R}_\omega^\perp := \mathcal{V}_n \mathcal{R}_n^\perp \mathcal{V}_n^{-1} \tag{14.96}$$

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 $s_0 \leq s \leq S$, estimates (5.30)-(5.33)*, with $\mu(\mathtt{b})$ defined in* (14.9). *Notice that these latter estimates hold on the whole $\mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$.*

This result allows to deduce Theorem 5.6, which is the key step for a Nash-Moser iterative scheme.

15 Proof of Theorem 4.1

We consider the finite-dimensional subspaces

$$E_n := \Big\{ \mathfrak{I}(\varphi) = (\Theta, I, z)(\varphi), \quad \Theta = \Pi_n \Theta, \ I = \Pi_n I, \ z = \Pi_n z \Big\}$$

where Π_n is the projector

$$\Pi_n := \Pi_{K_n} : z(\varphi, x) = \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{S}_0^c} z_{\ell,j} e^{\mathrm{i}(\ell \cdot \varphi + jx)} \ \mapsto \Pi_n z(\varphi, x) := \sum_{|(\ell,j)| \le K_n} z_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + jx)} \tag{15.1}$$

with $K_n = K_0^{\chi^n}$ (see (5.28)) and we denote with the same symbol $\Pi_n p(\varphi) := \sum_{|\ell| \le K_n} p_\ell e^{\mathrm{i}\ell\cdot\varphi}$. 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 15.1 we introduce the following constants:

$$\mathtt{a}_1 := \max\{6\sigma_1 + 13, \chi p(\tau+1)(4\mathtt{d}+1) + \chi(\mu(\mathtt{b}) + 2\sigma_1) + 1\}, \qquad \mathtt{a}_2 := \chi^{-1}\mathtt{a}_1 - \mu(\mathtt{b}) - 2\sigma_1, \tag{15.2}$$

$$\mu_1 := 3(\mu(\mathtt{b}) + 2\sigma_1) + 1, \qquad \mathtt{b}_1 := \mathtt{a}_1 + \mu(\mathtt{b}) + 3\sigma_1 + 3 + \chi^{-1}\mu_1, \qquad \chi = 3/2, \tag{15.3}$$

$$\sigma_1 := \max\{\bar\sigma\,, s_0 + 2k_0 + 5\}, \qquad S := s_0 + \mathtt{b}_1 \tag{15.4}$$

where $\bar\sigma := \bar\sigma(\tau, \nu, k_0) > 0$ is defined in Theorem 5.6, $s_0 + 2k_0 + 5$ is the largest loss of regularity in the estimates of the Hamiltonian vector field X_P in Lemma 5.1, $\mu(\mathtt{b})$ is defined in (14.9), $\mathtt{b}$ is the constant $\mathtt{b} := [\mathtt{a}] + 2 \in \mathbb{N}$ where $\mathtt{a}$ is defined in (14.7). The constants $\mathtt{b}_1$, μ_1 appear in $(\mathcal{P}3)_n$ of Theorem 15.1 below: $\mathtt{b}_1$ gives the maximal Sobolev regularity $S = s_0 + \mathtt{b}_1$ which has to be controlled along the Nash Moser iteration and μ_1 gives the rate of divergence of the high norms $\|\tilde W_n\|_{s_0+\mathtt{b}_1}^{k_0,\gamma}$. The constant $\mathtt{a}_1$ appears in (15.10) and gives the rate of convergence of $\mathcal{F}(\tilde U_n)$ in low norm.

The exponent p in (5.27) which links the scale $(N_n)_{n\ge0}$ of the reducibility scheme (Theorem 14.4) and the scale $(K_n)_{n\ge0}$ of the Nash-Moser iteration ($N_n = K_n^p$) is required to satisfy

$$p\mathtt{a} > (\chi - 1)\mathtt{a}_1 + \chi\sigma_1 = \frac{1}{2}\mathtt{a}_1 + \frac{3}{2}\sigma_1\,. \tag{15.5}$$

By (14.7), $\mathtt{a} \ge \chi(\tau+1)(4\mathtt{d}+1)+1$. Hence, by the definition of $\mathtt{a}_1$ in (15.2), there exists $p := p(\tau, \nu, k_0)$ such that (15.5) holds. For example we fix $p := 3(\mu(\mathtt{b}) + 3\sigma_1 + 1)/\mathtt{a}$.

Given $W = (\mathfrak{I}, \beta)$ where $\mathfrak{I} = \mathfrak{I}(\lambda)$ is the periodic component of a torus as in (4.12), 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 (4.13).

Theorem 15.1. (Nash-Moser) *There exist δ_0, $C_* > 0$, such that, if*

$$K_0^{\tau_3}\varepsilon\gamma^{-2M-3} < \delta_0, \quad \tau_3 := \max\{p\tau_2, 2\sigma_1 + \mathtt{a}_1 + 4\}, \quad K_0 := \gamma^{-1}, \quad \gamma := \varepsilon^a, \quad 0 < a < \frac{1}{\tau_3 + 2M + 3}\,, \tag{15.6}$$

where the constant M is defined in (14.8) and $\tau_2 := \tau_2(\tau, \nu)$ is defined in Theorem 14.3, then, for all $n \ge 0$:

$(\mathcal{P}1)_n$ *there exists a k_0 times differentiable function $\tilde W_n : \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] \to E_{n-1} \times \mathbb{R}^\nu$, $\lambda = (\omega, \mathtt{h}) \mapsto \tilde W_n(\lambda) := (\tilde{\mathfrak{I}}_n, \tilde\alpha_n - \omega)$, for $n \ge 1$, and $\tilde W_0 := 0$, satisfying*

$$\|\tilde W_n\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \le C_*\varepsilon\gamma^{-1}\,. \tag{15.7}$$

Let $\tilde U_n := U_0 + \tilde W_n$ where $U_0 := (\varphi, 0, 0, \omega)$. The difference $\tilde H_n := \tilde U_n - \tilde U_{n-1}$, $n \ge 1$, satisfies

$$\|\tilde H_1\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \le C_*\varepsilon\gamma^{-1}\,, \quad \|\tilde H_n\|_{s_0+\mu(\mathtt{b})+\sigma_1}^{k_0,\gamma} \le C_*\varepsilon\gamma^{-1}K_{n-1}^{-\mathtt{a}_2}\,, \quad \forall n \ge 2. \tag{15.8}$$

$(\mathcal{P}2)_n$ *Setting $\tilde\imath_n := (\varphi, 0, 0) + \tilde{\mathfrak{I}}_n$, we define*

$$\mathcal{G}_0 := \Omega \times [\mathtt{h}_1, \mathtt{h}_2]\,, \quad \mathcal{G}_{n+1} := \mathcal{G}_n \cap \mathbf{\Lambda}_{n+1}^\gamma(\tilde\imath_n)\,, \quad n \ge 0\,, \tag{15.9}$$

where $\mathbf{\Lambda}_{n+1}^\gamma(\tilde\imath_n)$ is defined in (14.95). Then, for all $\lambda \in \mathcal{G}_n$, setting $K_{-1} := 1$, we have

$$\|\mathcal{F}(\tilde U_n)\|_{s_0}^{k_0,\gamma} \le C_*\varepsilon K_{n-1}^{-\mathtt{a}_1}\,. \tag{15.10}$$

$(\mathcal{P}3)_n$ (High norms). $\|\tilde W_n\|^{k_0,\gamma}_{s_0+\mathtt{b}_1} \le C_* \varepsilon \gamma^{-1} K_{n-1}^{\mu_1}$ *for all* $\lambda \in \mathcal{G}_n$.

Proof. The proof is the same as Theorem 8.2 in [21]. It is based on an iterative Nash-Moser scheme and uses the almost-approximate inverse at each approximate quasi-periodic solution provided by Theorem 5.6. $\square$

We now complete the proof of Theorem 4.1. Let $\gamma = \varepsilon^a$ with $a \in (0, a_0)$ and $a_0 := 1/(2M + 3 + \tau_3)$ where τ_3 is defined in (15.6). Then the smallness condition given by the first inequality in (15.6) holds for $0 < \varepsilon < \varepsilon_0$ small enough and Theorem 15.1 applies. By (15.8) the sequence of functions

$$\tilde W_n = \tilde U_n - (\varphi, 0, 0, \omega) := (\tilde{\mathfrak{I}}_n, \tilde\alpha_n - \omega) = \big(\tilde\imath_n - (\varphi, 0, 0), \tilde\alpha_n - \omega\big)$$

is a Cauchy sequence in $\| \ \|^{k_0,\gamma}_{s_0}$ and then it converges to a function $W_\infty := \lim_{n\to+\infty} \tilde W_n$. We define

$$U_\infty := (i_\infty, \alpha_\infty) = (\varphi, 0, 0, \omega) + W_\infty \,, \quad W_\infty : \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2] \to H^{s_0}_\varphi \times H^{s_0}_\varphi \times H^{s_0}_{\varphi,x} \times \mathbb{R}^\nu \,.$$

By (15.7) and (15.8) we also deduce that

$$\|U_\infty - U_0\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \le C_* \varepsilon \gamma^{-1} \,, \quad \|U_\infty - \tilde U_n\|^{k_0,\gamma}_{s_0+\mu(\mathtt{b})+\sigma_1} \le C \varepsilon \gamma^{-1} K_n^{-\mathtt{a}_2} \,, \quad n \ge 1 \,. \tag{15.11}$$

Moreover by Theorem 15.1-$(\mathcal{P}2)_n$, we deduce that $\mathcal{F}(\lambda, U_\infty(\lambda)) = 0$ for all λ belonging to

$$\bigcap_{n\ge 0} \mathcal{G}_n = \mathcal{G}_0 \cap \bigcap_{n\ge 1} \mathbf{\Lambda}^\gamma_n(\tilde\imath_{n-1}) \overset{(14.95)}{=} \mathcal{G}_0 \cap \Big[\bigcap_{n\ge1} \mathtt{\Lambda}^\gamma_n(\tilde\imath_{n-1})\Big] \cap \Big[\bigcap_{n\ge1} \mathtt{\Lambda}^{\gamma,I}_n(\tilde\imath_{n-1})\Big] \,, \tag{15.12}$$

where $\mathcal{G}_0 = \mathtt{\Omega} \times [\mathtt{h}_1, \mathtt{h}_2]$ is defined in (15.9). By the first inequality in (15.11) we deduce (4.16) and (4.17).

It remains to prove that the Cantor set $\mathcal{C}^\gamma_\infty$ in (4.20) is contained in $\bigcap_{n\ge0} \mathcal{G}_n$. We first consider the set

$$\mathcal{G}_\infty := \mathcal{G}_0 \cap \Big[\bigcap_{n\ge1} \mathtt{\Lambda}^{2\gamma}_n(i_\infty)\Big] \cap \Big[\bigcap_{n\ge1} \mathtt{\Lambda}^{2\gamma,I}_n(i_\infty)\Big] \,. \tag{15.13}$$

Lemma 15.2. $\mathcal{G}_\infty \subseteq \bigcap_{n\ge0} \mathcal{G}_n$, *where* $\mathcal{G}_n$ *is defined in* (15.9).

Proof. See Lemma 8.6 of [21]. $\square$

Then we define the "final eigenvalues"

$$\mu^\infty_j := \mu^0_j(i_\infty) + r^\infty_j \,, \quad j \in \mathbb{N}^+ \setminus \mathbb{S}^+ \,, \tag{15.14}$$

where $\mu^0_j(i_\infty)$ are defined in (14.5) (with $\mathtt{m}_{\frac12}, r_j$ depending on i_∞) and

$$r^\infty_j := \lim_{n\to+\infty} r^n_j(i_\infty) \,, \quad j \in \mathbb{N}^+ \setminus \mathbb{S}^+ \,, \tag{15.15}$$

with r^n_j given in Theorem 14.3-$(\mathbf{S1})_n$. Note that the sequence $(r^n_j(i_\infty))_{n\in\mathbb{N}}$ is a Cauchy sequence in $| \ |^{k_0,\gamma}$ by (14.22). As a consequence its limit function $r^\infty_j(\omega, \mathtt{h})$ is well defined, it is k_0 times differentiable and satisfies

$$|r^\infty_j - r^n_j(i_\infty)|^{k_0,\gamma} \le C \varepsilon \gamma^{-2(M+1)} |j|^{-2\mathtt{m}} N_{n-1}^{-\mathtt{a}} \,, \ n \ge 0 \,. \tag{15.16}$$

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 (15.4)). The latter estimate, (15.14), (14.5) and (12.78) imply (4.18)-(4.19) with $\mathfrak{r}^\infty_j := r_j + r^\infty_j$ and $\mathtt{m}^\infty_{\frac12} := \mathtt{m}_{\frac12}(i_\infty)$.

Lemma 15.3. *The final Cantor set* $\mathcal{C}^\gamma_\infty$ *in* (4.20) *satisfies* $\mathcal{C}^\gamma_\infty \subseteq \mathcal{G}_\infty$, *where* $\mathcal{G}_\infty$ *is defined in* (15.13).

Proof. By (15.13), we have to prove that $\mathcal{C}^\gamma_\infty \subseteq \Lambda^{2\gamma}_n(i_\infty)$, $\forall n \in \mathbb{N}$. We argue by induction. For $n = 0$ the inclusion is trivial, since $\Lambda^{2\gamma}_0(i_\infty) = \Omega \times [\mathtt{h}_1, \mathtt{h}_2] = \mathcal{G}_0$. Now assume that $\mathcal{C}^\gamma_\infty \subseteq \Lambda^{2\gamma}_n(i_\infty)$ for some $n \geq 0$. For all $\lambda \in \mathcal{C}^\gamma_\infty \subseteq \Lambda^{2\gamma}_n(i_\infty)$, by (14.20), (15.14), (15.16), we get

$$|(\mu^n_j - \mu^n_{j'})(i_\infty) - (\mu^\infty_j - \mu^\infty_{j'})| \leq C\varepsilon\gamma^{-2(M+1)} N^{-\mathtt{a}}_{n-1}\big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big)$$

Therefore, for any $|\ell|, |j - j'| \leq N_n$ with $(\ell, j, j') \neq (0, j, j)$ (recall (4.20)) we have

$$\begin{aligned}|\omega \cdot \ell + \mu^n_j(i_\infty) - \mu^n_{j'}(i_\infty)| &\geq |\omega \cdot \ell + \mu^\infty_j - \mu^\infty_{j'}| - C\varepsilon\gamma^{-2(M+1)} N^{-\mathtt{a}}_{n-1}\big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big)\\ &\geq 4\gamma\langle\ell\rangle^{-\tau} j^{-\mathtt{d}} j'^{-\mathtt{d}} - C\varepsilon\gamma^{-2(M+1)} N^{-\mathtt{a}}_{n-1}\big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big)\\ &\geq 2\gamma\langle\ell\rangle^{-\tau} j^{-\mathtt{d}} j'^{-\mathtt{d}}\end{aligned}$$

provided $C\varepsilon\gamma^{-2M-3} N^{-\mathtt{a}}_{n-1} N^\tau_n \big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big) j^{\mathtt{d}} j'^{\mathtt{d}} \leq 1$. Since $\mathtt{m} > \mathtt{d}$ (see (14.7)), one has $(j + N_n)^{\mathtt{d}} j^{\mathtt{d}-2\mathtt{m}} \lesssim_{\mathtt{d}} N^{\mathtt{d}}_n$ for all $j \geq 1$. Hence, using $|j - j'| \leq N_n$,

$$\big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big) j^{\mathtt{d}} j'^{\mathtt{d}} = \frac{j'^{\mathtt{d}}}{j^{2\mathtt{m}-\mathtt{d}}} + \frac{j^{\mathtt{d}}}{j'^{2\mathtt{m}-\mathtt{d}}} \leq \frac{(j + N_n)^{\mathtt{d}}}{j^{2\mathtt{m}-\mathtt{d}}} + \frac{(j' + N_n)^{\mathtt{d}}}{j'^{2\mathtt{m}-\mathtt{d}}} \lesssim_{\mathtt{d}} N^{\mathtt{d}}_n.$$

Therefore, for some $C_1 > 0$, one has, for any $n \geq 0$,

$$C\varepsilon\gamma^{-2M-3} N^{-\mathtt{a}}_{n-1} N^\tau_n \big(j^{-2\mathtt{m}} + j'^{-2\mathtt{m}}\big) j^{\mathtt{d}} j'^{\mathtt{d}} \leq C_1\varepsilon\gamma^{-2M-3} N^{-\mathtt{a}}_{n-1} N^{\tau+\mathtt{d}}_n \leq 1$$

for ε small enough, by (14.7), (15.6) 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 14.3). In conclusion $\mathcal{C}^\gamma_\infty \subseteq \Lambda^{2\gamma}_{n+1}(i_\infty)$ (for the second Melnikov conditions with the + sign in (14.26) we apply the same argument). Similarly we prove that $\mathcal{C}^\gamma_\infty \subseteq \Lambda^{2\gamma, I}_n(i_\infty)$ for all $n \in \mathbb{N}$. □

Lemmata 15.2, 15.3 imply $\mathcal{C}^\gamma_\infty \subseteq \bigcap_{n\geq 0} \mathcal{G}_n$, where $\mathcal{G}_n$ is defined in (15.9). This concludes the proof of Theorem 4.1.

A Dirichlet-Neumann operator

Let $\eta \in \mathcal{C}^\infty(\mathbb{T})$. It is well-known (see e.g. [47], [5], [40]) that the Dirichlet-Neumann operator is a *pseudo-differential* operator of the form

$$G(\eta) = G(0) + \mathcal{R}_G(\eta), \qquad \text{where} \quad G(0) = |D| \tanh(\mathtt{h}|D|) \tag{A.1}$$

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 $\lambda := (\omega, \mathtt{h}) \in \mathbb{R}^\nu \times [\mathtt{h}_1, \mathtt{h}_2]$. For simplicity of notation we sometimes omit to write the dependence with respect to φ and λ.

In the sequel we use the following notation. Let X and Y be Banach spaces and $B \subset X$ be a bounded open set. We denote by $\mathcal{C}^1_b(B, Y)$ the space of the $\mathcal{C}^1$ functions $B \to Y$ bounded and with bounded derivatives.

Proposition A.1. (Dirichlet-Neumann) *Assume that $\partial^k_\lambda \eta(\lambda, \cdot, \cdot)$ is $\mathcal{C}^\infty$ for all $|k| \leq k_0$. There exists $\delta(s_0, k_0) > 0$ such that, if*

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

then the Dirichlet-Neumann operator $G(\eta)$ may be written as in (A.1) *where $\mathcal{R}_G(\eta)$ is an integral operator with $\mathcal{C}^\infty$ kernel K_G (see* (2.54)*) which satisfies, for all $m, s, \alpha \in \mathbb{N}$, the estimate*

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

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 $\mathcal{C}^1_b$.

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

$$\big\{(X,Y) : -\mathtt{h} - c \le Y \le 0\big\}\,, \tag{A.4}$$

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

$$x = U(X,Y) = X + p(X,Y), \quad y = V(X,Y) = Y + q(X,Y)\,. \tag{A.5}$$

Remark A.2. If (A.5) is a conformal map then the system obtained transforming (1.3) is simply (A.32) (the Laplace operator and the Neumann boundary conditions are transformed into themselves). ☐

We require that $q(X,Y)$ and $p(X,Y)$ are 2π-periodic in X, so that (A.5) 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

$$V(X, -\mathtt{h} - c) = -\mathtt{h} \qquad \Leftrightarrow \qquad q(X, -\mathtt{h} - c) = c\,, \quad \forall X \in \mathbb{R}\,, \tag{A.6}$$

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

$$V(X,0) = \eta(U(X,0)) \qquad \Leftrightarrow \qquad q(X,0) = \eta(X + p(X,0))\,. \tag{A.7}$$

The diffeomorphism (A.5) 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 $U_X = V_Y$, $U_Y = -V_X$, namely $p_X = q_Y$, $p_Y = -q_X$. The functions (U,V), i.e. (p,q), are harmonic conjugate. Moreover, (A.6) and the Cauchy-Riemann equations imply that

$$U_Y(X, -\mathtt{h} - c) = p_Y(X, -\mathtt{h} - c) = 0\,. \tag{A.8}$$

Given any periodic function

$$\mathtt{p}(X) = \mathtt{p}_0 + \sum_{k \neq 0} \mathtt{p}_k e^{\mathrm{i}kX}\,, \tag{A.9}$$

the unique function $p(X,Y)$ that is 2π-periodic in X and solves $\Delta p = 0$, $p(X,0) = \mathtt{p}(X)$, $p_Y(X, -\mathtt{h} - c) = 0$ is

$$p(X,Y) = \sum_{k \in \mathbb{Z}} \mathtt{p}_k \frac{\cosh(|k|(Y + \mathtt{h} + c))}{\cosh(|k|(\mathtt{h} + c))}\, e^{\mathrm{i}kX}. \tag{A.10}$$

The unique function $q(X,Y)$ that is 2π-periodic in X and solves $\Delta q = 0$, (A.6) and $p_X = q_Y$, $p_Y = -q_X$ is

$$q(X,Y) = c + \sum_{k \neq 0} \mathrm{i}\mathtt{p}_k \frac{\mathrm{sign}(k)}{\cosh(|k|(\mathtt{h} + c))} \sinh(|k|(Y + \mathtt{h} + c)) e^{\mathrm{i}kX}\,. \tag{A.11}$$

We still have to impose (A.7). By (A.11) we have

$$q(X,0) = c + \sum_{k \neq 0} \mathrm{i}\,\mathrm{sign}(k) \tanh(|k|(\mathtt{h} + c)) \mathtt{p}_k e^{\mathrm{i}kX} = c - \mathcal{H} \tanh((\mathtt{h} + c)|D|) \mathtt{p}(X) \tag{A.12}$$

where $\mathtt{p}(X)$ is defined in (A.9) and $\mathcal{H}$ is the Hilbert transform defined as the Fourier multiplier in (2.32). By (A.12), since $p(X,0) = \mathtt{p}(X)$, condition (A.7) amounts to solve

$$c - \mathcal{H} \tanh((\mathtt{h} + c)|D|) \mathtt{p}(X) = \eta(X + \mathtt{p}(X))\,. \tag{A.13}$$

Remark A.3. If we had required $c = 0$ (fixing the strip of the straight domain (A.4)), equation (A.13) would, in general, have no solution. For example, if $\eta(x) = \eta_0 \neq 0$, then $-\mathcal{H} \tanh(\mathtt{h}|D|) \mathtt{p}(X) = \eta_0$ 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 (A.13) is equivalent to

$$c = \langle \eta(X + \mathtt{p}(X)) \rangle \,, \quad -\mathcal{H} \tanh((\mathtt{h}+c)|D|)\mathtt{p}(X) = \pi_0^\perp \eta(X + \mathtt{p}(X)) \tag{A.14}$$

where $\langle f \rangle = f_0 = \pi_0 f$ is the average in X of any function f, π_0 is defined in (2.33), 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

$$c = \langle \eta(X + \mathtt{p}(X)) \rangle \,, \quad \mathtt{p}(X) = \frac{\mathcal{H}}{\tanh((\mathtt{h}+c)|D|)} [\eta(X + \mathtt{p}(X))] \,. \tag{A.15}$$

Since $\mathcal{H}^2 = -\pi_0^\perp$, if $\mathtt{p}$ solves the second equation in (A.15), then $\mathtt{p}$ also solves the second equation in (A.14).

Lemma A.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 (A.15) satisfying*

$$\|\mathtt{p}\|_s^{k_0,\gamma}, \|c\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|\eta\|_{s+k_0}^{k_0,\gamma} \,, \quad \forall s \geq s_0 \,. \tag{A.16}$$

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

$$\Phi(\mathtt{p}) := \mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)[\eta(\cdot + \mathtt{p}(\cdot))] \,, \qquad \text{where} \quad \mathtt{f}(\xi) := \frac{1}{\tanh(\xi)}, \quad \xi \neq 0 \,, \tag{A.17}$$

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

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

where the cut-off $\chi(\xi)$ is defined in (2.16). For all $n \in \mathbb{N}$, there is a constant $C_n(\mathtt{h}_1) > 0$ such that $|\partial_y^n g(y,\xi)| \leq C_n(\mathtt{h}_1)$ for all $y \geq \mathtt{h}_1/2$, $\xi \in \mathbb{R}$. 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 as g. Now $|c(\lambda,\varphi)| \leq \|\eta\|_{L^\infty} \leq C\|\eta\|_{s_0}$, and therefore $\mathtt{h} + c(\lambda,\varphi) \geq \mathtt{h}_1/2$ 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 a consequence, we have the following estimates for pseudo-differential norms (recall Definition 2.9) of the Fourier multiplier in (A.17): for all $s \geq s_0$,

$$|\!|\!|\mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)|\!|\!|_{0,s,0}^{k_0,\gamma} \,, \; |\!|\!|\mathcal{H}|D|\mathtt{f}'\big((\mathtt{h}+c)|D|\big)|\!|\!|_{0,s,0}^{k_0,\gamma} \lesssim_{s,k_0} 1 + \|c\|_s^{k_0,\gamma} \,. \tag{A.18}$$

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

$$\|c\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|\eta\|_{s+k_0}^{k_0,\gamma} + \|\mathtt{p}\|_s^{k_0,\gamma} \|\eta\|_{s_0+k_0+1}^{k_0,\gamma} \,. \tag{A.19}$$

Therefore by (A.18), (A.19) we get, for all $s \geq s_0$,

$$|\!|\!|\mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)|\!|\!|_{0,s,0}^{k_0,\gamma}, \; |\!|\!|\mathcal{H}|D|\mathtt{f}'\big((\mathtt{h}+c)|D|\big)|\!|\!|_{0,s,0}^{k_0,\gamma} \lesssim_{s,k_0} 1 + \|\eta\|_{s+k_0}^{k_0,\gamma} + \|\mathtt{p}\|_s^{k_0,\gamma} \|\eta\|_{s_0+k_0+1}^{k_0,\gamma} \,. \tag{A.20}$$

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

Step 1: Contraction in low norm. For any $\|\mathtt{p}\|_{2s_0+1}^{k_0,\gamma} \leq r \leq \delta(s_0,k_0)$, by (2.69), (A.20), (2.11), and using the bound $\|\eta\|_{s_0+k_0+1}^{k_0,\gamma} \leq 1$, we have, $\forall s \geq s_0$,

$$\|\Phi(\mathtt{p})\|_s^{k_0,\gamma} \lesssim_{s,k_0} \|\eta\|_{s+k_0}^{k_0,\gamma} + \|\eta\|_{s_0+k_0+1}^{k_0,\gamma} \|\mathtt{p}\|_s^{k_0,\gamma} \,. \tag{A.21}$$

We fix $r := 2C(s_0,k_0)\|\eta\|^{k_0,\gamma}_{2s_0+k_0+1}$ and we assume that $r \leq 1$. Then, using (A.21) with $s = 2s_0+1$, one deduces that 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

$$\Phi'(\mathtt{p})[\tilde{\mathtt{p}}] = \mathcal{A}(\mathtt{m}\,\tilde{\mathtt{p}})\,, \tag{A.22}$$

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

$$\mathcal{A}(h) := \langle h\rangle \mathcal{H}\mathtt{f}'((\mathtt{h}+c)|D|)|D|[\eta(X+\mathtt{p}(X))] + \mathcal{H}\mathtt{f}((\mathtt{h}+c)|D|)[h], \quad \mathtt{m} := \eta_x(X+\mathtt{p}(X))\,. \tag{A.23}$$

To obtain (A.22)-(A.23), note that $\partial_{\mathtt{p}} c[\tilde{\mathtt{p}}] = \langle \mathtt{m}\tilde{\mathtt{p}}\rangle$. By (2.11), for all $s \geq s_0$,

$$\|\mathtt{m}\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0+1} + \|\mathtt{p}\|^{k_0,\gamma}_s \|\eta\|^{k_0,\gamma}_{s_0+k_0+2}\,. \tag{A.24}$$

By (2.69), (A.20), (2.11), using the bounds $\|\eta\|^{k_0,\gamma}_{s_0+k_0+1} \leq 1$ and $\|\mathtt{p}\|^{k_0,\gamma}_{s_0} \leq 1$, we get, for all $s \geq s_0$,

$$|\!|\!|\mathcal{A}|\!|\!|^{k_0,\gamma}_{0,s,0} \lesssim_{s,k_0} 1 + \|\eta\|^{k_0,\gamma}_{s+k_0} + \|\mathtt{p}\|^{k_0,\gamma}_s \|\eta\|^{k_0,\gamma}_{s_0+k_0+1}\,. \tag{A.25}$$

By (A.22), (2.44), (A.24), (A.25) we deduce that, for all $s \geq s_0$,

$$|\!|\!|\Phi'(\mathtt{p})|\!|\!|^{k_0,\gamma}_{0,s,0} \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0+1} + \|\mathtt{p}\|^{k_0,\gamma}_s \|\eta\|^{k_0,\gamma}_{s_0+k_0+2}\,. \tag{A.26}$$

In particular, by (A.26) at $s = 2s_0+1$, and (2.69), we get

$$\|\Phi'(\mathtt{p})[\tilde{\mathtt{p}}]\,\|^{k_0,\gamma}_{2s_0+1} \leq C(s_0,k_0)\|\eta\|^{k_0,\gamma}_{2s_0+k_0+2}\|\tilde{\mathtt{p}}\|^{k_0,\gamma}_{2s_0+1} \leq \frac{1}{2}\|\tilde{\mathtt{p}}\|^{k_0,\gamma}_{2s_0+1} \tag{A.27}$$

provided $C(s_0,k_0)\|\eta\|^{k_0,\gamma}_{2s_0+k_0+2} \leq 1/2$. 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 (A.21), using that $\mathtt{p} = \Phi(\mathtt{p})$ there is $C(s_0,k_0) > 0$ such that if $C(s_0,k_0)\|\eta\|^{k_0,\gamma}_{s_0+k_0+1} \leq 1/2$ for all $s \in [s_0, 2s_0+1]$, one has $\|\mathtt{p}\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0}$. Using also (A.19) one deduces $\|c\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0}$ for all $s \in [s_0, 2s_0+1]$. Thus we have proved (A.16) 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 (A.16) arguing by induction on s. Assume that, for a given $s \geq 2s_0+1$, we have already proved that

$$\|\mathtt{p}\|^{k_0,\gamma}_s\,,\ \|c\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0}\,. \tag{A.28}$$

We want to prove that (A.28) holds for $s+1$. We have to estimate $\|\mathtt{p}\|^{k_0,\gamma}_{s+1} \simeq \max\{\|\mathtt{p}\|^{k_0,\gamma}_s,\ \|\partial_X\mathtt{p}\|^{k_0,\gamma}_s,\ \|\partial_{\varphi_i}\mathtt{p}\|^{k_0,\gamma}_s,\ i = 1,\ldots,\nu\}$. Using the definition (A.17) 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

$$\mathtt{p}_X = \mathcal{H}\mathtt{f}\big((\mathtt{h}+c)|D|\big)[\eta_x(X+\mathtt{p}(X))(1+\mathtt{p}_X)] = \Phi'(\mathtt{p})[\mathtt{p}_X] + \mathcal{A}(\mathtt{m}) \tag{A.29}$$

where the operator $\Phi'(\mathtt{p})$ is given by (A.22) and $\mathcal{A}, \mathtt{m}$ are defined in (A.23) (note that $\langle \eta_x(X+\mathtt{p}(X))(1+\mathtt{p}_X(X))\rangle = 0$). By (A.26) at $s = s_0$, for $\|\eta\|^{k_0,\gamma}_{s_0+k_0+2} \leq \delta(s_0,k_0)$ small enough, condition (2.52) for $A = -\Phi'(\mathtt{p})$ (with $\alpha = 0$) holds. Therefore the operator $\mathrm{Id} - \Phi'(\mathtt{p})$ is invertible and, by (2.53) (with $\alpha = 0$), (A.28) and (2.69), its inverse satisfies, for all $s \geq s_0$,

$$\|(\mathrm{Id} - \Phi'(\mathtt{p}))^{-1}h\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|h\|^{k_0,\gamma}_s + \|\eta\|^{k_0,\gamma}_{s+k_0+1}\|h\|^{k_0,\gamma}_{s_0}\,. \tag{A.30}$$

By (A.29), we deduce that $\mathtt{p}_X = (\mathrm{Id} - \Phi'(\mathtt{p}))^{-1}\mathcal{A}(\mathtt{m})$. By (2.69), (A.24)-(A.25) and (A.28), we get $\|\mathcal{A}(\mathtt{m})\|^{k_0,\gamma}_s \lesssim_s \|\eta\|^{k_0,\gamma}_{s+k_0+1}$. Hence, by (A.30), using $\|\eta\|^{k_0,\gamma}_{s_0+k_0+2} \leq 1$, we get

$$\|\mathtt{p}_X\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0+1}\,. \tag{A.31}$$

We similar arguments we get $\|\partial_{\varphi_i}\mathtt{p}\|^{k_0,\gamma}_s \lesssim_{s,k_0} \|\eta\|^{k_0,\gamma}_{s+k_0+1}$, $i = 1,\ldots,\nu$, and using (A.28), (A.31), we deduce (A.28) at $s+1$ for $\mathtt{p}$. By (A.19), the same estimate holds for c, and the induction step is proved. This completes the proof of (A.16).

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 $\mathcal{C}^1_b$ follows by the implicit function theorem. $\square$

Notice that (A.2) implies the smallness condition of Lemma A.4. Now we transform (1.3) via the conformal diffeomorphism

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

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

where c and $\mathtt{p}$ are the solutions of (A.15) provided by Lemma A.4. 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 $U_X = V_Y$, $U_Y = -V_X$ (or equivalently $p_X = q_Y$, $p_Y = -q_X$) and (A.6)-(A.8),

$$\Delta\phi = 0 \text{ in } \{-\mathtt{h}-c < Y < 0\}\,, \quad \phi(X,0) = (P\psi)(X)\,, \quad \phi_Y(X,-\mathtt{h}-c) = 0\,. \tag{A.32}$$

We calculate explicitly the solution ϕ of (A.32), which is (see (A.10))

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

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

$$\phi_Y(X,0) = \sum_{k\neq 0} \widehat{(P\psi)}_k \tanh(|k|(\mathtt{h}+c))|k| e^{\mathrm{i}kX} = |D|\tanh((\mathtt{h}+c)|D|)(P\psi)(X)\,. \tag{A.33}$$

Lemma A.5. *$G(\eta) = \partial_x P^{-1}\mathcal{H}\tanh((\mathtt{h}+c)|D|)\,P$.*

Proof. The proof is the same as the one of Lemma 2.40 in [21]. The only difference is that formula (A.33) in the case of infinite depth is given by $\phi_Y(X,0) = |D|(P\psi)(X)$. □

Proof of Proposition A.1 concluded. By Lemma A.5 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, using the decomposition (7.41),

$$\begin{aligned}\mathcal{R}_G^{(1)}(\eta) &:= \partial_x\big(P^{-1}\mathcal{H}\tanh((\mathtt{h}+c)|D|)P - \mathcal{H}\tanh((\mathtt{h}+c)|D|)\big)\\ &= \partial_x(P^{-1}\mathcal{H}P - \mathcal{H}) + \partial_x(P^{-1}\mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c})P - \mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c}))\,.\end{aligned} \tag{A.34}$$

The second term $\mathcal{R}_G^{(2)}(\eta)$ is

$$\mathcal{R}_G^{(2)}(\eta) := \partial_x\mathcal{H}\big(\tanh((\mathtt{h}+c)|D|) - \tanh(\mathtt{h}|D|)\big) = \partial_x\mathcal{H}\mathrm{Op}(r_{\mathtt{h}+c} - r_{\mathtt{h}}) = c\,\partial_x\mathcal{H}\mathrm{Op}(\breve{r}_{\mathtt{h},c}) \in OPS^{-\infty}\,, \tag{A.35}$$

where

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

Estimate (A.3) directly follows estimating (A.34) and (A.35) by Lemmata 2.17, 2.18, and using Lemma A.4. The differentiablility of the map $\{\|\eta\|_{s_1+6} < \delta(s_1)\} \to H^{s_1}(\mathbb{T}^\nu\times\mathbb{T}\times\mathbb{T})$, $\eta\mapsto K_G(\eta)$ follows by the differentiability of the map $\{\|\eta\|_{s_1+2} < \delta(s_1)\} \to H^{s_1}_\varphi \times H^{s_1}$, $\eta\mapsto(c(\eta),\mathtt{p}(\eta))$ proved in Lemma A.4.

B Whitney differentiable functions

The following definition is the one in Section 2.3, Chapter VI of [59], for Banach-valued functions.

Definition B.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 $f^{(j)} : F \to Y$, $j \in \mathbb{N}^n$, $0 \leq |j| \leq k$, with $f^{(0)} = f$, and a constant $M > 0$ such that if $R_j(x,y)$ is defined by*

$$f^{(j)}(x) = \sum_{\ell \in \mathbb{N}^n : |j+\ell| \leq k} \frac{1}{\ell!} f^{(j+\ell)}(y)\,(x-y)^{\ell} + R_j(x,y), \quad x, y \in F, \tag{B.1}$$

then

$$\|f^{(j)}(x)\|_Y \leq M, \quad \|R_j(x,y)\|_Y \leq M|x-y|^{\rho-|j|}, \quad \forall x, y \in F, \ |j| \leq k. \tag{B.2}$$

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* (B.2) *holds, namely*

$$\|f\|_{\mathrm{Lip}(\rho,F,Y)} := \inf\{M > 0 : \text{(B.2) } holds\}. \tag{B.3}$$

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 (B.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

$$C_1 \|f\|_{W^{k+1,\infty}(\mathbb{R}^n,Y)} \leq \|f\|_{\mathrm{Lip}(k+1,\mathbb{R}^n,Y)} \leq C_2 \|f\|_{W^{k+1,\infty}(\mathbb{R}^n,Y)} \tag{B.4}$$

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 B.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,*

$$\|\mathcal{E}_k f\|_{\mathrm{Lip}(\rho,\mathbb{R}^n,Y)} \leq C \|f\|_{\mathrm{Lip}(\rho,F,Y)}, \quad \forall f \in \mathrm{Lip}(\rho, F, Y), \tag{B.5}$$

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. $\square$

Clearly, since $\mathcal{E}_k f$ is an extension of f, one has

$$\|f\|_{\mathrm{Lip}(\rho,F,Y)} \leq \|\mathcal{E}_k f\|_{\mathrm{Lip}(\rho,\mathbb{R}^n,Y)} \leq C \|f\|_{\mathrm{Lip}(\rho,F,Y)}. \tag{B.6}$$

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 B.2 does not depend on the index of the space (namely s).

Lemma B.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 B.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], which relies on a nontrivial decomposition in cubes of the domain $\mathbb{R}^n$ only. $\square$

Thanks to the equivalence (B.6), Lemma B.3, and (B.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 B.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

$$\lambda = \gamma\mu, \qquad \gamma^{|j|} u^{(j)}(\lambda) = \gamma^{|j|} u^{(j)}(\gamma\mu) =: U^{(j)}(\mu) = U^{(j)}(\gamma^{-1}\lambda), \qquad U := \mathcal{R}_\gamma u\,. \tag{B.7}$$

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

$$\|u\|_{s,F}^{k+1,\gamma} = \|U\|_{s,\gamma^{-1}F}^{k+1,1}\,. \tag{B.8}$$

Under the rescaling $\mathcal{R}_\gamma$, (B.4) gives the equivalence of the two norms

$$\|f\|_{W^{k+1,\infty,\gamma}(\mathbb{R}^{\nu+1}, H^s)} := \sum_{|\alpha| \leq k+1} \gamma^{|\alpha|} \|\partial_\lambda^\alpha f\|_{L^\infty(\mathbb{R}^{\nu+1}, H^s)} \sim_{\nu,k} \|f\|_{s,\mathbb{R}^{\nu+1}}^{k+1,\gamma}\,. \tag{B.9}$$

Moreover, given $u \in \mathrm{Lip}(k+1, F, s, \gamma)$, its extension

$$\tilde{u} := \mathcal{R}_\gamma^{-1} \mathcal{E}_k \mathcal{R}_\gamma u \in \mathrm{Lip}(k+1, \mathbb{R}^{\nu+1}, s, \gamma) \quad \text{satisfies} \quad \|u\|_{s,F}^{k+1,\gamma} \sim_{\nu,k} \|\tilde{u}\|_{s,\mathbb{R}^{\nu+1}}^{k+1,\gamma}\,. \tag{B.10}$$

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 $0 \leq |j| \leq k$, $\lambda, \lambda_0 \in F$, and the usual smoothing estimates $\|\Pi_N f\|_s \leq N^\alpha \|f\|_{s-\alpha}$ and $\|\Pi_N^\perp f\|_s \leq N^{-\alpha}\|f\|_{s+\alpha}$ for Sobolev functions. $\square$

Proof of Lemma 2.3. Inequality (2.8) follows from the classical interpolation inequality $\|u\|_s \leq \|u\|_{s_0}^\theta \|u\|_{s_1}^{1-\theta}$, $s = \theta s_0 + (1-\theta)s_1$ for Sobolev functions, and from the Definition 2.1 of Whitney-Sobolev norms, since

$$\begin{aligned}
\gamma^{|j|}\|u^{(j)}(\lambda)\|_s &\leq (\gamma^{|j|}\|u^{(j)}(\lambda)\|_{s_0})^\theta (\gamma^{|j|}\|u^{(j)}(\lambda)\|_{s_1})^{1-\theta} \leq (\|u\|_{s_0,F}^{k+1,\gamma})^\theta (\|u\|_{s_1,F}^{k+1,\gamma})^{1-\theta}, \\
\gamma^{k+1}\|R_j(\lambda,\lambda_0)\|_s &\leq (\gamma^{k+1}\|R_j(\lambda,\lambda_0)\|_{s_0})^\theta (\gamma^{k+1}\|R_j(\lambda,\lambda_0)\|_{s_1})^{1-\theta} \leq (\|u\|_{s_0,F}^{k+1,\gamma})^\theta (\|u\|_{s_1,F}^{k+1,\gamma})^{1-\theta} |\lambda - \lambda_0|^{k+1-|j|}.
\end{aligned}$$

Inequality (2.9) follows from (2.8) by using the asymmetric Young inequality (like in Lemma 2.2 in [21]). $\square$

Proof of Lemma 2.4. By (B.9)-(B.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). $\square$

For any $\rho > 0$, we define the $\mathcal{C}^\infty$ function $h_\rho : \mathbb{R} \to \mathbb{R}$,

$$h_\rho(y) := \frac{\chi_\rho(y)}{y} = \frac{\chi(y\rho^{-1})}{y}\,, \quad \forall y \in \mathbb{R} \setminus \{0\}, \quad h_\rho(0) := 0\,, \tag{B.11}$$

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

$$h_\rho(y) = \frac{1}{y}\,,\ \forall |y| \ge \frac{2\rho}{3}\,, \qquad |h_\rho(y)| \le \frac{3}{\rho}\,,\ \forall y \in \mathbb{R}\,. \tag{B.12}$$

To prove Lemma 2.5, we use the following preliminary lemma.

Lemma B.4. *Let $f : \mathbb{R}^{\nu+1} \to \mathbb{R}$ and $\rho > 0$. Then the function*

$$g(\lambda) := h_\rho(f(\lambda)), \quad \forall \lambda \in \mathbb{R}^{\nu+1}\,, \tag{B.13}$$

where h_ρ is defined in (B.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

$$\gamma^{|\alpha|} |\partial_\lambda^\alpha f(\lambda)| \le M\,, \quad \forall \alpha \in \mathbb{N}^{\nu+1}\,, \quad 1 \le |\alpha| \le k+1\,, \tag{B.14}$$

for some $M \ge \rho$, then the function g is in $W^{k+1,\infty}(\mathbb{R}^{\nu+1}, \mathbb{R})$ and

$$\gamma^{|\alpha|} |\partial_\lambda^\alpha g(\lambda)| \le C_k \frac{M^{k+1}}{\rho^{k+2}}\,, \quad \forall \alpha \in \mathbb{N}^{\nu+1}, \quad 0 \le |\alpha| \le k+1. \tag{B.15}$$

Proof. Formula (B.15) for $\alpha = 0$ holds by (B.12). For $|\alpha| \ge 1$, we use the Faà di Bruno formula and (B.14). $\square$

Proof of Lemma 2.5. The function $(\omega \cdot \partial_\varphi)^{-1}_{ext} u$ defined in (2.15) is

$$\big((\omega \cdot \partial_\varphi)^{-1}_{ext} u\big)(\lambda, \varphi, x) = -\mathrm{i} \sum_{(\ell, j) \in \mathbb{Z}^{\nu+1}} g_\ell(\lambda) u_{\ell,j}(\lambda)\, e^{\mathrm{i}(\ell \cdot \varphi + jx)}\,,$$

where $g_\ell(\lambda) = h_\rho(\omega \cdot \ell)$ in (B.13) with $\rho = \gamma \langle \ell \rangle^{-\tau}$ and $f(\lambda) = \omega \cdot \ell$. The function $f(\lambda)$ satisfies (B.14) with $M = \gamma |\ell|$. Hence $g_\ell(\lambda)$ satisfies (B.15), namely

$$\gamma^{|\alpha|} |\partial_\lambda^\alpha g_\ell(\lambda)| \le C_k \gamma^{-1} \langle \ell \rangle^\mu \quad \forall \alpha \in \mathbb{N}^{\nu+1},\ 0 \le |\alpha| \le k+1, \tag{B.16}$$

where $\mu = k+1+(k+2)\tau$ is defined in (2.18). By the product rule and using (B.16), we deduce $\gamma^{|\alpha|} \|\partial_\lambda^\alpha((\omega \cdot \partial_\varphi)^{-1}_{ext} u)(\lambda)\|_s \le C_k \gamma^{-1} \|u\|^{k+1,\gamma}_{s+\mu, \mathbb{R}^{\nu+1}}$ and therefore (2.17). The proof is concluded by observing that the restriction of $(\omega \cdot \partial_\varphi)^{-1}_{ext} u$ to F gives $(\omega \cdot \partial_\varphi)^{-1} u$ as defined in (2.14), and (2.18) follows by (B.10). $\square$

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 (B.10). Then we observe that the composition $\mathtt{f}(\tilde u)$ is an extension of $\mathtt{f}(u)$, and therefore one has the inequality $\|\mathtt{f}(u)\|^{k+1,\gamma}_{s,F} \le \|\mathtt{f}(\tilde u)\|^{k+1,\gamma}_{s,\mathbb{R}^{\nu+1}} \sim \|\mathtt{f}(\tilde u)\|_{W^{k+1,\infty,\gamma}(\mathbb{R}^{\nu+1}, H^s)}$ by (B.9). Then (2.19) follows by the Moser composition estimates for $\|\ \|^{k+1,\gamma}_{s,\mathbb{R}^{\nu+1}}$ (see for instance Lemma 2.31 in [21]), together with the equivalence of the norms in (B.9)-(B.10). $\square$

C A Nash-Moser-Hörmander implicit function theorem

Let $(E_a)_{a \ge 0}$ be a decreasing family of Banach spaces with continuous injections $E_b \hookrightarrow E_a$,

$$\|u\|_{E_a} \le \|u\|_{E_b} \quad \text{for } a \le b. \tag{C.1}$$

Set $E_\infty = \cap_{a \ge 0} E_a$ 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,

$$\|S_j u\|_{E_a} \le C \|u\|_{E_a} \qquad \text{for all } a; \tag{C.2}$$

$$\|S_j u\|_{E_b} \le C 2^{j(b-a)} \|S_j u\|_{E_a} \qquad \text{if } a < b; \tag{C.3}$$

$$\|u - S_j u\|_{E_b} \le C 2^{-j(a-b)} \|u - S_j u\|_{E_a} \qquad \text{if } a > b; \tag{C.4}$$

$$\|(S_{j+1} - S_j) u\|_{E_b} \le C 2^{j(b-a)} \|(S_{j+1} - S_j) u\|_{E_a} \qquad \text{for all } a, b. \tag{C.5}$$

Set

$$R_0 u := S_1 u, \qquad R_j u := (S_{j+1} - S_j)u, \quad j \geq 1. \tag{C.6}$$

We also assume that

$$\|u\|_{E_a}^2 \leq C \sum_{j=0}^{\infty} \|R_j u\|_{E_a}^2 \quad \forall a \geq 0, \tag{C.7}$$

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 C.1 ([10])**. (Existence)** *Let $a_1, a_2, \alpha, \beta, a_0, \mu$ be real numbers with*

$$0 \leq a_0 \leq \mu \leq a_1, \qquad a_1 + \frac{\beta}{2} < \alpha < a_1 + \beta, \qquad 2\alpha < a_1 + a_2. \tag{C.8}$$

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

$$\begin{aligned} \|\Phi''(u)[v,w]\|_{F_a} \leq M_1(a)\big(\|v\|_{E_{a+\mu}} \|w\|_{E_{a_0}} + \|v\|_{E_{a_0}} \|w\|_{E_{a+\mu}}\big) \\ + \{M_2(a)\|u\|_{E_{a+\mu}} + M_3(a)\} \|v\|_{E_{a_0}} \|w\|_{E_{a_0}} \end{aligned} \tag{C.9}$$

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

$$\|\Psi(v)g\|_{E_a} \leq L_1(a)\|g\|_{F_{a+\beta-\alpha}} + \{L_2(a)\|v\|_{E_{a+\beta}} + L_3(a)\}\|g\|_{F_0} \quad \forall a \in [a_1, a_2], \tag{C.10}$$

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

$$\sum_{j=0}^{\infty} \|R_j g\|_{F_\beta}^2 \leq A^2 \|g\|_{F_\beta}^2, \quad \|g\|_{F_\beta} \leq \delta, \tag{C.11}$$

there exists $u \in E_\alpha$ solving $\Phi(u) = \Phi(0) + g$. The solution u satisfies

$$\|u\|_{E_\alpha} \leq C L_{123}(a_2)(1 + A)\|g\|_{F_\beta}, \tag{C.12}$$

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

$$\delta = 1/B, \quad B = C' L_{123}(a_2) \max\big\{1/\delta_1, 1 + A, (1 + A) L_{123}(a_2) M_{123}(a_2 - \mu)\big\} \tag{C.13}$$

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* (C.9) *holds for all $a \in [0, a_2 + c - \mu]$, $\Psi(v)$ maps F_∞ to E_{a_2+c}, and* (C.10) *holds for all $a \in [a_1, a_2 + c]$. If g satisfies* (C.11) *and, in addition, $g \in F_{\beta+c}$ with*

$$\sum_{j=0}^{\infty} \|R_j g\|_{F_{\beta+c}}^2 \leq A_c^2 \|g\|_{F_{\beta+c}}^2 \tag{C.14}$$

for some A_c, then the solution u belongs to $E_{\alpha+c}$, with

$$\|u\|_{E_{\alpha+c}} \leq C_c \big\{\mathcal{G}_1 (1 + A)\|g\|_{F_\beta} + \mathcal{G}_2 (1 + A_c)\|g\|_{F_{\beta+c}}\big\} \tag{C.15}$$

where

$$\mathcal{G}_1 := \tilde L_3 + \tilde L_{12}(\tilde L_3 \tilde M_{12} + L_{123}(a_2)\tilde M_3)(1 + z^N), \quad \mathcal{G}_2 := \tilde L_{12}(1 + z^N), \tag{C.16}$$

$$z := L_{123}(a_1) M_{123}(0) + \tilde L_{12} \tilde M_{12}, \tag{C.17}$$

$\tilde L_{12} := \tilde L_1 + \tilde L_2$, $\tilde L_i := L_i(a_2 + c)$, $i = 1,2,3$; $\tilde M_{12} := \tilde M_1 + \tilde M_2$, $\tilde M_i := M_i(a_2 + c - \mu)$, $i = 1,2,3$; *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 [34]. 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 (C.12), (C.15)). Theorem C.1 has the advantage of making explicit all the constants (unlike [34]), which is necessary to deduce the quantitative Theorem 7.3.

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*, J. Funct. Anal. 273 (2017), no. 12, 3875-3900.

[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., *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, volume 9 of Panoramas et Synthèses. Société Mathématique de France, Paris, 2000.

[25] Craig W., Nicholls D., *Travelling two and three dimensional capillary gravity water waves.* SIAM J. Math. Anal., 32(2):323–359 (electronic), 2000.

[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] Eliasson L. H., Kuksin S., *KAM for nonlinear Schrödinger equation*, Annals of Math, (2) 172, no. 1, 371-435, 2010.

[30] 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.

[31] Feola R., Procesi M., *Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations.* J. Diff. Eq., 259, no. 7, 3389-3447, 2015.

[32] Feola R., Procesi M., *KAM for quasi-linear autonomous NLS*, preprint 2018, arXiv:1705.07287.

[33] 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.

[34] Hörmander L., *The boundary problems of physical geodesy.* Arch. Rat. Mech. Anal. 62, no. 1, 1-52, 1976.
[35] Hörmander L., *The analysis of linear partial differential operators III.* Springer-Verlag, Berlin, 1990.
[36] Ifrim M., Tataru D., *Two dimensional gravity water waves with constant vorticity: I. Cubic lifespan*, arXiv:1510.07732.
[37] Ionescu A., Pusateri F., *Global solutions for the gravity water waves system in 2d*, Invent. Math., 199, 3, 653-804, 2015.
[38] Iooss G., Plotnikov P., *Existence of multimodal standing gravity waves.* J. Math. Fluid Mech., 7(suppl. 3):S349–S364, 2005.
[39] Iooss G., Plotnikov P., *Multimodal standing gravity waves: a completely resonant system.* J. Math. Fluid Mech., 7(suppl. 1):S110–S126, 2005.
[40] 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.
[41] Iooss G., Plotnikov P., *Asymmetrical tridimensional traveling gravity waves.* Arch. Rat. Mech. Anal., 200(3):789-880, 2011.
[42] 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.
[43] Kappeler T., Pöschel J., *KdV & KAM*, Springer-Verlag, 2003.
[44] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional Anal. i Prilozhen. 2, 22-37, 95, 1987.
[45] Kuksin S., *Analysis of Hamiltonian PDEs*, volume 19 of Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, 2000.
[46] Lannes D., *Well-posedness of the water-waves equations*, J. Amer. Math. Soc., 3, 605–654, 18, 2005.
[47] Lannes D., *The water waves problem: mathematical analysis and asymptotics.* Mathematical Surveys and Monographs, 188, 2013.
[48] Levi-Civita T., *Détermination rigoureuse des ondes permanentes d' ampleur finie.* Math. Ann., 93 , pp. 264-314, 1925.
[49] Liu J., Yuan X., *A KAM theorem for Hamiltonian partial differential equations with unbounded perturbations.* Comm. Math. Phys., 307(3), 629–673, 2011.
[50] Métivier G., *Para-differential Calculus and Applications to the Cauchy Problem for Nonlinear Systems.* Pubblicazioni Scuola Normale Pisa, 5, 2008.
[51] Montalto R., *Quasi-periodic solutions of forced Kirchhoff equation.* NoDEA, Nonlinear Differ. Equ. Appl., 24(1), 9, 2017.
[52] Moser J., *Convergent series expansions for quasi-periodic motions*, Math.Ann. 169, 136-176, 1967.
[53] Plotnikov P., Toland J., *Nash-Moser theory for standing water waves.* Arch. Ration. Mech. Anal., 159(1):1–83, 2001.
[54] Pöschel J., *A KAM-Theorem for some nonlinear PDEs*, Ann. Sc. Norm. Pisa, 23, (1996) 119-148.
[55] Procesi C., Procesi M., *A KAM algorithm for the completely resonant nonlinear Schrödinger equation*, Advances in Mathematics, volume 272, 399-470, 2015.
[56] Pjartli A.S., *Diophantine approximations of submanifolds of a Euclidean space*, Funktsional. Anal. i Prilozhen. 3 (4) (1969) 59-62.
[57] Robbin J., Salamon D., *The exponential Vandermonde matrix*, Linear Algebra Appl. 317, 2000, 225.
[58] Rüssmann H., *Invariant tori in non-degenerate nearly integrable Hamiltonian systems*, Regul. Chaotic Dyn. 6 (2), 119-204, 2001.
[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