Large KAM tori for quasi-linear perturbations of KdV

Massimiliano Berti, Thomas Kappeler, Riccardo Montalto

Abstract. We prove the persistence of "most" finite gap solutions of the KdV equation on the circle under sufficiently small and smooth *quasi-linear* Hamiltonian perturbations. The proof makes use of suitable symplectic coordinates, introduced by Kappeler-Montalto [15], in the vicinity of any finite-gap manifold, which admit a pseudo-differential expansion. Then we implement a Nash-Moser iteration scheme. A key step is to diagonalize the linearized operators, obtained at any approximate quasi-periodic solution, with sharp asymptotic estimates of their eigenvalues.

Keywords: KdV equation, KAM for PDE, Birkhoff coordinates, quasi-periodic solutions, finite-gap solutions.
MSC 2010: 37K55, 35Q53, 37K10.

Contents

1 Introduction

The aim of this paper is to investigate the stability of finite gap solutions of the KdV equation – also referred to as space periodic multi-solitons – under *quasi-linear* Hamiltonian perturbations

$$\partial_t u = -\partial_x^3 u + 6u\partial_x u + \varepsilon\partial_x \nabla P(u), \quad x \in \mathbb{T}_1 := \mathbb{R}/\mathbb{Z}, \tag{1.1}$$

where $\varepsilon \in (0,1)$ is a small parameter and ∇P is the L^2-gradient of

$$P(u) := \int_{\mathbb{T}_1} f(x, u(x), u_x(x))\, dx, \quad u_x := \partial_x u, \quad f \in \mathcal{C}^\infty(\mathbb{T}_1 \times \mathbb{R} \times \mathbb{R}, \mathbb{R}). \tag{1.2}$$

Equation (1.1) then reads

$$\partial_t u = \partial_x \nabla H_\varepsilon(u) \tag{1.3}$$

with Hamiltonian

$$H_\varepsilon(u) := H^{kdv}(u) + \varepsilon P(u), \qquad H^{kdv}(u) := \int_{\mathbb{T}_1} \frac{1}{2}u_x^2(x) + u^3(x)\, dx. \tag{1.4}$$

As phase spaces for the Hamiltonian PDE (1.3) we choose $H_0^s(\mathbb{T}_1)$, $s \geq 0$, where

$$H_0^s(\mathbb{T}_1) := \Big\{ u \in H^s(\mathbb{T}_1) : \int_{\mathbb{T}_1} u(x)\, dx = 0 \Big\}, \quad L_0^2(\mathbb{T}_1) \equiv H_0^0(\mathbb{T}_1), \tag{1.5}$$

$H^s(\mathbb{T}_1)$ denotes the Sobolev space of real valued functions

$$H^s(\mathbb{T}_1) := \Big\{ u(x) = \sum_{n\in\mathbb{Z}} u_n e^{\mathrm{i}2\pi nx} \; : \; \|u\|_{H_x^s} := \big(\sum_{n\in\mathbb{Z}} \langle n\rangle^{2s} |u_n|^2\big)^{\frac{1}{2}} < \infty, \; \overline{u_n} = u_{-n} \;\; \forall n \in \mathbb{Z} \Big\} \tag{1.6}$$

and $\langle n\rangle := \max\{1, |n|\}$ for any $n \in \mathbb{Z}$. We also write $L^2(\mathbb{T}_1)$ for $H^0(\mathbb{T}_1)$. The symplectic form on the phase space $L_0^2(\mathbb{T}_1)$ is

$$\mathcal{W}_{L_0^2}(u, v) := \int_{\mathbb{T}_1} (\partial_x^{-1} u) v\, dx, \quad \forall u, v \in L_0^2(\mathbb{T}_1), \tag{1.7}$$

where the operator ∂_x^{-1} is defined in (2.19). The Hamiltonian vector field $X_H(u) = \partial_x \nabla H(u)$ associated with the Hamiltonian H, is characterized by $dH(u)[\cdot] = \mathcal{W}_{L_0^2}(X_H, \cdot)$.

In order to state our main result, let us first describe the dynamics of equation (1.1), when $\varepsilon = 0$. It is proved in Kappeler-Pöschel [16] that the KdV equation $\partial_t u = -\partial_x^3 u + 6uu_x$ admits global analytic Birkhoff coordinates, which we describe below. As a consequence, all its solutions are periodic, quasi-periodic or almost periodic in time. The quasi-periodic solutions of the KdV equation are referred to as *finite gap solutions* or alternatively *space periodic multi-solitons.*

Birkhoff coordinates: For any $s \geq 0$, let $h_0^s := \{ z = (z_n)_{n\in\mathbb{Z}} \in h^s : z_0 = 0 \}$ where

$$h^s := \Big\{ z = (z_n)_{n\in\mathbb{Z}}, \; z_n \in \mathbb{C} \, : \, \|z\|_s^2 := \sum_{n\in\mathbb{Z}} \langle n\rangle^{2s} |z_n|^2 < \infty, \; \overline{z_n} = z_{-n}, \; \forall n \in \mathbb{Z} \Big\},$$

endowed with the standard Poisson bracket defined by

$$\{z_n, z_k\} = \mathrm{i}2\pi n\, \delta_{k,-n}, \quad \forall\, n, k \in \mathbb{Z}.$$

By $\mathcal{F}$ we denote the Fourier transform, $\mathcal{F} : L^2(\mathbb{T}_1) \to h^0$, $u \mapsto (u_n)_{n\in\mathbb{Z}}$, where $u_n := \int_{\mathbb{T}_1} u(x) e^{-\mathrm{i}2\pi nx}\, dx$ for any $n \in \mathbb{Z}$ and by $\mathcal{F}^{-1} : h^0 \to L^2(\mathbb{T}_1)$ its inverse.

Theorem 1.1. *([16]) There exists a real analytic diffeomorphism $\Psi^{kdv} : h_0^0 \to H_0^0(\mathbb{T}_1)$ so that:*
(i) for any $s \in \mathbb{Z}_{\geq 0}$, $\Psi^{kdv}(h_0^s) \subseteq H_0^s(\mathbb{T}_1)$ and $\Psi^{kdv} : h_0^s \to H_0^s(\mathbb{T}_1)$ is a real analytic symplectic diffeomorphism.
(ii) $H^{kdv} \circ \Psi^{kdv} : h_0^1 \to \mathbb{R}$ is a real analytic function of the actions $I_k := \frac{1}{2\pi k} z_k z_{-k}$, $k \geq 1$. The KdV Hamiltonian, viewed as a function of the actions $(I_k)_{k\geq 1}$, is denoted by $\mathcal{H}_o^{kdv}$.
(iii) $\Psi^{kdv}(0) = 0$ and the differential $d_0\Psi^{kdv}$ of Ψ^{kdv} at 0 is the inverse Fourier transform $\mathcal{F}^{-1}$.

As a consequence of Theorem 1.1, the KdV equation, expressed in the Birkhoff coordinates $(z_n)_{n\neq 0}$, reads

$$\partial_t z_n = \mathrm{i}\omega_n^{kdv}((I_k)_{k\geq 1}) z_n\,,\ \forall n \in \mathbb{Z}\setminus\{0\}\,,\quad \omega_{\pm m}^{kdv}((I_k)_{k\geq 1}) := \pm\partial_{I_m}\mathcal{H}_o^{kdv}((I_k)_{k\geq 1})\,,\ \forall m\geq 1\,, \tag{1.8}$$

and its solutions are

$$z_n(t) = \exp\big(\mathrm{i}\omega_n^{kdv}((I_k^{(0)})_{k\geq 1})\ t\big) z_n(0)\,,\ \forall n\in\mathbb{Z}\setminus\{0\}\,,\quad I_k^{(0)} := \frac{1}{2\pi k} z_k(0) z_{-k}(0)\,,\ \forall k\geq 1\,.$$

Finite gap solutions: Let us consider a finite set $\mathbb{S}_+ \subset \mathbb{N}_+ := \{1,2,\dots\}$ and define

$$\mathbb{S} := \mathbb{S}_+ \cup (-\mathbb{S}_+)\,,\quad \mathbb{S}_+^\perp := \mathbb{N}_+\setminus\mathbb{S}_+\,,\quad \mathbb{S}^\perp := \mathbb{S}_+^\perp \cup (-\mathbb{S}_+^\perp) \subset \mathbb{Z}\setminus\{0\}\,.$$

A $\mathbb{S}-$gap solution of the KdV equation is a solution of the form

$$z_n(t) = \exp\big(\mathrm{i}\omega_n^{kdv}(\nu,0)t\big) z_n(0)\,,\ z_n(0)\neq 0\,,\ \forall n\in\mathbb{S}\,,\quad z_n(t) = 0\,,\ \forall n\in\mathbb{S}^\perp\,, \tag{1.9}$$

where $\nu := (I_k^{(0)})_{k\in\mathbb{S}_+} \in \mathbb{R}_{>0}^{\mathbb{S}_+}$ and, by a slight abuse of notation, we write

$$\omega_n^{kdv}(I,(I_k)_{k\in\mathbb{S}_+^\perp}) := \omega_n^{kdv}((I_k)_{k\geq 1})\,,\quad I := (I_k)_{k\in\mathbb{S}_+}\in\mathbb{R}_{>0}^{\mathbb{S}_+}\,. \tag{1.10}$$

Such solutions are quasi-periodic in time with frequency vector

$$\omega^{kdv}(\nu) := \big(\omega_n^{kdv}(\nu,0)\big)_{n\in\mathbb{S}_+} \in \mathbb{R}^{\mathbb{S}_+}\,, \tag{1.11}$$

parametrized by $\nu\in\mathbb{R}_{>0}^{\mathbb{S}_+}$. The map $\nu\mapsto\omega^{kdv}(\nu)$ is a local analytic diffeomorphism, see Remark 3.9. When written in action-angle coordinates,

$$\theta := (\theta_n)_{n\in\mathbb{S}_+}\in\mathbb{T}^{\mathbb{S}_+}\,,\quad I = (I_n)_{n\in\mathbb{S}_+}\in\mathbb{R}_{>0}^{\mathbb{S}_+}\,,\quad z_n = \sqrt{2\pi n I_n}e^{-\mathrm{i}\theta_n}\,,\ n\in\mathbb{S}_+\,,$$

instead of the complex Birkhoff coordinates z_n, the $\mathbb{S}-$gap solution (1.9) reads

$$\theta(t) = \theta^{(0)} - \omega^{kdv}(\nu)t\,,\qquad I(t) = \nu\,,\qquad z_n(t) = 0\,,\quad \forall n\in\mathbb{S}^\perp\,. \tag{1.12}$$

Motivated by questions raised by S. Kuksin and V. Zakharov, our aim is to study the stability of these finite gap solutions under quasi-linear perturbations.

In the whole paper $\Xi\subset\mathbb{R}_{>0}^{\mathbb{S}_+}$ is the closure of a bounded open set so that ω^{kdv} defined in (1.11) is a diffeomorphism onto its image. Moreover we require that, for some $\delta>0$ small enough,

$$\Xi + B_{\mathbb{S}_+}(\delta) \subseteq \mathbb{R}_{>0}^{\mathbb{S}_+} \tag{1.13}$$

where $B_{\mathbb{S}_+}(\delta)$ denotes the ball of radius δ in $\mathbb{R}^{\mathbb{S}_+}$ centered at the origin.

The main result of this paper (Theorem 1.2) is that, for ε small enough, and for ν in a subset of Ξ of *large* Lebesgue measure, there is a quasi-periodic solution of equation (1.1), close to the finite gap solution of the KdV equation $q(\theta^{(0)}-\omega^{kdv}(\nu)t,x;\nu)$ where, for any $\nu\in\Xi$,

$$q(\cdot\,,\cdot\,;\nu) : \mathbb{T}^{\mathbb{S}_+} \to H_0^0(\mathbb{T}_1)\,,\ \varphi\mapsto q(\varphi,\cdot\,;\nu)\,,\qquad q(\varphi,x\,;\nu) := \Psi^{kdv}(\varphi,\,\nu,\,0)(x)\,,\ \forall x\in\mathbb{T}_1\,, \tag{1.14}$$

and $\Psi^{kdv}(\theta,\,\nu,\,0)(x) := \Psi^{kdv}((z_n)_{n\in\mathbb{S}},\,0)(x)$. Notice that the function $q(\varphi,x)\equiv q(\varphi,x;\nu)$ is in $\mathcal{C}^\infty(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)$, actually it is real analytic.

Let $H^s\equiv H^s_{\varphi,x}$, $s\geq 0$, denote the Sobolev space $H^s(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)$ of periodic, real valued functions

$$H^s := \Big\{ f = \sum_{(\ell,j)\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{Z}} f_{\ell,j}\, e^{\mathrm{i}(\ell\cdot\varphi+2\pi jx)}\ :\ \|f\|_s^2 := \sum_{(\ell,j)\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{Z}} |f_{\ell,j}|^2\langle\ell,j\rangle^{2s} < \infty,\ \overline{f_{\ell,j}} = f_{-(\ell,j)}\Big\} \tag{1.15}$$

where $\langle\ell,j\rangle := \max\{1,|\ell|,|j|\}$. For $s > (|\mathbb{S}_+|+1)/2$ the embedding $H^s(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)\subset C^0(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)$ holds and $H^s(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)$ is an algebra.

The main result of this paper is the following one:

Theorem 1.2. *Let f be a function in $\mathcal{C}^\infty(\mathbb{T}_1 \times \mathbb{R} \times \mathbb{R}, \mathbb{R})$. Then there exist $\bar{s} > (|\mathbb{S}_+| + 1)/2$ and $\varepsilon_0 \in (0, 1)$ so that, for any $\varepsilon \in (0, \varepsilon_0)$, there exists a measurable subset $\Xi_\varepsilon \subseteq \Xi$ with the following properties:*

$$\lim_{\varepsilon \to 0} |\Xi \setminus \Xi_\varepsilon| = 0$$

and, for any $\nu \in \Xi_\varepsilon$, there exists a quasi-periodic solution $u_\varepsilon(\omega_\varepsilon(\nu)t, x; \nu)$ of equation (1.1) with $u_\varepsilon(\cdot, \cdot; \nu)$ in $H^{\bar{s}}(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1)$ and frequency vector $\omega_\varepsilon(\nu) \in \mathbb{R}^{\mathbb{S}_+}$ so that

$$\lim_{\varepsilon \to 0} \|u_\varepsilon(\cdot, \cdot; \nu) - q(\cdot, \cdot; \nu)\|_{\bar{s}} = 0 \,, \quad \lim_{\varepsilon \to 0} \omega_\varepsilon(\nu) = -\omega^{kdv}(\nu) \,,$$

where $q(\varphi, x; \nu)$ is defined in (1.14) and $\omega^{kdv}(\nu)$ in (1.11). The solution $u_\varepsilon(\omega_\varepsilon(\nu)t, x; \nu)$ is linearly stable.

Remark 1.3. *Actually the same result holds for any density f of class $\mathcal{C}^{s_*}$ for $s_* \in \mathbb{N}$ large enough. We assume f to be $\mathcal{C}^\infty-$smooth merely for simplicity of notation.*

Ideas of the proof: Theorem 1.2 is proved by using a Nash-Moser scheme. One of the main issues concerns the invertibility of the linearized Hamiltonian operator $\omega \cdot \partial_\varphi - \partial_x d\nabla \mathcal{H}_\varepsilon(u(\varphi, x))$ where $u(\omega t, x)$ is an approximate quasi-periodic solution of (1.3), close to the finite gap solutions (1.14). Since the perturbation in (1.1) is quasi-linear, i.e. the perturbed vector field $\partial_x \nabla P(u)$ might contain $\partial_x^3 u$ and hence it is of the same order as the KdV vector field $\partial_x \nabla H^{kdv}(u)$, the KAM reducibility schemes known in literature cannot be applied directly. A key ingredient of the proof are special canonical coordinates, constructed in Kappeler-Montalto [15],

$$\Psi : (\theta, y, w) \mapsto \Psi(\theta, y, w) \in L^2_0(\mathbb{T}_1) \,,$$

defined in a neighborhood of the finite gap manifold $\mathbb{T}_1^{\mathbb{S}_+} \times \{\nu\} \times \{0\}$ in $\mathbb{T}_1^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+}_{>0} \times L^2_\perp(\mathbb{T}_1)$ where

$$L^2_\perp(\mathbb{T}_1) := \Big\{ w = \sum_{n \in \mathbb{S}^\perp} w_n e^{\mathrm{i} 2\pi n x} \in L^2_0(\mathbb{T}_1) \Big\} \tag{1.16}$$

which admits a pseudo-differential expansion. They have the following main properties that we describe in detail in Section 3.1:

(i) For $\varepsilon = 0$, the manifold of $\mathbb{S}-$gap solutions in the range of Ψ is characterized by the equation $w = 0$ and the linearized equation along the manifold $\{w = 0, I = \text{const.}\}$ is in diagonal form with coefficients only depending on I, see Theorem 3.1-**(AE3)**.

(ii) When expressed in these coordinates, the linearized Hamiltonian vector field admits an expansion in terms of pseudo-differential operators, see Section 3.2.

Thanks to their pseudo-differential nature, these coordinates allow us to deal with quasi-linear perturbations. We first perform preliminary transformations (which are Fourier integral operators generated as flows of linear transport PDEs and pseudo-differential maps) which diagonalize the above mentioned linearized operators up to a pseudo-differential operator of order zero plus a regularizing remainder (see Section 6). At this point, using the properties of the KdV frequencies, we are able to perform a KAM reducibility scheme in order to complete the diagonalization. This strategy has been carried out for small amplitude solutions of KdV in [3], [2]. In that case one can directly use the differential structure of (1.1)-(1.2). The novelty of Theorem 1.2 consists in the fact that the unperturbed solutions are not required to be small.

Related work: The first KAM theorems for perturbations of large finite gap solutions of the KdV equation (1.1) were established by Kuksin in [19], see also [20], and by Kappeler-Pöschel in [16], in the case the perturbation is semi-linear, namely the density $f(x, u)$ in (1.2) does not depend on u_x. The key idea is to exploit that the frequencies of KdV grow asymptotically as $\sim j^3$ as $|j| \to +\infty$, and therefore one can impose second order Melnikov non-resonance conditions of the form $|\omega \cdot \ell + j^3 - i^3| \geq (j^2 + i^2)/2$, $i \neq j$, which gain 2 space derivatives (outside the diagonal $i = j$), sufficient to compensate the loss of one space derivative produced by the vector field $\varepsilon \partial_x (\partial_u f)(x, u)$. Subsequently, Liu-Yuan in [13] proved KAM results for semilinear perturbations of small amplitude solutions of the derivative NLS and Benjamin-Ono equations whereas Zhang-Gao-Yuan [21] proved analogous results for the reversible derivative NLS $\mathrm{i}u_t + u_{xx} = |u_x|^2 u$.

More recently, Berti-Biasco-Procesi [5]-[6] proved existence and stability of small quasi-periodic solutions of autonomous derivative Klein-Gordon equations of the form $y_{tt} - y_{xx} + my = g(x, y, y_x, y_t)$ satisfying reversibility conditions.

In all of the work mentioned above, the perturbations are required to be semilinear. Concerning quasi-linear perturbations, the first KAM results of small amplitude solutions of the KdV equation were established by Baldi-Berti-Montalto in [3], [2], by using pseudo-differential calculus, see also [4], [11], [12]. The frequency-amplitude modulation is obtained in [3] by a weak-Birkhoff normal form analysis. Due to the purely differential structure of (1.1), the required tools of pseudo-differential calculus in [3], [2] mainly concern multiplication operators and Fourier multipliers. In order to obtain KAM type results for the water waves equations, more advanced techniques have been developed in Berti-Montalto [9] and Baldi-Berti-Haus-Montalto [1].

For studying perturbations of large finite gap solutions, the Birkhoff coordinates are a natural setting. Existence of large KAM tori for semilinear perturbations of the cubic NLS has been obtained in [8], exploiting that the Birkhoff map is a one smoothing perturbation of the Fourier transform (see [18]). This property is not sufficient to deal with quasi-linear perturbations. It would be useful to exploit some pseudo-differential property of the coordinates, that, however, the Birkhoff map does not seem to possess. As an alternative, Kappeler-Montalto [15] constructed, in the vicinity of finite gap manifolds of KdV, symplectic coordinates admitting a pseudo-differential expansion to any given order, up to a remainder satisfying tame estimates. In the present paper we use these symplectic variables as a starting point and then apply KAM techniques developed in [9], [1] to prove Theorem 1.2.

We expect that by the same methods of proof, a result analogous to Theorem 1.2 can be proved for finite gap solutions of the defocusing NLS equation as well as the defocusing mKdV equation.

Notation. We denote by $\mathbb{N} = \{0, 1, 2, \ldots\}$ the set of natural numbers and set $\mathbb{N}_+ = \{1, 2, \ldots\}$. Given a Banach space X with norm $\|\cdot\|_X$, we denote by by $H^s_\varphi X = H^s(\mathbb{T}^{\mathbb{S}_+}, X)$, $s \in \mathbb{N}$, the Sobolev space of functions $f : \mathbb{T}^{\mathbb{S}_+} \to X$ equipped with the norm

$$\|f\|_{H^s_\varphi X} := \|f\|_{L^2_\varphi X} + \max_{|\beta|=s} \|\partial_\varphi^\beta f\|_{L^2_\varphi X} . \tag{1.17}$$

We also denote $H^0_\varphi X = L^2_\varphi X$. We recall that the continuous Sobolev embedding theorem is stronger in the case X is a Hilbert space H. The corresponding theorems read as follows

$$H^s(\mathbb{T}^{\mathbb{S}_+}, X) \hookrightarrow \mathcal{C}^0(\mathbb{T}^{\mathbb{S}_+}, X) , \quad \forall s > |\mathbb{S}_+| , \qquad H^s(\mathbb{T}^{\mathbb{S}_+}, H) \hookrightarrow \mathcal{C}^0(\mathbb{T}^{\mathbb{S}_+}, H) , \quad \forall s > |\mathbb{S}_+|/2 . \tag{1.18}$$

Let $H^s_x := H^s(\mathbb{T}_1)$, $s \geq 0$, and denote by $\big(f, g\big)_{L^2_x}$ the L^2-inner product on $L^2_x \equiv H^0_x$,

$$\big(f, g\big)_{L^2_x} := \int_{\mathbb{T}_1} f(x) g(x)\, dx . \tag{1.19}$$

Furthermore, we denote by $\Pi_\perp$ the L^2-orthogonal projector onto the subspace $L^2_\perp(\mathbb{T}_1)$, defined in (1.16), and by $\Pi_0^\perp$ the one onto the subspace of functions with zero average. We set

$$H^s_\perp(\mathbb{T}_1) := H^s(\mathbb{T}_1) \cap L^2_\perp(\mathbb{T}_1) \tag{1.20}$$

and

$$H^s_\perp \equiv H^s_\perp(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1) := \big\{ u \in H^s(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1) : u(\varphi, \cdot) \in L^2_\perp(\mathbb{T}_1) \big\} , \tag{1.21}$$

which is an algebra for $s \geq s_0 := [\frac{|\mathbb{S}_+|+1}{2}] + 1$. The space $H^0_\perp$ is also denoted by $L^2_\perp$. Let

$$\mathcal{E}_s := \mathbb{T}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times H^s_\perp(\mathbb{T}_1) , \quad \mathcal{E} \equiv \mathcal{E}_0 , \qquad E_s := \mathbb{R}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times H^s_\perp(\mathbb{T}_1) , \quad E \equiv E_0 , \tag{1.22}$$

where $H^s_\perp(\mathbb{T}_1)$ is defined in (1.20). Elements of $\mathcal{E}$ are denoted by $\mathfrak{x} = (\theta, y, w)$ and the ones of its tangent space E by $\widehat{\mathfrak{x}} = (\widehat{\theta}, \widehat{y}, \widehat{w})$. For $s < 0$, we consider the Sobolev space $H^s_\perp(\mathbb{T}_1)$ of distributions, and the spaces $\mathcal{E}_s$ and E_s are defined in a similar way as in (1.22). Notice that $H^{-s}_\perp(\mathbb{T}_1)$ is the dual space of $H^s_\perp(\mathbb{T}_1)$. On E, we denote by $\langle \cdot, \cdot \rangle$ the inner product, defined by

$$\big\langle (\widehat{\theta}_1, \widehat{y}_1, \widehat{w}_1), (\widehat{\theta}_2, \widehat{y}_2, \widehat{w}_2) \big\rangle := \widehat{\theta}_1 \cdot \widehat{\theta}_2 + \widehat{y}_1 \cdot \widehat{y}_2 + \big(\widehat{w}_1, \widehat{w}_2\big)_{L^2_x} . \tag{1.23}$$

By a slight abuse of notation, $\Pi_\perp$ also denotes the projector of E_s onto its third component,

$$\Pi_\perp : E_s \to H^s_\perp(\mathbb{T}_1)\,, \quad (\widehat{\theta}, \widehat{y}, \widehat{w}) \mapsto \widehat{w}\,.$$

For any $0 < \delta < 1$, we denote by $B_{\mathbb{S}_+}(\delta)$ the open ball in $\mathbb{R}^{\mathbb{S}_+}$ of radius δ centered at 0 and by $B^s_\perp(\delta)$, $s \geq 0$, the corresponding one in $H^s_\perp(\mathbb{T}_1)$ where we also write $B_\perp(\delta)$ for $B^0_\perp(\delta)$. These balls are used to define the following open neighborhoods in $\mathcal{E}_s$, $s \in \mathbb{N}$,

$$\mathcal{V}^s(\delta) := \mathbb{T}_1^{\mathbb{S}_+} \times B_{\mathbb{S}_+}(\delta) \times B^s_\perp(\delta)\,, \qquad \mathcal{V}(\delta) \equiv \mathcal{V}^0(\delta)\,, \qquad 0 < \delta < 1\,. \tag{1.24}$$

The space of bounded linear operators between Banach spaces X_1, X_2 is denoted by $\mathcal{B}(X_1, X_2)$ and endowed with the operator norm. For two linear operators A, B we denote by $[A, B]$ their commutator, $[A, B] := AB - BA$. We denote $A^\top$ the transposed with respect to the scalar product (1.19).
Throughout the paper, $\Omega \subseteq \mathbb{R}^{\mathbb{S}_+}$ denotes a parameter set of frequency vectors. Given any function $f : \Omega \to X$, we denote by $\Delta_\omega f$ the difference function

$$\Delta_\omega f : \Omega \times \Omega \to X\,, \quad (\omega_1, \omega_2) \mapsto f(\omega_1) - f(\omega_2)\,.$$

Acknowledgements. Part of this work was written during the stay of M. Berti at FIM. We thank FIM for the kind hospitality. In addition, the research was partially supported by PRIN 2015KB9WPT005 (M.B.), by the Swiss National Science Foundation (T.K., R.M.), and by INDAM-GNFM (R.M.).

2 Preliminaries

2.1 Function spaces and linear operators

In the paper we consider real or complex functions $u(\varphi, x; \omega)$, $(\varphi, x) \in \mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1$, depending on a parameter $\omega \in \Omega$ in a Lipschitz way, where Ω is a subset of $\mathbb{R}^{\mathbb{S}_+}$. Given $0 < \gamma < 1$ and $s \geq 0$, we define the norm

$$\begin{aligned} &\|u\|^{\mathrm{Lip}(\gamma)}_{s,\Omega} := \|u\|^{\mathrm{Lip}(\gamma)}_s := \|u\|^{\sup}_s + \gamma \|u\|^{\mathrm{lip}}_s \\ &\|u\|^{\sup} := \sup_{\omega \in \Omega} \|u(\omega)\|_s, \quad \|u\|^{\mathrm{lip}}_s := \sup_{\omega_1, \omega_2 \in \Omega\,, \omega_1 \neq \omega_2} \frac{\|u(\omega_1) - u(\omega_2)\|_s}{|\omega_1 - \omega_2|} \end{aligned} \tag{2.1}$$

where $\|\ \|_s$ is the norm of the Sobolev space H^s defined in (1.15). For a function $u : \Omega \to \mathbb{C}$, the sup norm and the Lipschitz semi-norm are denoted by $|u|^{\sup}$ and, respectively $|u|^{\mathrm{lip}}$. Correpondingly, we write $|u|^{\mathrm{Lip}(\gamma)} := |u|^{\sup} + \gamma |u|^{\mathrm{lip}}$.

By Π_N, $N \in \mathbb{N}_+$, we denote the *smoothing* operators on H^s,

$$(\Pi_N u)(\varphi, x) := \sum_{\langle \ell, j \rangle \leq N} u_{\ell,j} e^{\mathrm{i}(\ell \cdot \varphi + 2\pi j x)}\,, \qquad \Pi_N^\perp := \mathrm{Id} - \Pi_N\,. \tag{2.2}$$

They satisfy, for any $\alpha \geq 0$, $s \in \mathbb{R}$, the estimates

$$\|\Pi_N u\|^{\mathrm{Lip}(\gamma)}_s \leq N^\alpha \|u\|^{\mathrm{Lip}(\gamma)}_{s-\alpha}\,, \qquad \|\Pi_N^\perp u\|^{\mathrm{Lip}(\gamma)}_s \leq N^{-\alpha} \|u\|^{\mathrm{Lip}(\gamma)}_{s+\alpha}\,. \tag{2.3}$$

Furthermore the following interpolation inequalities hold: for any $0 \leq s_1 < s_2$ and $0 < \theta < 1$,

$$\|u\|^{\mathrm{Lip}(\gamma)}_{\theta s_1 + (1-\theta) s_2} \leq 2 (\|u\|^{\mathrm{Lip}(\gamma)}_{s_1})^\theta (\|u\|^{\mathrm{Lip}(\gamma)}_{s_2})^{1-\theta}\,. \tag{2.4}$$

Multiplication and composition with Sobolev functions satisfy the following tame estimates.

Lemma 2.1. (Product and composition) *(i) For any $s \geq s_0 = [(|\mathbb{S}_+| + 1)/2] + 1$*

$$\|uv\|^{\mathrm{Lip}(\gamma)}_s \leq C(s) \|u\|^{\mathrm{Lip}(\gamma)}_s \|v\|^{\mathrm{Lip}(\gamma)}_{s_0} + C(s_0) \|u\|^{\mathrm{Lip}(\gamma)}_{s_0} \|v\|^{\mathrm{Lip}(\gamma)}_s\,. \tag{2.5}$$

(ii) Let $\beta(\cdot,\cdot;\omega):\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1\to\mathbb{R}$ with $\|\beta\|^{\mathrm{Lip}(\gamma)}_{2s_0+2}\le\delta(s_0)$ small enough. Then the composition operator $\mathcal{B}:u\mapsto\mathcal{B}u$, $(\mathcal{B}u)(\varphi,x):=u(\varphi,x+\beta(\varphi,x))$ satisfies, for any $s\ge s_0+1$,

$$\|\mathcal{B}u\|^{\mathrm{Lip}(\gamma)}_s\lesssim_s\|u\|^{\mathrm{Lip}(\gamma)}_{s+1}+\|\beta\|^{\mathrm{Lip}(\gamma)}_s\|u\|^{\mathrm{Lip}(\gamma)}_{s_0+2}\,. \tag{2.6}$$

The function $\breve{\beta}$, obtained by solving $y=x+\beta(\varphi,x)$ for x, $x=y+\breve{\beta}(\varphi,y)$, satisfies

$$\|\breve{\beta}\|^{\mathrm{Lip}(\gamma)}_s\lesssim_s\|\beta\|^{\mathrm{Lip}(\gamma)}_{s+1}\,,\quad\forall s\ge s_0\,. \tag{2.7}$$

(iii) Let $\alpha(\cdot;\omega):\mathbb{T}^{\mathbb{S}_+}\to\mathbb{R}$ with $\|\alpha\|^{\mathrm{Lip}(\gamma)}_{2s_0+2}\le\delta(s_0)$ small enough. Then the composition operator $\mathcal{A}:u\mapsto\mathcal{A}u$, $(\mathcal{A}u)(\varphi,x):=u(\varphi+\alpha(\varphi)\omega,x)$ satisfies, for any $s\ge s_0+1$,

$$\|\mathcal{A}u\|^{\mathrm{Lip}(\gamma)}_s\lesssim_s\|u\|^{\mathrm{Lip}(\gamma)}_{s+1}+\|\alpha\|^{\mathrm{Lip}(\gamma)}_s\|u\|^{\mathrm{Lip}(\gamma)}_{s_0+2}\,. \tag{2.8}$$

The function $\breve{\alpha}$, obtained by solving $\vartheta=\varphi+\alpha(\varphi)\omega$ for φ, $\varphi=\vartheta+\breve{\alpha}(\vartheta)\omega$, satisfies

$$\|\breve{\alpha}\|^{\mathrm{Lip}(\gamma)}_s\lesssim_s\|\alpha\|^{\mathrm{Lip}(\gamma)}_{s+1}\,,\quad\forall s\ge s_0\,. \tag{2.9}$$

Proof. Item (i) follows from (2.72) in [9] and (ii)-(iii) follow from [9, Lemma 2.30]. □

If ω is diophantine, namely

$$|\omega\cdot\ell|\ge\frac{\gamma}{|\ell|^\tau}\,,\quad\forall\ell\in\mathbb{Z}^{\mathbb{S}_+}\setminus\{0\}\,,$$

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

$$(\omega\cdot\partial_\varphi)^{-1}u=\sum_{j\in\mathbb{Z},\ell\in\mathbb{Z}^{\mathbb{S}_+}\setminus\{0\}}\frac{u_{\ell,j}}{\mathrm{i}\omega\cdot\ell}e^{\mathrm{i}(\ell\cdot\varphi+2\pi jx)}\,,$$

and it satisfies the estimate (cf. e.g. [8, Lemma 2.2])

$$\|(\omega\cdot\partial_\varphi)^{-1}u\|^{\mathrm{Lip}(\gamma)}_s\le C\gamma^{-1}\|u\|^{\mathrm{Lip}(\gamma)}_{s+2\tau+1}\,. \tag{2.10}$$

We also record Moser's 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 φ and x play the same role, we state it for the Sobolev space $H^s(\mathbb{T}^d)$, (cf. e.g. [9, Lemma 2.31]).

Lemma 2.2. (Composition operator) *Let $f\in\mathcal{C}^\infty(\mathbb{T}^d\times\mathbb{R}^n,\mathbb{C})$. If $v(\cdot;\omega)\in H^s(\mathbb{T}^d,\mathbb{R}^n)$, $\omega\in\Omega$, is a family of Sobolev functions satisfying $\|v\|^{\mathrm{Lip}(\gamma)}_{s_0(d)}\le 1$ where $s_0(d)>d/2$, then, for any $s\ge s_0(d)$,*

$$\|\mathtt{f}(v)\|^{\mathrm{Lip}(\gamma)}_s\le C(s,f)(1+\|v\|^{\mathrm{Lip}(\gamma)}_s)\,. \tag{2.11}$$

Moreover, if $f(\varphi,x,0)=0$, then $\|\mathtt{f}(v)\|^{\mathrm{Lip}(\gamma)}_s\le C(s,f)\|v\|^{\mathrm{Lip}(\gamma)}_s$.

Linear operators. Throughout the paper we consider φ-dependent families of linear operators $A:\mathbb{T}^{\mathbb{S}_+}\to\mathcal{L}(L^2(\mathbb{T}_1,\mathbb{C}))$, $\varphi\mapsto A(\varphi)$, acting on complex valued functions $u(x)$ of the space variable x. 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. as an element in $\mathcal{L}(L^2(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1,\mathbb{C}))$ defined by

$$A[u](\varphi,x)\equiv(Au)(\varphi,x):=(A(\varphi)u(\varphi,\cdot))(x)\,. \tag{2.12}$$

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

When u in (2.12) is expanded in its Fourier series,

$$u(\varphi,x)=\sum_{j\in\mathbb{Z}}u_j(\varphi)e^{2\pi \mathrm{i} jx}=\sum_{j\in\mathbb{Z},\ell\in\mathbb{Z}^{\mathbb{S}_+}}u_{\ell,j}e^{\mathrm{i}(\ell\cdot\varphi+2\pi jx)}\,, \tag{2.13}$$

one obtains

$$(Au)(\varphi,x)=\sum_{j,j'\in\mathbb{Z}}A_j^{j'}(\varphi)u_{j'}(\varphi)e^{\mathrm{i}2\pi jx}=\sum_{j\in\mathbb{Z},\ell\in\mathbb{Z}^{\mathbb{S}_+}}\ \sum_{j'\in\mathbb{Z},\ell'\in\mathbb{Z}^{\mathbb{S}_+}}A_j^{j'}(\ell-\ell')u_{\ell',j'}e^{\mathrm{i}(\ell\cdot\varphi+2\pi jx)}\,. \tag{2.14}$$

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

Definition 2.3. *Given a linear operator A as in* (2.14) *we define the following operators:*

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*

$$(\Pi_N A)_j^{j'}(\ell-\ell'):=\begin{cases}A_j^{j'}(\ell-\ell') & \text{if}\quad \langle\ell-\ell'\rangle\le N\\ 0 & \text{otherwise}\,.\end{cases} \tag{2.15}$$

3. $\langle\partial_\varphi\rangle^b A$, $b\in\mathbb{R}$, *whose matrix elements are* $\langle\ell-\ell'\rangle^b A_j^{j'}(\ell-\ell')$.
4. $\partial_{\varphi_m}A(\varphi)=[\partial_{\varphi_m},A]$ (DIFFERENTIATED OPERATOR) *whose matrix elements are* $\mathrm{i}(\ell_m-\ell'_m)A_j^{j'}(\ell-\ell')$.

Definition 2.4. (Hamiltonian and symplectic operators) *(i) A φ-dependent family of linear operators $X(\varphi)$, $\varphi\in\mathbb{T}^{\mathbb{S}_+}$, densily defined in $L^2_0(\mathbb{T}_1)$, is* HAMILTONIAN *if $X(\varphi)=\partial_x G(\varphi)$ for some real linear operator $G(\varphi)$ which is self-adjoint with respect to the L^2-inner product. We also say that $\omega\cdot\partial_\varphi-\partial_x G(\varphi)$ is Hamiltonian.*
(ii) A φ-dependent family of linear operators $A(\varphi):L^2_0(\mathbb{T}_1)\to L^2_0(\mathbb{T}_1)$, $\forall\varphi\in\mathbb{T}^{\mathbb{S}_+}$, is SYMPLECTIC *if*

$$\mathcal{W}_{L^2_0}(A(\varphi)u,A(\varphi)v)=\mathcal{W}_{L^2_0}(u,v)\,,\quad \forall u,v\in L^2_0(\mathbb{T}_1)\,,$$

where the symplectic 2-form $\mathcal{W}_{L^2_0}$ is defined in (1.7).

Under a φ-dependent family of symplectic transformations $\Phi(\varphi)$, $\varphi\in\mathbb{T}^{\mathbb{S}_+}$, the linear Hamiltonian operator $\omega\cdot\partial_\varphi-\partial_x G(\varphi)$ transforms into another Hamiltonian one.

Lemma 2.5. *A family of operators $R(\varphi)$, $\varphi\in\mathbb{T}^{\mathbb{S}_+}$, expanded as $R(\varphi)=\sum_{\ell\in\mathbb{Z}^{\mathbb{S}_+}}R(\ell)e^{\mathrm{i}\ell\cdot\varphi}$, is*
(i) SELF-ADJOINT *if and only if $\overline{R_j^{j'}(\ell)}=R_{j'}^{j}(-\ell)$;*
(ii) REAL *if and only if $\overline{R_{j'}^{j}(\ell)}=R_{-j'}^{-j}(-\ell)$;*
(iii) REAL AND SELF-ADJOINT *if and only if $R_j^{j'}(\ell)=R_{-j'}^{-j}(\ell)$.*

Lemma 2.6. *Let $X:H_0^{s+3}(\mathbb{T}_1)\to H_0^s(\mathbb{T}_1)$ be a linear Hamiltonian vector field of the form*

$$X=\sum_{k=0}^{2}a_{3-k}(x)\partial_x^{3-k}+\text{bounded operator} \tag{2.16}$$

where $a_{3-k}\in\mathcal{C}^\infty(\mathbb{T}_1)$. Then $a_2=2(a_3)_x$.

Proof. Since X is a linear Hamiltonian vector field it has the form $X=\partial_x\mathcal{A}$ where $\mathcal{A}$ is a densely defined operator on $L^2_0(\mathbb{T}_1)$ satisfying $\mathcal{A}=\mathcal{A}^\top$. Therefore, using (2.16),

$$\begin{aligned}\mathcal{A}&=\partial_x^{-1}X=a_3(x)\partial_{xx}+\big(-(a_3)_x+a_2\big)\partial_x+\dots\\ \mathcal{A}^\top&=-X^\top\partial_x^{-1}=a_3(x)\partial_{xx}+\big(3(a_3)_x-a_2\big)\partial_x+\dots\,.\end{aligned}$$

The identity $\mathcal{A}=\mathcal{A}^\top$ implies that $a_2=2(a_3)_x$. $\square$

2.2 Pseudo-differential operators

In this section we recall properties of pseudo-differential operators on the torus used in this paper, following [9]. Note however that $x \in \mathbb{T}_1$ and not in $\mathbb{R}/(2\pi\mathbb{Z})$.

Definition 2.7. *We say that $a : \mathbb{T}_1 \times \mathbb{R} \to \mathbb{C}$ is a symbol of order $m \in \mathbb{R}$ if, for any $\alpha, \beta \in \mathbb{N}$,*

$$\left|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)\right| \leq C_{\alpha,\beta} \langle \xi \rangle^{m-\beta}, \quad \forall (x,\xi) \in \mathbb{T}_1 \times \mathbb{R}. \tag{2.17}$$

The set of such symbols is denoted by S^m. Given $a \in S^m$, we denote by A the operator, which maps a one periodic function $u(x) = \sum_{j\in\mathbb{Z}} u_j e^{\mathrm{i}jx}$ to

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

The operator A is referred to as the PSEUDO-DIFFERENTIAL OPERATOR (ΨDO) *of order m, associated to the symbol a, and is also denoted by $Op(a)$ or $a(x,D)$ where $D = \frac{1}{\mathrm{i}}\partial_x$. Furthermore we denote by OPS^m the set of pseudo-differential operators $a(x,D)$ with $a(x,\xi) \in S^m$ and set $OPS^{-\infty} := \cap_{m\in\mathbb{R}} OPS^m$.*

When the symbol a is independent of ξ, the operator $A = \mathrm{Op}(a)$ is the multiplication operator by the function $a(x)$, i.e., $A : u(x) \mapsto a(x)u(x)$ and we also write a for A. More generally, we consider symbols $a(\varphi, x, \xi; \omega)$, depending in addition on the variable $\varphi \in \mathbb{T}^{\mathbb{S}_+}$ and the parameter ω, where a is $\mathcal{C}^\infty$ in φ and Lipschitz continuous with respect to ω. By a slight abuse of notation, we denote the class of such symbols of order m also by S^m. Alternatively, we denote A by $A(\varphi)$ or $\mathrm{Op}(a(\varphi,\cdot))$.

Given an even cut off function $\chi_0 \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$, satisfying

$$0 \leq \chi_0 \leq 1, \qquad \chi_0(\xi) = 0, \quad \forall |\xi| < \frac{1}{2}, \qquad \chi_0(\xi) = 1, \quad \forall |\xi| \geq \frac{2}{3}, \tag{2.18}$$

we define, for any $m \in \mathbb{Z}$, $\partial_x^m = \mathrm{Op}(\chi_0(\xi)(\mathrm{i}2\pi\xi)^m)$, so that

$$\partial_x^m[e^{\mathrm{i}2\pi jx}] = (\mathrm{i}2\pi j)^m e^{\mathrm{i}2\pi jx}, \quad j \in \mathbb{Z}\setminus\{0\}, \quad \partial_x^m[1] = 0. \tag{2.19}$$

Note that $\partial_x^0[u](x) = u(x) - u_0$, hence ∂_x^0 is not the identity operator.

Now we recall the norm of a symbol $a(\varphi,x,\xi;\omega)$ in S^m, introduced in [9, Definition 2.11], which controls the regularity in (φ,x) and the decay in ξ of a and its derivatives $\partial_\xi^\beta a \in S^{m-\beta}$, $0 \leq \beta \leq \alpha$, in the Sobolev norm $\|\ \|_s$. By a slight abuse of terminology, we refer to it as the norm of the corresponding pseudo-differential operator. Unlike [9] we consider the difference quotient instead of the derivative with respect to ω, and write $|\ |_{m,s,\alpha}^{1,\gamma}$ instead of $|\ |_{m,s,\alpha}^{\mathrm{Lip}(\gamma)}$.

Definition 2.8. *Let $A(\omega) := a(\varphi,x,D;\omega) \in OPS^m$ be a family of pseudo-differential operators with symbols $a(\varphi,x,\xi;\omega) \in S^m$, $m \in \mathbb{R}$. For $\gamma \in (0,1)$, $\alpha \in \mathbb{N}$, $s \geq 0$, we define the* WEIGHTED ΨDO NORM *of A as*

$$|A|_{m,s,\alpha}^{\mathrm{Lip}(\gamma)} := \sup_{\omega\in\Omega} |A(\omega)|_{m,s,\alpha} + \gamma \sup_{\substack{\omega_1,\omega_2\in\Omega \\ \omega_1\neq\omega_2}} \frac{|A(\omega_1) - A(\omega_2)|_{m,s,\alpha}}{|\omega_1 - \omega_2|}$$

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

Notice that for any $s \leq s'$, $\alpha \leq \alpha'$, and $m \leq m'$,

$$|\cdot|_{m,s,\alpha}^{\mathrm{Lip}(\gamma)} \leq |\cdot|_{m,s',\alpha}^{\mathrm{Lip}(\gamma)}, \qquad |\cdot|_{m,s,\alpha}^{\mathrm{Lip}(\gamma)} \leq |\cdot|_{m,s,\alpha'}^{\mathrm{Lip}(\gamma)}, \qquad |\cdot|_{m',s,\alpha}^{\mathrm{Lip}(\gamma)} \leq |\cdot|_{m,s,\alpha}^{\mathrm{Lip}(\gamma)}. \tag{2.20}$$

For a Fourier multiplier $g(D;\omega)$ with symbol $g \in S^m$, one has

$$|\mathrm{Op}(g)|_{m,s,\alpha}^{\mathrm{Lip}(\gamma)} = |\mathrm{Op}(g)|_{m,0,\alpha}^{\mathrm{Lip}(\gamma)} \leq C(m,\alpha,g), \quad \forall s \geq 0, \tag{2.21}$$

and, for a function $a(\varphi,x;\omega)$,

$$|\mathrm{Op}(a)|_{0,s,\alpha}^{\mathrm{Lip}(\gamma)} = |\mathrm{Op}(a)|_{0,s,0}^{\mathrm{Lip}(\gamma)} \lesssim \|a\|_s^{\mathrm{Lip}(\gamma)}. \tag{2.22}$$

Composition. If $A = a(\varphi, x, D; \omega) \in OPS^m$, $B = b(\varphi, x, D; \omega) \in OPS^{m'}$ then the composition $AB := A \circ B$ is a pseudo-differential operator with a symbol $\sigma_{AB}(\varphi, x, \xi; \omega)$ in $S^{m+m'}$ which, for any $N \geq 0$, admits the asymptotic expansion

$$\sigma_{AB}(\varphi, x, \xi; \omega) = \sum_{\beta=0}^{N} \frac{1}{\mathrm{i}^\beta \beta!} \partial_\xi^\beta a(\varphi, x, \xi; \omega)\, \partial_x^\beta b(\varphi, x, \xi; \omega) + r_N(\varphi, x, \xi; \omega) \tag{2.23}$$

with remainder $r_N \in S^{m+m'-N-1}$. We record the following tame estimate for the composition of two pseudo-differential operators, proved in [9, Lemma 2.13].

Lemma 2.9. (Composition) *Let $A = a(\varphi, x, D; \omega)$, $B = b(\varphi, x, D; \omega)$ be pseudo-differential operators with symbols $a(\varphi, x, \xi; \omega) \in S^m$, $b(\varphi, x, \xi; \omega) \in S^{m'}$, $m, m' \in \mathbb{R}$. Then $A(\omega) \circ B(\omega)$ is the pseudo-differential operator of order $m + m'$, associated to the symbol $\sigma_{AB}(\varphi, x, \xi; \omega)$ which satisfies, for any $\alpha \in \mathbb{N}$, $s \geq s_0$,*

$$|AB|^{\mathrm{Lip}(\gamma)}_{m+m',s,\alpha} \lesssim_{m,\alpha} C(s)|A|^{\mathrm{Lip}(\gamma)}_{m,s,\alpha}|B|^{\mathrm{Lip}(\gamma)}_{m',s_0+\alpha+|m|,\alpha} + C(s_0)|A|^{\mathrm{Lip}(\gamma)}_{m,s_0,\alpha}|B|^{\mathrm{Lip}(\gamma)}_{m',s+\alpha+|m|,\alpha} \,. \tag{2.24}$$

Moreover, for any integer $N \geq 1$, the remainder $R_N := \mathrm{Op}(r_N)$ in (2.23) satisfies

$$\begin{aligned} |R_N|^{\mathrm{Lip}(\gamma)}_{m+m'-N-1,s,\alpha} \lesssim_{m,N,\alpha} & C(s)|A|^{\mathrm{Lip}(\gamma)}_{m,s,N+1+\alpha}|B|^{\mathrm{Lip}(\gamma)}_{m',s_0+2(N+1)+|m|+\alpha,\alpha} \\ & + C(s_0)|A|^{\mathrm{Lip}(\gamma)}_{m,s_0,N+1+\alpha}|B|^{\mathrm{Lip}(\gamma)}_{m',s+2(N+1)+|m|+\alpha,\alpha}. \end{aligned} \tag{2.25}$$

By (2.23) the commutator $[A, B]$ of two pseudo-differential operators $A = a(x, D) \in OPS^m$ and $B = b(x, D) \in OPS^{m'}$ is a pseudo-differential operator of order $m + m' - 1$, and Lemma 2.9 then leads to the following lemma, cf. [9, Lemma 2.15].

Lemma 2.10. (Commutator) *If $A = a(\varphi, x, D; \omega) \in OPS^m$ and $B = b(\varphi, x, D; \omega) \in OPS^{m'}$, $m, m' \in \mathbb{R}$, then the commutator $[A, B] := AB - BA$ is the pseudo-differential operator of order $m + m' - 1$ associated to the symbol $\sigma_{AB}(\varphi, x, \xi; \omega) - \sigma_{BA}(\varphi, x, \xi; \omega) \in S^{m+m'-1}$ which for any $\alpha \in \mathbb{N}$ and $s \geq s_0$ satisfies*

$$\begin{aligned} |[A, B]|^{\mathrm{Lip}(\gamma)}_{m+m'-1,s,\alpha} \lesssim_{m,m',\alpha} & C(s)|A|^{\mathrm{Lip}(\gamma)}_{m,s+2+|m'|+\alpha,\alpha+1}|B|^{\mathrm{Lip}(\gamma)}_{m',s_0+2+|m|+\alpha,\alpha+1} \\ & + C(s_0)|A|^{\mathrm{Lip}(\gamma)}_{m,s_0+2+|m'|+\alpha,\alpha+1}|B|^{\mathrm{Lip}(\gamma)}_{m',s+2+|m|+\alpha,\alpha+1} \,. \end{aligned} \tag{2.26}$$

In the case of operators of the special form $a\partial_x^m$, Lemma 2.9 and Lemma 2.10 simplify as follows:

Lemma 2.11. (Composition and commutator of homogeneous symbols) *Let $A = a\partial_x^m$, $B = b\partial_x^{m'}$ where $m, m' \in \mathbb{Z}$ and $a(\varphi, x; \omega)$, $b(\varphi, x; \omega)$ are $\mathcal{C}^\infty-$smooth functions with respect to (φ, x) and Lipschitz with respect to $\omega \in \Omega$. Then there exist combinatorial constants $K_{n,m} \in \mathbb{R}$, $0 \leq n \leq N$, with $K_{0,m} = 1$ and $K_{1,m} = m$ so that the following holds:*
(i) For any $N \in \mathbb{N}$, the composition $A \circ B$ is in $OPS^{m+m'}$ and admits the asymptotic expansion

$$A \circ B = \sum_{n=0}^{N} K_{n,m}\, a\, (\partial_x^n b) \partial_x^{m+m'-n} + \mathcal{R}_N(a, b)$$

where the remainder $\mathcal{R}_N(a, b)$ is in $OPS^{m+m'-N-1}$. Furthermore there is a constant $\sigma_N(m) > 0$ so that, for any $s \geq s_0$, $\alpha \in \mathbb{N}$,

$$|\mathcal{R}_N(a, b)|^{\mathrm{Lip}(\gamma)}_{m+m'-N-1,s,\alpha} \lesssim_{m,m',s,N,\alpha} \|a\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_N(m)}\|b\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_N(m)} + \|a\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_N(m)}\|b\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_N(m)} \,.$$

(ii) For any $N \in \mathbb{N}_+$, the commutator $[A, B]$ is in $OPS^{m+m'-1}$ and admits the asymptotic expansion

$$[A, B] = \sum_{n=1}^{N} \big(K_{n,m} a(\partial_x^n b) - K_{n,m'}(\partial_x^n a) b\big) \partial_x^{m+m'-n} + \mathcal{Q}_N(a, b)$$

where the remainder $\mathcal{Q}_N(a, b)$ is in $OPS^{m+m'-N-1}$. Furthermore, there is a constant $\sigma_N(m, m') > 0$ so that, for any $s \geq s_0$, $\alpha \in \mathbb{N}$,

$$|\mathcal{Q}_N(a, b)|^{\mathrm{Lip}(\gamma)}_{m+m'-N-1,s,\alpha} \lesssim_{m,m',s,N,\alpha} \|a\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_N(m,m')}\|b\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_N(m,m')} + \|a\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_N(m,m')}\|b\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_N(m,m')} \,.$$

Proof. See formula (2.23) and Lemma 2.9. $\square$

We finally give the following result on the exponential of a pseudo-differential operator of order 0.

Lemma 2.12. (Exponential map) *If $A := \mathrm{Op}(a(\varphi, x, \xi; \omega))$ is in OPS^0, then $\sum_{k\geq 0} \frac{1}{k!}\sigma_{A^k}(\varphi, x, \xi; \omega)$ is a symbol of order 0 and hence the corresponding pseudo-differential operator, denoted by $\Phi = \exp(A)$, is in OPS^0, and for any $s \geq s_0$, $\alpha \in \mathbb{N}$, there is a constant $C(s, \alpha) > 0$ so that*

$$|\Phi - \mathrm{Id}|^{\mathrm{Lip}(\gamma)}_{0,s,\alpha} \leq |A|^{\mathrm{Lip}(\gamma)}_{0,s+\alpha,\alpha} \exp\big(C(s,\alpha)|A|^{\mathrm{Lip}(\gamma)}_{0,s_0+\alpha,\alpha}\big)\,. \tag{2.27}$$

Proof. Iterating (2.24), for any $s \geq s_0$, $\alpha \in \mathbb{N}$, there is a constant $C(s, \alpha) > 0$ such that

$$|A^k|^{\mathrm{Lip}(\gamma)}_{0,s,\alpha} \leq C(s,\alpha)^{k-1}(|A|^{\mathrm{Lip}(\gamma)}_{0,s_0+\alpha,\alpha})^{k-1}|A|^{\mathrm{Lip}(\gamma)}_{0,s+\alpha,\alpha}\,, \quad \forall k \geq 1\,. \tag{2.28}$$

Therefore

$$\begin{aligned}|\Phi - \mathrm{Id}|^{\mathrm{Lip}(\gamma)}_{0,s,\alpha} &\leq \sum_{k\geq 1}\frac{1}{k!}|A^k|^{\mathrm{Lip}(\gamma)}_{0,s,\alpha} \overset{(2.28)}{\leq} |A|^{\mathrm{Lip}(\gamma)}_{0,s+\alpha,\alpha}\sum_{k\geq 1}\frac{1}{k!}C(s,\alpha)^{k-1}(|A|^{\mathrm{Lip}(\gamma)}_{0,s_0+\alpha,\alpha})^{k-1} \\ &\leq |A|^{\mathrm{Lip}(\gamma)}_{0,s+\alpha,\alpha}\exp\big(C(s,\alpha)|A|^{\mathrm{Lip}(\gamma)}_{0,s_0+\alpha,\alpha}\big)\end{aligned}.$$

This shows that $\sum_{k\geq 0}\frac{1}{k!}\sigma_{A^k}(\varphi, x, \xi; \omega)$ is a symbol in S^0 and that the estimate (2.27) holds. $\square$

2.3 Lip(γ)-tame and modulo-tame operators

In this section we recall the notion and the main properties of Lip(γ)-σ-tame and Lip(γ)-modulo-tame operators. We refer to [9, Section 2.2] where this notion was introduced, with the only difference that here we consider difference quotients instead of first order derivatives with respect to the parameter ω.

Definition 2.13. (**Lip(γ)-σ-tame**) *Let $\sigma \geq 0$. A linear operator $A := A(\omega)$ as in (2.12) is Lip(γ)-σ-tame if there exist $S > s_1 \geq s_0$ and a non-decreasing function $[s_1, S] \to [0, +\infty)$, $s \mapsto \mathfrak{M}_A(s)$, so that, for any $s_1 \leq s \leq S$ and $u \in H^{s+\sigma}$,*

$$\sup_{\omega\in\Omega}\|A(\omega)u\|_s + \gamma \sup_{\substack{\omega_1,\omega_2\in\Omega\\ \omega_1\neq\omega_2}}\Big\|\frac{A(\omega_1) - A(\omega_2)}{|\omega_1 - \omega_2|}u\Big\|_s \leq \mathfrak{M}_A(s_1)\|u\|_{s+\sigma} + \mathfrak{M}_A(s)\|u\|_{s_1+\sigma}\,. \tag{2.29}$$

When σ is zero, we simply write Lip(γ)-tame instead of Lip(γ)-0-tame. We say that $\mathfrak{M}_A(s)$ is a TAME CONSTANT *of the operator A. Note that $\mathfrak{M}_A(s)$ is not uniquely determined and that it may also depend on the "loss of derivatives" σ. We will not indicate this dependence.*

Representing the operator A by its matrix elements $\big(A_j^{j'}(\ell - \ell')\big)_{\ell,\ell'\in\mathbb{Z}^{\mathbb{S}_+}, j,j'\in\mathbb{Z}}$ as in (2.14), we have, for all $j' \in \mathbb{Z}$, $\ell' \in \mathbb{Z}^{\mathbb{S}_+}$, for all $\omega_1, \omega_2 \in \Omega$, $\omega_1 \neq \omega_2$,

$$\sum\nolimits_{\ell,j}\langle \ell, j\rangle^{2s_1}\Big(\big|A_j^{j'}(\ell-\ell')\big|^2 + \gamma^2\Big|\frac{\Delta_\omega A_j^{j'}(\ell-\ell')}{|\omega_1-\omega_2|}\Big|^2\Big) \lesssim \big(\mathfrak{M}_A(s_1)\big)^2\langle \ell', j'\rangle^{2(s_1+\sigma)} \tag{2.30}$$

where we recall that $\Delta_\omega f = f(\omega_1) - f(\omega_2)$.

Lemma 2.14. (Composition) *Let A, B be, respectively, Lip(γ)-σ_A-tame and Lip(γ)-σ_B-tame operators with tame constants $\mathfrak{M}_A(s)$ and $\mathfrak{M}_B(s)$. Then the composition $A \circ B$ is Lip(γ)-($\sigma_A + \sigma_B$)-tame with a tame constant satisfying*

$$\mathfrak{M}_{AB}(s) \lesssim \mathfrak{M}_A(s)\mathfrak{M}_B(s_1 + \sigma_A) + \mathfrak{M}_A(s_1)\mathfrak{M}_B(s + \sigma_A)\,.$$

Proof. See [9, Lemma 2.20]. $\square$

We now discuss the action of a Lip(γ)-σ-tame operator $A(\omega)$ on a family of Sobolev functions $u(\omega) \in H^s$.

Lemma 2.15. (Action on H^s) *Let $A := A(\omega)$ be a* Lip(γ)*-σ-tame operator with tame constant $\mathfrak{M}_A(s)$. Then, for any family of Sobolev functions $u := u(\omega) \in H^{s+\sigma}$, Lipschitz with respect to ω, one has*

$$\|Au\|_s^{\mathrm{Lip}(\gamma)} \lesssim \mathfrak{M}_A(s_1)\|u\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \mathfrak{M}_A(s)\|u\|_{s_1+\sigma}^{\mathrm{Lip}(\gamma)}\,.$$

Proof. See [9, Lemma 2.22]. ☐

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

Lemma 2.16. *Let $a(\varphi, x, \xi; \omega) \in S^0$ be a family of symbols that are Lipschitz with respect to ω. If $A = a(\varphi, x, D; \omega)$ satisfies $|A|_{0,s,0}^{\mathrm{Lip}(\gamma)} < +\infty$, $s \geq s_0$, then A is* Lip(γ)*-tame with a tame constant satisfying*

$$\mathfrak{M}_A(s) \leq C(s)|A|_{0,s,0}^{\mathrm{Lip}(\gamma)}\,. \tag{2.31}$$

As a consequence

$$\|Au\|_s^{\mathrm{Lip}(\gamma)} \leq C(s_0)|A|_{0,s_0,0}^{\mathrm{Lip}(\gamma)}\|u\|_s^{\mathrm{Lip}(\gamma)} + C(s)|A|_{0,s,0}^{\mathrm{Lip}(\gamma)}\|u\|_{s_0}^{\mathrm{Lip}(\gamma)}\,. \tag{2.32}$$

Proof. See [9, Lemma 2.21] for the proof of (2.31). The estimate (2.32) then follows from Lemma 2.15. ☐

In the KAM reducibility scheme of Section 7, we need to consider Lip(γ)-tame operators A which satisfy a stronger condition, referred to Lip(γ)-modulo-tame operators.

Definition 2.17. (Lip(γ)-modulo-tame) *Let $S > s_1 \geq s_0$. A linear operator $A := A(\omega)$ as in* (2.12) *is* Lip(γ)*-modulo-tame if there exists a non-decreasing function $[s_1, S] \to [0, +\infty)$, $s \mapsto \mathfrak{M}_A^\sharp(s)$, such that the majorant operators $|A(\omega)|$ (see Definition 2.3) satisfy, for any $s_1 \leq s \leq S$ and $u \in H^s$,*

$$\sup_{\omega\in\Omega}\|\,|A(\omega)|u\|_s + \gamma \sup_{\substack{\omega_1,\omega_2\in\Omega\\ \omega_1\neq\omega_2}}\Big\|\,\frac{|A(\omega_1) - A(\omega_2)|}{|\omega_1 - \omega_2|}u\Big\|_s \leq \mathfrak{M}_A^\sharp(s_1)\|u\|_s + \mathfrak{M}_A^\sharp(s)\|u\|_{s_1}\,. \tag{2.33}$$

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

If A, B are Lip(γ)-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 2.18. *An operator A that is* Lip(γ)*-modulo-tame with modulo-tame constant $\mathfrak{M}_A^\sharp(s)$ is also* Lip(γ)*-tame and $\mathfrak{M}_A^\sharp(s)$ is a tame constant for A.*

Proof. See [9, Lemma 2.24]. ☐

The class of operators which are Lip(γ)-modulo-tame is closed under sum and composition.

Lemma 2.19. (Sum and composition) *Let A, B be* Lip(γ)*-modulo-tame operators with modulo-tame constants respectively $\mathfrak{M}_A^\sharp(s)$ and $\mathfrak{M}_B^\sharp(s)$. Then $A + B$ is* Lip(γ)*-modulo-tame with a modulo-tame constant satisfying*

$$\mathfrak{M}_{A+B}^\sharp(s) \leq \mathfrak{M}_A^\sharp(s) + \mathfrak{M}_B^\sharp(s)\,. \tag{2.34}$$

The composed operator $A \circ B$ is Lip(γ)*-modulo-tame with a modulo-tame constant satisfying*

$$\mathfrak{M}_{AB}^\sharp(s) \leq C\left(\mathfrak{M}_A^\sharp(s)\mathfrak{M}_B^\sharp(s_1) + \mathfrak{M}_A^\sharp(s_1)\mathfrak{M}_B^\sharp(s)\right) \tag{2.35}$$

where $C \geq 1$ is a constant. Assume in addition that $\langle\partial_\varphi\rangle^{\mathtt{b}}A$, $\langle\partial_\varphi\rangle^{\mathtt{b}}B$ (see Definition 2.3) are Lip(γ)*-modulo-tame with a modulo-tame constants, respectively, $\mathfrak{M}_{\langle\partial_\varphi\rangle^{\mathtt{b}}A}^\sharp(s)$ and $\mathfrak{M}_{\langle\partial_\varphi\rangle^{\mathtt{b}}B}^\sharp(s)$. Then $\langle\partial_\varphi\rangle^{\mathtt{b}}(AB)$ is* Lip(γ)*-modulo-tame with a modulo-tame constant satisfying, for some $C(\mathtt{b}) \geq 1$,*

$$\begin{aligned}\mathfrak{M}_{\langle\partial_\varphi\rangle^{\mathtt{b}}(AB)}^\sharp(s) \leq C(\mathtt{b})\Big(&\mathfrak{M}_{\langle\partial_\varphi\rangle^{\mathtt{b}}A}^\sharp(s)\mathfrak{M}_B^\sharp(s_1) + \mathfrak{M}_{\langle\partial_\varphi\rangle^{\mathtt{b}}A}^\sharp(s_1)\mathfrak{M}_B^\sharp(s)\\ &+ \mathfrak{M}_A^\sharp(s)\mathfrak{M}_{\langle\partial_\varphi\rangle^{\mathtt{b}}B}^\sharp(s_1) + \mathfrak{M}_A^\sharp(s_1)\mathfrak{M}_{\langle\partial_\varphi\rangle^{\mathtt{b}}B}^\sharp(s)\Big)\,.\end{aligned} \tag{2.36}$$

Proof. See [9, Lemma 2.25]. $\square$

Iterating (2.35)-(2.36) we obtain that, for any $n \geq 2$,

$$\mathfrak{M}^\sharp_{A^n}(s) \leq \big(2C\mathfrak{M}^\sharp_A(s_1)\big)^{n-1}\mathfrak{M}^\sharp_A(s)\,, \tag{2.37}$$

and

$$\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}A^n}(s) \leq (4C(\mathtt{b})C)^{n-1}\Big(\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}A}(s)\big[\mathfrak{M}^\sharp_A(s_1)\big]^{n-1} + \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}A}(s_1)\mathfrak{M}^\sharp_A(s)\big[\mathfrak{M}^\sharp_A(s_1)\big]^{n-2}\Big)\,. \tag{2.38}$$

As an application of (2.37)-(2.38) we obtain the following

Lemma 2.20. (Exponential map) *Let A and $\langle\partial_\varphi\rangle^{\mathtt{b}}A$ be* $\mathrm{Lip}(\gamma)$*-modulo-tame operators and assume that* $\mathfrak{M}^\sharp_A : [s_1, S] \to [0, +\infty)$ *is a modulo-tame constant satisfying*

$$\mathfrak{M}^\sharp_A(s_1) \leq 1\,. \tag{2.39}$$

Then the operators $\Phi^{\pm 1} := \exp(\pm A)$, $\Phi^{\pm1} - \mathrm{Id}$ *and* $\langle\partial_\varphi\rangle^{\mathtt{b}}(\Phi^{\pm1} - \mathrm{Id})$ *are* $\mathrm{Lip}(\gamma)$*-modulo-tame with modulo-tame constants satisfying, for any* $s_1 \leq s \leq S$,

$$\begin{gathered}\mathfrak{M}^\sharp_{\Phi^{\pm1}-\mathrm{Id}}(s) \lesssim \mathfrak{M}^\sharp_A(s)\,,\\ \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}(\Phi^{\pm1}-\mathrm{Id})}(s) \lesssim_{\mathtt{b}} \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}A}(s) + \mathfrak{M}^\sharp_A(s)\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}A}(s_1)\,.\end{gathered} \tag{2.40}$$

Proof. In view of the identity $\Phi^{\pm1} - \mathrm{Id} = \sum_{n\geq1}\frac{(\pm A)^n}{n!}$ and the assumption (2.39) the claimed estimates follow by (2.37)-(2.38). $\square$

Lemma 2.21. (Smoothing) *Suppose that* $\langle\partial_\varphi\rangle^{\mathtt{b}}A$, $\mathtt{b} \geq 0$, *is* $\mathrm{Lip}(\gamma)$*-modulo-tame. Then the operator* $\Pi^\perp_N A$ *(see Definition 2.3) is* $\mathrm{Lip}(\gamma)$*-modulo-tame with a modulo-tame constant satisfying*

$$\mathfrak{M}^\sharp_{\Pi^\perp_N A}(s) \leq N^{-\mathtt{b}}\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}A}(s)\,, \quad \mathfrak{M}^\sharp_{\Pi^\perp_N A}(s) \leq \mathfrak{M}^\sharp_A(s)\,. \tag{2.41}$$

Proof. See [9, Lemma 2.27]. $\square$

Lemma 2.22. *Let* $a_1(\cdot;\omega), a_2(\cdot;\omega)$ *be functions in* $\mathcal{C}^\infty(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1,\mathbb{C})$ *and* $\omega \in \Omega$. *Consider the linear operator* $\mathcal{R}$ *defined by* $\mathcal{R}h := a_1 \cdot (a_2, h)_{L^2_x}$, *for any* $h \in L^2_x$. *Then for any* $\lambda \in \mathbb{N}^{\mathbb{S}_+}$ *and* $n_1, n_2 \geq 0$, *the operator* $\langle D\rangle^{n_1}\partial_\varphi^\lambda\mathcal{R}\langle D\rangle^{n_2}$ *is* $\mathrm{Lip}(\gamma)$*-tame with a tame constant satisfying, for some* $\sigma \equiv \sigma(n_1, n_2, \lambda) > 0$,

$$\mathfrak{M}_{\langle D\rangle^{n_1}\partial_\varphi^\lambda\mathcal{R}\langle D\rangle^{n_2}}(s) \lesssim_{S,n_1,n_2,\lambda} (\max_{i=1,2}\|a_i\|_{s+\sigma})\cdot(\max_{i=1,2}\|a_i\|_{s_0+\sigma})\,.$$

Proof. For any $n_1, n_2 \geq 0$, $\lambda \in \mathbb{N}^{\mathbb{S}_+}$, $h \in L^2_x$, one has

$$\langle D\rangle^{n_1}\partial_\varphi^\lambda\mathcal{R}\langle D\rangle^{n_2}h = \sum_{\lambda_1+\lambda_2=\lambda} c_{\lambda_1,\lambda_2}\langle D\rangle^{n_1}[\partial_\varphi^{\lambda_1}a_1]\,\big(\langle D\rangle^{n_2}[\partial_\varphi^{\lambda_2}a_2]\,,\,h\big)_{L^2_x}$$

where we used that the operator $\langle D\rangle$ is symmetric. The lemma then follows by (2.5). $\square$

2.4 Tame estimates

In this section we record various tame estimates for compositions of functions and operators with a torus embedding $\breve{\imath} : \mathbb{T}^{\mathbb{S}_+} \to \mathcal{E}_s$ of the form (cf. (1.22))

$$\breve{\imath}(\varphi) = (\varphi, 0, 0) + \iota(\varphi)\,, \quad \iota(\varphi) = (\Theta(\varphi), y(\varphi), w(\varphi))\,,$$

with norm $\|\iota\|_s^{\mathrm{Lip}(\gamma)} := \|\Theta\|^{\mathrm{Lip}(\gamma)}_{H^s_\varphi} + \|y\|^{\mathrm{Lip}(\gamma)}_{H^s_\varphi} + \|w\|^{\mathrm{Lip}(\gamma)}_s$. We shall use that the Sobolev norm in (1.15) is equivalent to

$$\|\ \|_s = \|\ \|_{H^s_{\varphi,x}} \sim_s \|\ \|_{H^s_\varphi L^2_x} + \|\ \|_{L^2_\varphi H^s_x} \tag{2.42}$$

and the interpolation estimate (which is a consequence of Young's inequality)

$$\|w\|_{H^s_\varphi H^\sigma_x} \le \|w\|_{H^{s+\sigma}_\varphi L^2_x} + \|w\|_{L^2_\varphi H^{s+\sigma}_x} \lesssim_{s,\sigma} \|w\|_{s+\sigma}\,. \tag{2.43}$$

Given a Banach space X with norm $\|\ \|_X$, we consider the space $\mathcal{C}^s(\mathbb{T}^{\mathbb{S}_+}, X)$, $s \in \mathbb{N}$, of $\mathcal{C}^s-$smooth maps $f : \mathbb{T}^{\mathbb{S}_+} \to X$ equipped with the norm

$$\|f\|_{\mathcal{C}^s_\varphi X} := \sum_{0\le|\alpha|\le s} \|\partial^\alpha_\varphi f\|^{\sup}_X\,, \qquad \|\partial^\alpha_\varphi f\|^{\sup}_X := \sup_{\varphi\in\mathbb{T}^{\mathbb{S}_+}} \|\partial^\alpha_\varphi f(\varphi)\|_X\,. \tag{2.44}$$

By the Sobolev embedding $\|f\|_{\mathcal{C}^s_\varphi X} \lesssim_{s_1} \|f\|_{H^{s+s_1}_\varphi X}$ for $s_1 > |\mathbb{S}_+|$, whereas if X is a Hilbert space, the latter estimate is valid for $s_1 > |\mathbb{S}_+|/2$. On the scale of Banach spaces $\mathcal{C}^s(\mathbb{T}^{\mathbb{S}_+}, X)$ the following interpolation inequalities hold: for any $0 \le k \le s$,

$$\|f\|_{\mathcal{C}^k_\varphi X} \lesssim_s \|f\|^{1-\frac{k}{s}}_{\mathcal{C}^0_\varphi X} \|f\|^{\frac{k}{s}}_{\mathcal{C}^s_\varphi X}\,. \tag{2.45}$$

Recall that $\mathcal{E}_s, E_s$ are defined in (1.22) and $\mathcal{V}^s(\delta)$ in (1.24). Let Ω be an open bounded subset of $\mathbb{R}^{\mathbb{S}_+}$.

Lemma 2.23. *Let $\sigma > 0$ and assume that, for any $s \ge 0$, the map $a : (\mathcal{V}^\sigma(\delta) \cap \mathcal{E}_{s+\sigma}) \times \Omega \to H^s(\mathbb{T}_1)$ is $\mathcal{C}^\infty$ with respect to $\mathfrak{x} = (\theta, y, w)$, $\mathcal{C}^1$ with respect to ω, and satisfies for any $\mathfrak{x} \in \mathcal{V}^\sigma(\delta) \cap \mathcal{E}_{s+\sigma}$, $\alpha \in \mathbb{N}^{\mathbb{S}_+}$ with $|\alpha| \le 1$, and $l \ge 1$, the tame estimates*

$$\begin{aligned} &\|\partial^\alpha_\omega a(\mathfrak{x};\omega)\|_{H^s_x} \lesssim_s 1 + \|w\|_{H^{s+\sigma}_x}\,, \\ &\|d^l \partial^\alpha_\omega a(\mathfrak{x};\omega)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H^s_x} \lesssim_{s,l,\alpha} \sum_{j=1}^{l} \Big(\|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma}} \prod_{n\ne j} \|\widehat{\mathfrak{x}}_n\|_{E_\sigma} \Big) + \|w\|_{H^{s+\sigma}_x} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_\sigma}\,. \end{aligned} \tag{2.46}$$

Then for any $\breve{\iota}$ with $\|\iota\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma} \le \delta$, the following tame estimates hold for any $s \ge 0$:
(i)

$$\begin{aligned} &\|a(\breve{\iota})\|^{\mathrm{Lip}(\gamma)}_s \lesssim_s 1 + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma}\,, \\ &\|da(\breve{\iota})[\widehat{\iota}_1]\|^{\mathrm{Lip}(\gamma)}_s \lesssim_s \|\widehat{\iota}_1\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma} + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma} \|\widehat{\iota}_1\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma}\,, \\ &\|d^2 a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|^{\mathrm{Lip}(\gamma)}_s \lesssim_s \|\widehat{\iota}_1\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma} \|\widehat{\iota}_2\|_{s_0+\sigma} + \|\widehat{\iota}_1\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma} \|\widehat{\iota}_2\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma} \\ &\qquad\qquad + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma} \|\widehat{\iota}_1\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma} \|\widehat{\iota}_2\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma}\,. \end{aligned} \tag{2.47}$$

(ii) If in addition $a(\theta, 0, 0; \omega) = 0$, then $\|a(\breve{\iota})\|^{\mathrm{Lip}(\gamma)}_s \lesssim_s \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma}$.
(iii) If in addition $a(\theta, 0, 0; \omega) = 0$, $\partial_y a(\theta, 0, 0; \omega) = 0$, and $\partial_w a(\theta, 0, 0; \omega) = 0$, then

$$\begin{aligned} &\|a(\breve{\iota})\|^{\mathrm{Lip}(\gamma)}_s \lesssim_s \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma} \|\iota\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma}\,, \\ &\|da(\breve{\iota})[\widehat{\iota}]\|^{\mathrm{Lip}(\gamma)}_s \lesssim_s \|\iota\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma} \|\widehat{\iota}\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma} + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\sigma} \|\widehat{\iota}\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma}\,. \end{aligned}$$

Proof. (*i*) It suffices to prove the estimates in (2.47) for $\|d^2 a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|_s$ and $\|d^2 a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|^{\mathrm{lip}}_s$ since the ones for $a(\breve{\iota})$ and $da(\breve{\iota})$ then follow by Taylor expansions. By the hypothesis (2.46) with $l = 2$, $\alpha = 0$, we have, for any $\varphi \in \mathbb{T}^{\mathbb{S}_+}$, $s \ge 0$,

$$\begin{aligned} \|d^2 a(\breve{\iota}(\varphi))[\widehat{\iota}_1(\varphi), \widehat{\iota}_2(\varphi)]\|_{H^s_x} &\lesssim_s \|\widehat{\iota}_1(\varphi)\|_{E_{s+\sigma}} \|\widehat{\iota}_2(\varphi)\|_{E_\sigma} + \|\widehat{\iota}_1(\varphi)\|_{E_\sigma} \|\widehat{\iota}_2(\varphi)\|_{E_{s+\sigma}} \\ &\quad + \|\iota(\varphi)\|_{E_{s+\sigma}} \|\widehat{\iota}_1(\varphi)\|_{E_\sigma} \|\widehat{\iota}_2(\varphi)\|_{E_\sigma}\,. \end{aligned} \tag{2.48}$$

Squaring the expressions on the left and right hand side of (2.48) and then integrating them with respect to φ, one concludes, using (2.42), (2.43), and the Sobolev embedding (1.18), that

$$\|d^2 a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|_{L^2_\varphi H^s_x} \lesssim_s \|\widehat{\iota}_1\|_{s+\sigma} \|\widehat{\iota}_2\|_{s_0+\sigma} + \|\widehat{\iota}_1\|_{s_0+\sigma} \|\widehat{\iota}_2\|_{s+\sigma} + \|\iota\|_{s+\sigma} \|\widehat{\iota}_1\|_{s_0+\sigma} \|\widehat{\iota}_2\|_{s_0+\sigma}\,. \tag{2.49}$$

In order to estimate $\|d^2a(\breve\iota)[\widehat\iota_1,\widehat\iota_2]\|_{H^s_\varphi L^2_x}$, we estimate $\|d^2a(\breve\iota)[\widehat\iota_1,\widehat\iota_2]\|_{\mathcal{C}^s_\varphi L^2_x}$. We claim that

$$\|d^2a(\breve\iota)[\widehat\iota_1,\widehat\iota_2]\|_{\mathcal{C}^s_\varphi L^2_x} \lesssim_s \|\widehat\iota_1\|_{s_0+\sigma}\|\widehat\iota_2\|_{s+s_0+\sigma} + \|\widehat\iota_1\|_{s+s_0+\sigma}\|\widehat\iota_2\|_{s_0+\sigma} + \|\iota\|_{s+s_0+\sigma}\|\widehat\iota_1\|_{s_0+\sigma}\|\widehat\iota_2\|_{s_0+\sigma} \tag{2.50}$$

so that the estimate for $\|d^2a(\breve\iota)[\widehat\iota_1,\widehat\iota_2]\|_s$ stated in (2.47) follows by (2.49), (2.50), and (2.42). The bound for $\|d^2a(\breve\iota)[\widehat\iota_1,\widehat\iota_2]\|_s^{\mathrm{lip}}$ is obtained in the same fashion.

Proof of (2.50). By the Leibnitz rule, for any $\beta\in\mathbb{N}^{\mathbb{S}_+}$, $0\le|\beta|\le s$,

$$\partial_\varphi^\beta\Big(d^2a(\breve\iota(\varphi))[\widehat\iota_1(\varphi),\widehat\iota_2(\varphi)]\Big) = \sum_{\beta_1+\beta_2+\beta_3=\beta} c_{\beta_1,\beta_2,\beta_3}\partial_\varphi^{\beta_1}(d^2a(\breve\iota(\varphi)))\big[\partial_\varphi^{\beta_2}\widehat\iota_1(\varphi),\partial_\varphi^{\beta_3}\widehat\iota_2(\varphi)\big] \tag{2.51}$$

where $c_{\beta_1,\beta_2,\beta_3}$ are combinatorial constants. Each term in the latter sum is estimated individually. For $1\le|\beta_1|\le s$ we have

$$\begin{aligned}&\partial_\varphi^{\beta_1}(d^2a(\breve\iota(\varphi)))\big[\partial_\varphi^{\beta_2}\widehat\iota_1(\varphi),\partial_\varphi^{\beta_3}\widehat\iota_2(\varphi)\big] = \\ &\sum_{\substack{1\le m\le|\beta_1|\\ \alpha_1+\cdots+\alpha_m=\beta_1}} c_{\alpha_1,\cdots,\alpha_m} d^{m+2}a(\breve\iota(\varphi))\big[\partial_\varphi^{\alpha_1}\breve\iota(\varphi),\cdots,\partial_\varphi^{\alpha_m}\breve\iota(\varphi),\partial_\varphi^{\beta_2}\widehat\iota_1(\varphi),\partial_\varphi^{\beta_3}\widehat\iota_2(\varphi)\big]\end{aligned}$$

for suitable combinatorial constants $c_{\alpha_1,\cdots,\alpha_m}$. Then, by (2.46) with $l=m+2$, $\alpha=0$, we have the bound

$$\begin{aligned}&\|\partial_\varphi^{\beta_1}(d^2a(\breve\iota))[\partial_\varphi^{\beta_2}\widehat\iota_1,\partial_\varphi^{\beta_3}\widehat\iota_2]\|_{\mathcal{C}^0_\varphi L^2_x} \lesssim_\beta \\ &\sum_{\substack{1\le m\le|\beta_1|\\ \alpha_1+\cdots+\alpha_m=\beta_1}} (1+\|\iota\|_{\mathcal{C}^{|\alpha_1|}_\varphi E_\sigma})\cdots(1+\|\iota\|_{\mathcal{C}^{|\alpha_m|}_\varphi E_\sigma})\|\widehat\iota_1\|_{\mathcal{C}^{|\beta_2|}_\varphi E_\sigma}\|\widehat\iota_2\|_{\mathcal{C}^{|\beta_3|}_\varphi E_\sigma}\,.\end{aligned} \tag{2.52}$$

Arguing as in the proof of the formula (75) in [8], for any $j=1,\ldots,m$, we have

$$(1+\|\iota\|_{\mathcal{C}^{|\alpha_j|}_\varphi E_\sigma}) \lesssim_\beta (1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})^{1-\frac{|\alpha_j|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma})^{\frac{|\alpha_j|}{|\beta|}}\,,$$

and, using the interpolation estimate (2.45), we get

$$\begin{aligned}&(1+\|\iota\|_{\mathcal{C}^{|\alpha_1|}_\varphi E_\sigma})\cdots(1+\|\iota\|_{\mathcal{C}^{|\alpha_m|}_\varphi E_\sigma})\|\widehat\iota_1\|_{\mathcal{C}^{|\beta_2|}_\varphi E_\sigma}\|\widehat\iota_2\|_{\mathcal{C}^{|\beta_3|}_\varphi E_\sigma} \\ &\lesssim_s \|\widehat\iota_1\|_{\mathcal{C}^0_\varphi E_\sigma}^{1-\frac{|\beta_2|}{|\beta|}}\|\widehat\iota_1\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma}^{\frac{|\beta_2|}{|\beta|}}\|\widehat\iota_2\|_{\mathcal{C}^0_\varphi E_\sigma}^{1-\frac{|\beta_3|}{|\beta|}}\|\widehat\iota_2\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma}^{\frac{|\beta_3|}{|\beta|}}\prod_{j=1}^m(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})^{1-\frac{|\alpha_j|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma})^{\frac{|\alpha_j|}{|\beta|}} \\ &\lesssim_s \|\widehat\iota_1\|_{\mathcal{C}^0_\varphi E_\sigma}^{1-\frac{|\beta_2|}{|\beta|}}\|\widehat\iota_1\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma}^{\frac{|\beta_2|}{|\beta|}}\|\widehat\iota_2\|_{\mathcal{C}^0_\varphi E_\sigma}^{1-\frac{|\beta_3|}{|\beta|}}\|\widehat\iota_2\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma}^{\frac{|\beta_3|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})^{m-\sum_{j=1}^m\frac{|\alpha_j|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma})^{\sum_{j=1}^m\frac{|\alpha_j|}{|\beta|}}\,.\end{aligned} \tag{2.53}$$

By (1.18), (2.43), $(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})^{m-1}\lesssim(1+\|\iota\|_{s_0+\sigma})^{m-1}\lesssim(1+\delta)^{m-1}$ and $\frac{\sum_{j=1}^m|\alpha_j|}{|\beta|}=\frac{|\beta_1|}{|\beta|}=1-\frac{|\beta_2|}{|\beta|}-\frac{|\beta_3|}{|\beta|}$, so that

$$\begin{aligned}(2.53) &\lesssim_s \|\widehat\iota_1\|_{\mathcal{C}^0_\varphi E_\sigma}^{\frac{|\beta_1|+|\beta_3|}{|\beta|}}\|\widehat\iota_1\|_{\mathcal{C}^s_\varphi E_\sigma}^{\frac{|\beta_2|}{|\beta|}}\|\widehat\iota_2\|_{\mathcal{C}^0_\varphi E_\sigma}^{\frac{|\beta_1|+|\beta_2|}{|\beta|}}\|\widehat\iota_2\|_{\mathcal{C}^s_\varphi E_\sigma}^{\frac{|\beta_3|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})^{\frac{|\beta_2|+|\beta_3|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^s_\varphi E_\sigma})^{\frac{|\beta_1|}{|\beta|}} \\ &\lesssim_s \Big(\|\widehat\iota_1\|_{\mathcal{C}^0_\varphi E_\sigma}\|\widehat\iota_2\|_{\mathcal{C}^0_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^s_\varphi E_\sigma})\Big)^{\frac{|\beta_1|}{|\beta|}}\Big(\|\widehat\iota_1\|_{\mathcal{C}^s_\varphi E_\sigma}\|\widehat\iota_2\|_{\mathcal{C}^0_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})\Big)^{\frac{|\beta_2|}{|\beta|}} \\ &\quad\times\Big(\|\widehat\iota_1\|_{\mathcal{C}^0_\varphi E_\sigma}\|\widehat\iota_2\|_{\mathcal{C}^s_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})\Big)^{\frac{|\beta_3|}{|\beta|}}\end{aligned}$$

and, by the iterated Young inequality with exponents $|\beta|/|\beta_1|$, $|\beta|/|\beta_2|$, $|\beta|/|\beta_3|$, we conclude that (2.53) is bounded by

$$\begin{aligned}&\|\widehat\iota_1\|_{\mathcal{C}^0_\varphi E_\sigma}\|\widehat\iota_2\|_{\mathcal{C}^0_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^s_\varphi E_\sigma})+\|\widehat\iota_1\|_{\mathcal{C}^s_\varphi E_\sigma}\|\widehat\iota_2\|_{\mathcal{C}^0_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})+\|\widehat\iota_1\|_{\mathcal{C}^0_\varphi E_\sigma}\|\widehat\iota_2\|_{\mathcal{C}^s_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma}) \\ &\overset{(1.18),(2.43)}{\lesssim_s} \|\iota\|_{s+s_0+\sigma}\|\widehat\iota_1\|_{s_0+\sigma}\|\widehat\iota_2\|_{s_0+\sigma} + +\|\widehat\iota_1\|_{s+s_0+\sigma}\|\widehat\iota_2\|_{s_0+\sigma}+\|\widehat\iota_1\|_{s_0+\sigma}\|\widehat\iota_2\|_{s+s_0+\sigma}\,.\end{aligned}$$

Note that (2.52) satisfies the same type of bound as (2.53). The term in (2.51) with $\beta_1 = 0$ is estimated in the same way and thus (2.50) is proved.

PROOF *(ii)*-*(iii)*. Let $\varphi \mapsto \breve{\imath}(\varphi) = (\theta(\varphi), y(\varphi), w(\varphi))$ be a torus embedding. If $a(\theta, 0, 0) = 0$, we write

$$a(\breve{\imath}) = \int_0^1 da(\breve{\imath}_t)[\widehat{\imath}]\, dt\,, \quad \breve{\imath}_t = (1-t)(\theta(\varphi), 0, 0) + t\breve{\imath}(\varphi)\,,\ \widehat{\imath} := (0, y(\varphi), w(\varphi))\,,$$

and, if $a(\theta, 0, 0), \partial_y a(\theta, 0, 0), \partial_w a(\theta, 0, 0)$ vanish, we write

$$a(\breve{\imath}) = \int_0^1 (1-t) d^2 a(\breve{\imath}_t)[\widehat{\imath}, \widehat{\imath}]\, dt\,.$$

Items *(ii)*-*(iii)* follow by item *(i)*, noting that $\|\widehat{\imath}\|_s^{\mathrm{Lip}(\gamma)} = \|(0, y(\cdot), w(\cdot))\|_s^{\mathrm{Lip}(\gamma)} \lesssim \|\iota\|_s^{\mathrm{Lip}(\gamma)}$ for any $s \geq 0$. $\square$

Given $M \in \mathbb{N}$, we define the constant

$$\mathfrak{s}_M := \max\{s_0, M+1\}\,. \tag{2.54}$$

Lemma 2.24. *Assume that, for any $M \geq 0$, there is $\sigma_M \geq 0$ so that:*

- **Assumption A.** *For any $s \geq 0$, the map*

$$\mathcal{R} : (\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}) \times \Omega \to \mathcal{B}(H^s(\mathbb{T}_1), H^{s+M+1}(\mathbb{T}_1))$$

is $\mathcal{C}^\infty$ with respect to $\mathfrak{x}$, $\mathcal{C}^1$ with respect to ω and, for any $\mathfrak{x} \in \mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}$, $\alpha \in \mathbb{N}^{\mathbb{S}_+}$ with $|\alpha| \leq 1$,

$$\|\partial_\omega^\alpha \mathcal{R}(\mathfrak{x}; \omega)[\widehat{w}]\|_{H_x^{s+M+1}} \lesssim_{s,M} \|\widehat{w}\|_{H_x^s} + \|w\|_{H_x^{s+\sigma_M}} \|\widehat{w}\|_{L_x^2}\,,$$

and, for any $l \geq 1$, $\|d^l \partial_\omega^\alpha \mathcal{R}(\mathfrak{x}; \omega)[\widehat{w}][\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{s+M+1}}$ is bounded by

$$\lesssim_{s,M,l} \|\widehat{w}\|_{H_x^s} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} + \|\widehat{w}\|_{L_x^2} \Big(\|w\|_{H_x^{s+\sigma_M}} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} + \sum_{j=1}^l \big(\|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{n \neq j} \|\widehat{\mathfrak{x}}_n\|_{E_{\sigma_M}} \big) \Big)\,.$$

- **Assumption B.** *For any $-M-1 \leq s \leq 0$, the map*

$$\mathcal{R} : \mathcal{V}^{\sigma_M}(\delta) \times \Omega \to \mathcal{B}(H^s(\mathbb{T}_1), H^{s+M+1}(\mathbb{T}_1))$$

is $\mathcal{C}^\infty$ w.r to $\mathfrak{x}$, $\mathcal{C}^1$ with respect to ω and, for any $\mathfrak{x} \in \mathcal{V}^{\sigma_M}(\delta)$, $\alpha \in \mathbb{N}^{\mathbb{S}_+}$ with $|\alpha| \leq 1$, and $l \geq 1$,

$$\|\partial_\omega^\alpha \mathcal{R}(\mathfrak{x}; \omega)[\widehat{w}]\|_{H_x^{s+M+1}} \lesssim_{s,M} \|\widehat{w}\|_{H_x^s}\,,$$

$$\|d^l \partial_\omega^\alpha \mathcal{R}(\mathfrak{x}; \omega)[\widehat{w}][\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{s+M+1}} \lesssim_{s,M,l} \|\widehat{w}\|_{H_x^s} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,.$$

Then for any $S \geq \mathfrak{s}_M$ and $\lambda \in \mathbb{N}^{\mathbb{S}_+}$, there is a constant $\sigma_M(\lambda) > 0$, so that for any $\breve{\imath}(\varphi) = (\varphi, 0, 0) + \iota(\varphi)$ with $\|\iota\|_{s_0+\sigma_M(\lambda)}^{\mathrm{Lip}(\gamma)} \leq \delta$ and any $n_1, n_2 \in \mathbb{N}$ satisfying $n_1 + n_2 \leq M+1$, the following holds:

(i) The operator $\langle D\rangle^{n_1} \partial_\varphi^\lambda (\mathcal{R} \circ \breve{\imath}) \langle D\rangle^{n_2}$ is $\mathrm{Lip}(\gamma)$-tame with a tame constant satisfying, for any $\mathfrak{s}_M \leq s \leq S$,

$$\mathfrak{M}_{\langle D\rangle^{n_1} \partial_\varphi^\lambda (\mathcal{R}\circ\breve{\imath}) \langle D\rangle^{n_2}}(s) \lesssim_{S,M,\lambda} 1 + \|\iota\|_{s+\sigma_M(\lambda)}^{\mathrm{Lip}(\gamma)}\,.$$

(ii) The operator $\langle D\rangle^{n_1} \partial_\varphi^\lambda (d\mathcal{R}(\breve{\imath})[\widehat{\imath}]) \langle D\rangle^{n_2}$ is $\mathrm{Lip}(\gamma)$-tame with a tame constant satisfying, for any $\mathfrak{s}_M \leq s \leq S$,

$$\mathfrak{M}_{\langle D\rangle^{n_1} \partial_\varphi^\lambda (d\mathcal{R}(\breve{\imath})[\widehat{\imath}]) \langle D\rangle^{n_2}}(s) \lesssim_{S,M,\lambda} \|\widehat{\imath}\|_{s+\sigma_M(\lambda)}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma_M(\lambda)}^{\mathrm{Lip}(\gamma)} \|\widehat{\imath}\|_{s_0+\sigma_M(\lambda)}^{\mathrm{Lip}(\gamma)}\,.$$

(iii) If in addition $\mathcal{R}(\theta, 0, 0; \omega) = 0$, then the operator $\langle D\rangle^{n_1} \partial_\varphi^\lambda (\mathcal{R} \circ \breve{\imath}) \langle D\rangle^{n_2}$ is $\mathrm{Lip}(\gamma)$-tame with a tame constant satisfying, for any $\mathfrak{s}_M \leq s \leq S$,

$$\mathfrak{M}_{\langle D\rangle^{n_1} \partial_\varphi^\lambda (\mathcal{R}\circ\breve{\imath}) \langle D\rangle^{n_2}}(s) \lesssim_{S,M,\lambda} \|\iota\|_{s+\sigma_M(\lambda)}^{\mathrm{Lip}(\gamma)}\,.$$

Proof. Since item (i) and (ii) can be proved in a similar way, we only prove (ii). For any given $n_1, n_2 \in \mathbb{N}$ with $n_1 + n_2 \leq M+1$, set $\mathcal{Q} := \langle D \rangle^{n_1} \mathcal{R} \langle D \rangle^{n_2}$. Assumption A implies that for any $s \geq M+1$ and any $\mathfrak{r} \in \mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}$, the operator $\mathcal{Q}(\mathfrak{r})$ is in $\mathcal{B}(H^s_x)$ and for any $\widehat{\mathfrak{r}}_1, \ldots, \widehat{\mathfrak{r}}_l \in E_{s+\sigma_M}$ with $l \geq 1$, and $\widehat{w} \in H^s_x$,

$$
\begin{aligned}
&\|\mathcal{Q}(\mathfrak{r})[\widehat{w}]\|_{H^s_x} \lesssim_{s,M} \|\widehat{w}\|_{H^s_x} + \|w\|_{H^{s+\sigma_M}_x} \|\widehat{w}\|_{H^{M+1}_x} \,, \\
&\|d^l \big(\mathcal{Q}(\mathfrak{r})[\widehat{w}]\big)[\widehat{\mathfrak{r}}_1, \ldots, \widehat{\mathfrak{r}}_l]\|_{H^s_x} \lesssim_{s,M,l} \|\widehat{w}\|_{H^s_x} \prod_{j=1}^{l} \|\widehat{\mathfrak{r}}_j\|_{E_{\sigma_M}} \\
&\quad + \|\widehat{w}\|_{H^{M+1}_x} \Big(\|\mathfrak{r}\|_{E_{s+\sigma_M}} \prod_{j=1}^{l} \|\widehat{\mathfrak{r}}_j\|_{E_{\sigma_M}} + \sum_{j=1}^{l} \|\widehat{\mathfrak{r}}_j\|_{E_{s+\sigma_M}} \prod_{n \neq j} \|\widehat{\mathfrak{r}}_n\|_{E_{\sigma_M}} \Big) \,.
\end{aligned}
\tag{2.55}
$$

Furthermore Assumption B implies that, for any $\mathfrak{r} \in \mathcal{V}^{\sigma_M}(\delta)$, the operator $\mathcal{Q}(\mathfrak{r})$ is in $\mathcal{B}(L^2_x)$ and for any $\widehat{\mathfrak{r}}_1, \ldots, \widehat{\mathfrak{r}}_l \in E_{\sigma_M}$, $l \geq 1$,

$$
\|\mathcal{Q}(\mathfrak{r})\|_{\mathcal{B}(L^2_x)} \lesssim_M 1 \,, \qquad \|d^l \mathcal{Q}(\mathfrak{r})[\widehat{\mathfrak{r}}_1, \ldots, \widehat{\mathfrak{r}}_l]\|_{\mathcal{B}(L^2_x)} \lesssim_{M,l} \prod_{j=1}^{l} \|\widehat{\mathfrak{r}}_j\|_{E_{\sigma_M}} \,. \tag{2.56}
$$

One computes by Leibniz's rule

$$
\partial_\varphi^\lambda \big(d\mathcal{Q}(\breve{\imath}(\varphi))[\widehat{\imath}(\varphi)] \big) = \sum_{\substack{0 \leq k \leq |\lambda| \\ \lambda_1 + \ldots + \lambda_{k+1} = \lambda}} c_{\lambda_1, \ldots, \lambda_{k+1}} d^{k+1} \mathcal{Q}(\breve{\imath}(\varphi))[\partial_\varphi^{\lambda_1} \breve{\imath}(\varphi), \ldots, \partial_\varphi^{\lambda_k} \breve{\imath}(\varphi), \partial_\varphi^{\lambda_{k+1}} \widehat{\imath}(\varphi)] \tag{2.57}
$$

where $c_{\lambda_1, \ldots, \lambda_{k+1}}$ are combinatorial constants.

Estimate of $\|\partial_\varphi^\lambda \big(d\mathcal{Q}(\breve{\imath}(\varphi))[\widehat{\imath}(\varphi)] \big)[\widehat{w}]\|_{L^2_\varphi H^s_x}$. By (2.55), we have, for $s \geq M+1$,

$$
\begin{aligned}
&\|d^{k+1} \mathcal{Q}(\breve{\imath}(\varphi))[\partial_\varphi^{\lambda_1} \breve{\imath}(\varphi), \ldots, \partial_\varphi^{\lambda_k} \breve{\imath}(\varphi), \partial_\varphi^{\lambda_{k+1}} \widehat{\imath}(\varphi)][\widehat{w}(\varphi)]\|_{H^s_x} \\
&\lesssim_{s,M,k} \|\widehat{w}(\varphi)\|_{H^s_x} \|\partial_\varphi^{\lambda_{k+1}} \widehat{\imath}(\varphi)\|_{E_{\sigma_M}} \prod_{n=1}^{k} \|\partial_\varphi^{\lambda_n} \breve{\imath}(\varphi)\|_{E_{\sigma_M}} \\
&+ \|\widehat{w}(\varphi)\|_{H^{M+1}_x} \Big(\|\imath(\varphi)\|_{E_{s+\sigma_M}} \|\partial_\varphi^{\lambda_{k+1}} \widehat{\imath}(\varphi)\|_{E_{\sigma_M}} \prod_{n=1}^{k} \|\partial_\varphi^{\lambda_n} \breve{\imath}(\varphi)\|_{E_{\sigma_M}} \\
&+ \sum_{j=1}^{k} \|\partial_\varphi^{\lambda_j} \breve{\imath}(\varphi)\|_{E_{s+\sigma_M}} \big(\prod_{n \neq j} \|\partial_\varphi^{\lambda_n} \breve{\imath}(\varphi)\|_{E_{\sigma_M}} \big) \|\partial_\varphi^{\lambda_{k+1}} \widehat{\imath}(\varphi)\|_{E_{\sigma_M}} + \|\partial_\varphi^{\lambda_{k+1}} \widehat{\imath}(\varphi)\|_{E_{s+\sigma_M}} \prod_{n=1}^{k} \|\partial_\varphi^{\lambda_n} \breve{\imath}(\varphi)\|_{E_{\sigma_M}} \Big) \,.
\end{aligned}
\tag{2.58}
$$

Note that by the Sobolev embedding and (2.43), for any $s \geq 0$, $\mu \in \mathbb{N}^{\mathbb{S}_+}$,

$$
\|\partial_\varphi^\mu \breve{\imath}(\varphi)\|_{E_s} \lesssim 1 + \|\partial_\varphi^\mu \imath\|_{C^0_\varphi E_s} \lesssim 1 + \|\imath\|_{s+s_0+|\mu|} \,. \tag{2.59}
$$

Hence (2.57)-(2.58) and $\|\cdot\|_{L^2_\varphi H^s_x} \lesssim \|\cdot\|_s$ imply that for any $\breve{\imath}$ with $\|\imath\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_M(\lambda)} \leq \delta$ and any $s \geq M+1$,

$$
\begin{aligned}
&\|\partial_\varphi^\lambda \big(d\mathcal{Q}(\breve{\imath}(\varphi))[\widehat{\imath}(\varphi)] \big)[\widehat{w}(\varphi)]\|_{L^2_\varphi H^s_x} \\
&\lesssim_{s,M,\lambda} \|\widehat{w}\|_s \|\widehat{\imath}\|_{s_0+\sigma_M(\lambda)} + \|\widehat{w}\|_{M+1} \big(\|\imath\|_{s+\sigma_M(\lambda)} \|\widehat{\imath}\|_{s_0+\sigma_M(\lambda)} + \|\widehat{\imath}\|_{s+\sigma_M(\lambda)} \big)
\end{aligned}
\tag{2.60}
$$

for some constant $\sigma_M(\lambda) > 0$.

Estimate of $\|\partial_\varphi^\lambda \big(d\mathcal{Q}(\breve{\imath}(\varphi))[\widehat{\imath}(\varphi)] \big)\|_{H^s_\varphi \mathcal{B}(L^2_x)}$. For any $s \in \mathbb{N}$, $\beta \in \mathbb{N}^{\mathbb{S}_+}$, $|\beta| \leq s$, we need to estimate $\|\partial_\varphi^{\beta+\lambda} \big(d\mathcal{Q}(\breve{\imath}(\varphi))[\widehat{\imath}(\varphi)] \big)\|_{L^2_\varphi \mathcal{B}(L^2_x)}$. As in (2.57) we have

$$
\partial_\varphi^{\beta+\lambda} \big(d\mathcal{Q}(\breve{\imath}(\varphi))[\widehat{\imath}(\varphi)] \big) = \sum_{\substack{0 \leq k \leq |\beta| + |\lambda| \\ \alpha_1 + \ldots + \alpha_{k+1} = \beta + \lambda}} c_{\alpha_1, \ldots, \alpha_{k+1}} d^{k+1} \mathcal{Q}(\breve{\imath}(\varphi))[\partial_\varphi^{\alpha_1} \breve{\imath}(\varphi), \ldots, \partial_\varphi^{\alpha_k} \breve{\imath}(\varphi), \partial_\varphi^{\alpha_{k+1}} \widehat{\imath}(\varphi)] \tag{2.61}
$$

where $c_{\alpha_1,\ldots,\alpha_{k+1}}$ are combinatorial constants. By (2.56) and (2.59) one obtains that

$$\|d^{k+1}\mathcal{Q}(\breve{\iota}(\varphi))[\partial_\varphi^{\alpha_1}\breve{\iota}(\varphi),\ldots,\partial_\varphi^{\alpha_k}\breve{\iota}(\varphi),\partial_\varphi^{\alpha_{k+1}}\widehat{\iota}(\varphi)]\|_{L^2_\varphi\mathcal{B}(L^2_x)} \lesssim_{\beta,\lambda} \prod_{j=1}^{k}(1+\|\iota\|_{|\alpha_j|+\eta_M})\|\widehat{\iota}\|_{|\alpha_{k+1}|+\eta_M} \tag{2.62}$$

for some $\eta_M>0$. Using the interpolation inequality (2.4), and arguing as in the proof of the formula (75) in [8], we have, for any $\breve{\iota}$ with $\|\iota\|_{\eta_M}\le 1$ and any $j=1,\ldots,k$,

$$\begin{aligned}
&1+\|\iota\|_{|\alpha_j|+\eta_M}\lesssim(1+\|\iota\|_{\eta_M})^{1-\frac{|\alpha_j|}{|\beta+\lambda|}}(1+\|\iota\|_{|\beta+\lambda|+\eta_M})^{\frac{|\alpha_j|}{|\beta+\lambda|}} \overset{\|\iota\|_{\eta_M}\le 1}{\lesssim} (1+\|\iota\|_{|\beta+\lambda|+\eta_M})^{\frac{|\alpha_j|}{|\beta+\lambda|}}\,,\\
&\|\widehat{\iota}\|_{|\alpha_{k+1}|+\eta_M}\lesssim\|\widehat{\iota}\|_{\eta_M}^{1-\frac{|\alpha_{k+1}|}{|\beta+\lambda|}}\|\widehat{\iota}\|_{|\beta+\lambda|+\eta_M}^{\frac{|\alpha_{k+1}|}{|\beta+\lambda|}}\,.
\end{aligned}$$

Then by (2.62) and since $\sum_{j=1}^{k}|\alpha_j|+|\alpha_{k+1}|=|\beta+\lambda|$, it follows that

$$\begin{aligned}
&\|d^{k+1}\mathcal{Q}(\breve{\iota}(\varphi))[\partial_\varphi^{\alpha_1}\breve{\iota}(\varphi),\ldots,\partial_\varphi^{\alpha_k}\breve{\iota}(\varphi),\partial_\varphi^{\alpha_{k+1}}\widehat{\iota}(\varphi)]\|_{L^2_\varphi\mathcal{B}(L^2_x)}\\
&\lesssim_{s,\lambda}(1+\|\iota\|_{|\beta+\lambda|+\eta_M})^{\frac{\sum_{j=1}^k|\alpha_j|}{|\beta+\lambda|}}\|\widehat{\iota}\|_{\eta_M}^{1-\frac{|\alpha_{k+1}|}{|\beta+\lambda|}}\|\widehat{\iota}\|_{|\beta+\lambda|+\eta_M}^{\frac{|\alpha_{k+1}|}{|\beta+\lambda|}}\\
&\lesssim_{s,\lambda}\Big((1+\|\iota\|_{|\beta+\lambda|+\eta_M})\|\widehat{\iota}\|_{\eta_M}\Big)^{\frac{\sum_{j=1}^k|\alpha_j|}{|\beta|+|\lambda|}}\|\widehat{\iota}\|_{|\beta+\lambda|+\eta_M}^{\frac{|\alpha_{k+1}|}{|\beta+\lambda|}}\\
&\lesssim_{s,\lambda}\|\widehat{\iota}\|_{|\beta+\lambda|+\eta_M}+\|\iota\|_{|\beta+\lambda|+\eta_M}\|\widehat{\iota}\|_{\eta_M}
\end{aligned} \tag{2.63}$$

where for the latter inequality we used Young's inequality with exponents $\frac{|\beta+\lambda|}{\sum_{j=1}^k|\alpha_j|},\frac{|\beta+\lambda|}{|\alpha_{k+1}|}$. Combining (2.61) and (2.63) we obtain

$$\|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])\|_{H^s_\varphi\mathcal{B}(L^2_x)}\lesssim_{s,M,\lambda}\|\widehat{\iota}\|_{s+|\lambda|+\eta_M}+\|\iota\|_{s+|\lambda|+\eta_M}\|\widehat{\iota}\|_{\eta_M}\,. \tag{2.64}$$

Estimate of $\|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])[\widehat{w}]\|_{H^s_\varphi L^2_x}$. Using that

$$\Big(\sum_{\ell\in\mathbb{Z}^{\mathbb{S}_+}}\|\widehat{A}(\ell)\|^2_{\mathcal{B}(L^2_x)}\langle\ell\rangle^{2s}\Big)^{1/2}\lesssim_{s_0}\|A\|_{H^{s+s_0}_\varphi\mathcal{B}(L^2_x)}$$

one deduces from [8, Lemma 2.12] that for any $\breve{\iota}$ with $\|\iota\|_{2s_0+|\lambda|+\eta_M}\le 1$ and any $s\ge s_0$,

$$\begin{aligned}
\|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])[\widehat{w}]\|_{H^s_\varphi L^2_x}&\lesssim_s\|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])\|_{H^{2s_0}_\varphi\mathcal{B}(L^2_x)}\|\widehat{w}\|_{H^s_\varphi L^2_x}+\|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])\|_{H^{s+s_0}_\varphi\mathcal{B}(L^2_x)}\|\widehat{w}\|_{H^{s_0}_\varphi L^2_x}\\
&\overset{(2.64)}{\lesssim_{s,M}}\|\widehat{w}\|_s\|\widehat{\iota}\|_{2s_0+|\lambda|+\eta_M}+\|\widehat{w}\|_{s_0}\big(\|\widehat{\iota}\|_{s+s_0+|\lambda|+\eta_M}+\|\iota\|_{s+s_0+|\lambda|+\eta_M}\|\widehat{\iota}\|_{2s_0+|\lambda|+\eta_M}\big)\,.
\end{aligned} \tag{2.65}$$

Increasing the constant $\sigma_M(\lambda)$ in (2.60) if needed, one infers from the estimates (2.60), (2.65) that for any $s\ge\mathfrak{s}_M=\max\{s_0,M+1\}$, $\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])$ satisfies

$$\|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])[\widehat{w}]\|_s\lesssim_{s,M,\lambda}\|\widehat{w}\|_s\|\widehat{\iota}\|_{s_0+\sigma_M(\lambda)}+\|\widehat{w}\|_{\mathfrak{s}_M}\big(\|\widehat{\iota}\|_{s+\sigma_M(\lambda)}+\|\iota\|_{s+\sigma_M(\lambda)}\|\widehat{\iota}\|_{s_0+\sigma_M(\lambda)}\big)\,. \tag{2.66}$$

Furthermore, arguing similarly, one can show that for any $\omega_1,\omega_2\in\Omega$, $\omega_1\ne\omega_2$, the operator $\partial_\varphi^\lambda\Delta_\omega(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])$ satisfies the estimate, for any $s\ge\mathfrak{s}_M$

$$\gamma\frac{\|\partial_\varphi^\lambda\Delta_\omega(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])[\widehat{w}]\|_s}{|\omega_1-\omega_2|}\lesssim_{s,M,\lambda}\|\widehat{w}\|_s\|\widehat{\iota}\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_M(\lambda)}+\|\widehat{w}\|_{\mathfrak{s}_M}\big(\|\widehat{\iota}\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M(\lambda)}+\|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M(\lambda)}\|\widehat{\iota}\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_M(\lambda)}\big)\,. \tag{2.67}$$

It then follows from (2.66) and (2.67) that there exists a tame constant $\mathfrak{M}_{\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])}(s)$ for $\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])$ satisfying the estimate stated in item (ii).
Proof of (iii). Since $\mathcal{R}(\theta,0,0)=0$, we can write

$$\mathcal{R}(\breve{\iota})=\int_0^1 d\mathcal{R}(\breve{\iota}_t)[\widehat{\iota}]\,dt\,,\quad \breve{\iota}_t=(1-t)(\theta(\varphi),0,0)+t\breve{\iota}(\varphi)\,,\quad \widehat{\iota}(\varphi):=(0,y(\varphi),w(\varphi))\,.$$

Since $\|\widehat{\iota}\|_s\lesssim\|\iota\|_s$ for any $s\ge 0$, item (iii) is thus a direct consequence of (ii). $\square$

2.5 Egorov type theorems

The main purpose of this section is to investigate operators obtained by conjugating a pseudo-differential operator of the form $a(\varphi,x)\partial_x^m$, $m\in\mathbb{Z}$, by the flow map of a transport equation. These results are used in Section 6.3.

Let $\Phi(\tau_0,\tau,\varphi)$ denote the flow of the transport equation

$$\partial_\tau\Phi(\tau_0,\tau,\varphi)=B(\tau,\varphi)\Phi(\tau_0,\tau,\varphi)\,,\quad \Phi(\tau_0,\tau_0,\varphi)=\mathrm{Id}\,, \tag{2.68}$$

where

$$B(\tau,\varphi):=\Pi_\perp\big(b(\tau,\varphi,x)\partial_x+b_x(\tau,\varphi,x)\big)\,,\quad b\equiv b(\tau,\varphi,x):=\frac{\beta(\varphi,x)}{1+\tau\beta_x(\varphi,x)}\,, \tag{2.69}$$

and the real valued function $\beta(\varphi,x)\equiv\beta(\varphi,x;\omega)$ is $\mathcal{C}^\infty$ with respect to the variables (φ,x) and Lipschitz with respect to the parameter $\omega\in\Omega$. For brevity we set $\Phi(\tau,\varphi):=\Phi(0,\tau,\varphi)$ and $\Phi(\varphi):=\Phi(0,1,\varphi)$. Note that $\Phi(\varphi)^{-1}=\Phi(1,0,\varphi)$ and that

$$\Phi(\tau_0,\tau,\varphi)=\Phi(\tau,\varphi)\circ\Phi(\tau_0,\varphi)^{-1}\,. \tag{2.70}$$

By standard hyperbolic estimates, equation (2.68) is well-posed. The flow $\Phi(\tau_0,\tau,\varphi)$ has the following properties.

Lemma 2.25. (Transport flow) *Let $\lambda_0\in\mathbb{N}$, $S>s_0$. For any $\lambda\in\mathbb{N}$ with $\lambda\le\lambda_0$, $n_1,n_2\in\mathbb{R}$ with $n_1+n_2=-\lambda-1$, and $s\ge s_0$, there exist constants $\sigma(\lambda_0,n_1,n_2)>0$, $\delta\equiv\delta(S,\lambda_0,n_1,n_2)\in(0,1)$ such that, if*

$$\|\beta\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma(\lambda_0,n_1,n_2)}\le\delta\,, \tag{2.71}$$

then for any $m\in\mathbb{S}_+$, $\langle D\rangle^{n_1}\partial_{\varphi_m}^\lambda\Phi(\tau_0,\tau,\varphi)\langle D\rangle^{n_2}$ is a $\mathrm{Lip}(\gamma)$-tame operator with a tame constant satisfying

$$\mathfrak{M}_{\langle D\rangle^{n_1}\partial_{\varphi_m}^\lambda\Phi(\tau_0,\tau,\varphi)\langle D\rangle^{n_2}}(s)\lesssim_{S,\lambda_0,n_1,n_2}1+\|\beta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma(\lambda_0,n_1,n_2)}\,,\qquad\forall s_0\le s\le S\,,\quad\forall\tau_0,\tau\in[0,1]\,. \tag{2.72}$$

In addition, if $n_1+n_2=-\lambda-2$, then $\langle D\rangle^{n_1}\partial_{\varphi_m}^\lambda(\Phi(\tau_0,\tau,\varphi)-\mathrm{Id})\langle D\rangle^{n_2}$ is $\mathrm{Lip}(\gamma)$-tame with a tame constant satisfying

$$\mathfrak{M}_{\langle D\rangle^{n_1}\partial_{\varphi_m}^\lambda(\Phi(\tau_0,\tau,\varphi)-\mathrm{Id})\langle D\rangle^{n_2}}(s)\lesssim_{S,\lambda_0,n_1,n_2}\|\beta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma(\lambda_0,n_1,n_2)}\,,\qquad\forall s_0\le s\le S\,,\quad\forall\tau_0,\tau\in[0,1]\,. \tag{2.73}$$

Furthermore, let $s_0<s_1<S$, $n_1,n_2\in\mathbb{R}$, $\lambda_0\in\mathbb{N}$, $\lambda\le\lambda_0$ with $n_1+n_2=-\lambda-1$, $m\in\mathbb{S}_+$. If β_1 and β_2 satisfy $\|\beta_i\|_{s_1+\sigma(n_1,n_2)}\le\delta$ for some $\sigma(n_1,n_2)>0$, and $\delta\in(0,1)$ small enough, then

$$\|\langle D\rangle^{n_1}\partial_{\varphi_m}^\lambda\Delta_{12}\Phi(\tau_0,\tau,\varphi)\langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})}\lesssim_{s_1,\lambda_0,n_1,n_2}\|\Delta_{12}\beta\|_{s_1+\sigma(n_1,n_2)}\,,\quad\tau_0,\tau\in[0,1]\,, \tag{2.74}$$

where $\Delta_{12}\beta:=\beta_2-\beta_1$ and $\Delta_{12}\Phi(\tau_0,\tau,\varphi):=\Phi(\tau_0,\tau,\varphi;\beta_2)-\Phi(\tau_0,\tau,\varphi;\beta_1)$.

Proof. The proof of (2.72) is similar to the one of Propositions A.7, A.10 and A.11 in [9]. In comparison to the latter results the main difference is that the vector field (2.69) is of order 1, whereas the vector field considered in [9] is of order $\frac{1}{2}$. Using (2.72) we now prove (2.73). By (2.68), one has that

$$\Phi(\tau_0,\tau,\varphi)-\mathrm{Id}=\int_{\tau_0}^{\tau}B(t,\varphi)\Phi(\tau_0,t,\varphi)\,dt\,.$$

Then, for any $\lambda\in\mathbb{N}$ with $\lambda\le\lambda_0$ and any $n_1,n_2\in\mathbb{R}$ with $n_1+n_2=-\lambda-2$, one has by Leibniz' rule

$$\begin{aligned}&\langle D\rangle^{n_1}\partial_{\varphi_m}^\lambda(\Phi(\tau_0,\tau,\varphi)-\mathrm{Id})\langle D\rangle^{n_2}\\&=\sum_{\lambda_1+\lambda_2=\lambda}c_{\lambda_1,\lambda_2}\int_{\tau_0}^{\tau}\big(\langle D\rangle^{n_1}\partial_{\varphi_m}^{\lambda_1}B(t,\varphi)\langle D\rangle^{n_2+\lambda_2+1}\big)\big(\langle D\rangle^{-n_2-\lambda_2-1}\partial_{\varphi_m}^{\lambda_2}\Phi(\tau_0,t,\varphi)\langle D\rangle^{n_2}\big)\,dt\\&=\sum_{\lambda_1+\lambda_2=\lambda}c_{\lambda_1,\lambda_2}\int_{\tau_0}^{\tau}\big(\langle D\rangle^{n_1}\partial_{\varphi_m}^{\lambda_1}B(t,\varphi)\langle D\rangle^{-1-n_1-\lambda_1}\big)\big(\langle D\rangle^{-n_2-\lambda_2-1}\partial_{\varphi_m}^{\lambda_2}\Phi(\tau_0,t,\varphi)\langle D\rangle^{n_2}\big)\,dt\end{aligned}$$

where c_{λ_1,λ_2} are combinatorial constants and we used that $n_2 + \lambda_2 + 1 = -1 - n_1 - \lambda_1$. Recalling the definition (2.69) of B, using Lemmata 2.9, 2.16, 2.27-(i), and (2.72), one has that for any $s \geq s_0$,

$$
\begin{aligned}
\mathfrak{M}_{\langle D\rangle^{n_1}\partial_{\varphi_m}^{\lambda_1} B\langle D\rangle^{-1-n_1-\lambda_1}}(s) &\lesssim_s |\langle D\rangle^{n_1} B \langle D\rangle^{-1-n_1-\lambda_1}|_{0,s+\lambda_1,0}^{\mathrm{Lip}(\gamma)} \lesssim_{s,\lambda_1,n_1} \|\beta\|_{s+\sigma(\lambda_1,n_1)}^{\mathrm{Lip}(\gamma)}\,, \\
\mathfrak{M}_{\langle D\rangle^{-1-n_2-\lambda_2}\partial_{\varphi_m}^{\lambda_2}\Phi(\tau_0,t,\varphi)\langle D\rangle^{n_2}}(s) &\lesssim_{s,\lambda_2,n_1,n_2} 1 + \|\beta\|_{s+\sigma(\lambda_2,n_1,n_2)}^{\mathrm{Lip}(\gamma)}\,.
\end{aligned}
\tag{2.75}
$$

Then (2.73) follows by (2.75), Lemma 2.14 and (2.71). The estimate (2.74) follows by similar arguments. $\square$

For what follows we need to study the solutions of the characteristic ODE $\partial_\tau x = -b(\tau,\varphi,x)$ associated to the transport operator defined in (2.69).

Lemma 2.26. (Characteristic flow) *The characteristic flow $\gamma^{\tau_0,\tau}(\varphi,x)$ defined by*

$$
\partial_\tau \gamma^{\tau_0,\tau}(\varphi,x) = -b\big(\tau,\varphi,\gamma^{\tau_0,\tau}(\varphi,x)\big)\,, \quad \gamma^{\tau_0,\tau_0}(\varphi,x) = x\,,
\tag{2.76}
$$

is given by

$$
\gamma^{\tau_0,\tau}(\varphi,x) = x + \tau_0\beta(\varphi,x) + \breve{\beta}(\tau,\varphi,x+\tau_0\beta(\varphi,x))\,,
\tag{2.77}
$$

where $y \mapsto y + \breve{\beta}(\tau,\varphi,y)$ is the inverse diffeomorphism of $x \mapsto x + \tau\beta(\varphi,x)$.

Proof. A direct computation proves that $\gamma^{0,\tau}(y) = y + \breve{\beta}(\tau,\varphi,y)$ and therefore $\gamma^{\tau,0}(x) = x + \tau\beta(\varphi,x)$. By the composition rule of the flow $\gamma^{\tau_0,\tau} = \gamma^{0,\tau}\circ\gamma^{\tau_0,0}$ we deduce (2.77). $\square$

Lemma 2.27. *There are $\sigma,\delta > 0$ such that, if $\|\beta\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} \leq \delta$, then*
(i) $\|b\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\beta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}$ for any $s \geq s_0$.
(ii) For any $\tau_0,\tau \in [0,1]$, $s \geq s_0$, we have $\|\gamma^{\tau_0,\tau}(\varphi,x) - x\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\beta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}$.
(iii) Let $s_1 > s_0$ and assume that $\|\beta_j\|_{s_1+\sigma} \leq \delta$, $j = 1,2$. Then $\Delta_{12}b := b(\cdot;\beta_2) - b(\cdot;\beta_1)$ and $\Delta_{12}\gamma^{\tau_0,\tau} := \gamma^{\tau_0,\tau}(\cdot;\beta_2) - \gamma^{\tau_0,\tau}(\cdot;\beta_1)$ can be estimated in terms of $\Delta_{12}\beta := \beta_2 - \beta_1$ as

$$
\|\Delta_{12}b\|_{s_1} \lesssim_{s_1} \|\Delta_{12}\beta\|_{s_1+\sigma}\,, \qquad \|\Delta_{12}\gamma^{\tau_0,\tau}\|_{s_1} \lesssim_{s_1} \|\Delta_{12}\beta\|_{s_1+\sigma}\,.
$$

Proof. Item (i) follows by the definition of b in (2.69) and Lemma 2.2. Item (ii) follows by (2.77) and Lemma 2.1. Item (iii) follows by similar arguments. $\square$

Now we prove the following Egorov type theorem, saying that the operator, obtained by conjugating $a(\varphi,x)\partial_x^m$, $m \in \mathbb{Z}$, with the time one flow $\Phi(\varphi) = \Phi(0,1,\varphi)$ of the transport equation (2.68), remains a pseudo-differential operator with a homogenous asymptotic expansion.

Proposition 2.28. (Egorov) *Let $N, \lambda_0 \in \mathbb{N}$, $S > s_0$ and assume that $\beta(\cdot;\omega), a(\cdot;\omega)$ are in $\mathcal{C}^\infty(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)$ and Lipschitz continuous with respect to $\omega \in \Omega$. Then there exist constants $\sigma_N(\lambda_0)$, $\sigma_N > 0$, $\delta(S,N,\lambda_0) \in (0,1)$, and $C_0 > 0$ such that, if*

$$
\|\beta\|_{s_0+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)} \leq \delta\,, \quad \|a\|_{s_0+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)} \leq C_0\,,
\tag{2.78}
$$

then the conjugated operator

$$
\mathcal{P}(\varphi) := \Phi(\varphi)\mathcal{P}_0(\varphi)\Phi(\varphi)^{-1}\,, \qquad \mathcal{P}_0 := a(\varphi,x;\omega)\partial_x^m\,, \quad m \in \mathbb{Z}\,,
$$

is a pseudo-differential operator of order m with an expansion of the form

$$
\mathcal{P}(\varphi) = \sum_{i=0}^{N} p_{m-i}(\varphi,x;\omega)\partial_x^{m-i} + \mathcal{R}_N(\varphi)
\tag{2.79}
$$

with the following properties:

1. *The principal symbol p_m of $\mathcal{P}$ is given by*

$$p_m(\varphi, x; \omega) = \Big([1 + \breve{\beta}_y(\varphi, y; \omega)]^m a(\varphi, y; \omega) \Big)|_{y = x + \beta(\varphi, x; \omega)} \tag{2.80}$$

where $y \mapsto y + \breve{\beta}(\varphi, y; \omega)$ denotes the inverse diffeomorphism of $x \mapsto x + \beta(\varphi, x; \omega)$.

2. *For any $s \geq s_0$ and $i = 1, \ldots, N$,*

$$\|p_m - a\|_s^{\mathrm{Lip}(\gamma)}\,,\ \|p_{m-i}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,N} \|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)} + \|a\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)} \|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)}\,. \tag{2.81}$$

3. *For any $\lambda \in \mathbb{N}$ with $\lambda \leq \lambda_0$, $n_1, n_2 \in \mathbb{N}$ with $n_1 + n_2 + \lambda_0 \leq N - 1 - m$, $k \in \mathbb{S}_+$, the pseudo-differential operator $\langle D \rangle^{n_1} \partial_{\varphi_k}^\lambda \mathcal{R}_N(\varphi) \langle D \rangle^{n_2}$ is* Lip(γ)*-tame with a tame constant satisfying, for any $s_0 \leq s \leq S$,*

$$\mathfrak{M}_{\langle D \rangle^{n_1} \partial_{\varphi_k}^\lambda \mathcal{R}_N(\varphi) \langle D \rangle^{n_2}}(s) \lesssim_{S,N,\lambda_0} \|\beta\|_{s+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)} + \|a\|_{s+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)} \|\beta\|_{s_0+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)}\,. \tag{2.82}$$

4. *Let $s_0 < s_1$ and assume that $\|\beta_j\|_{s_1+\sigma_N(\lambda_0)} \leq \delta$, $\|a_j\|_{s_1+\sigma_N(\lambda_0)} \leq C_0$, $j = 1, 2$. Then*

$$\|\Delta_{12} p_{m-i}\|_{s_1} \lesssim_{s_1,N} \|\Delta_{12} a\|_{s_1+\sigma_N} + \|\Delta_{12}\beta\|_{s_1+\sigma_N}\,, \quad i = 0, \ldots, N\,,$$

and, for any $\lambda \leq \lambda_0$, $n_1, n_2 \in \mathbb{N}$ with $n_1 + n_2 + \lambda_0 \leq N - 1 - m$, and $k \in \mathbb{S}_+$,

$$\|\langle D \rangle^{n_1} \partial_{\varphi_k}^\lambda \Delta_{12} \mathcal{R}_N(\varphi) \langle D \rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,N,n_1,n_2} \|\Delta_{12} a\|_{s_1+\sigma_N(\lambda_0)} + \|\Delta_{12}\beta\|_{s_1+\sigma_N(\lambda_0)}$$

where we refer to Lemma 2.25 for the meaning of Δ_{12}.

Proof. The orthogonal projector $\Pi_\perp$ is a Fourier multiplier of order 0, $\Pi_\perp = \mathrm{Op}(\chi_\perp(\xi))$, where $\chi_\perp$ is a $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ cut-off function which is equal to 1 on a neighborhood of $\mathbb{S}^\perp$ and vanishes in a neighborhood of $\mathbb{S} \cup \{0\}$. Then we decompose the operator $B(\tau, \varphi) = \Pi_\perp(b(\tau, \varphi, x)\partial_x + b_x(\tau, \varphi, x))$ as

$$\begin{aligned} &B(\tau, \varphi) = B_1(\tau, \varphi) + B_\infty(\tau, \varphi)\,, \\ &B_1(\tau, \varphi) := b(\tau, \varphi, x)\partial_x + b_x(\tau, \varphi, x)\,, \quad B_\infty(\tau, \varphi) := \mathrm{Op}(b_\infty(\tau, \varphi, x, \xi)) \in OPS^{-\infty} \end{aligned} \tag{2.83}$$

where for some $\sigma > 0$, B_∞ satisfies, for any $s, m \geq 0$ and $\alpha \in \mathbb{N}$, the estimate

$$|B_\infty|_{-m,s,\alpha}^{\mathrm{Lip}(\gamma)} \lesssim_{m,s,\alpha} \|\beta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\,. \tag{2.84}$$

The conjugated operator $\mathcal{P}(\tau, \varphi) := \Phi(\tau, \varphi)\mathcal{P}_0(\varphi)\Phi(\tau, \varphi)^{-1}$ solves the Heisenberg equation

$$\partial_\tau \mathcal{P}(\tau, \varphi) = [B(\tau, \varphi), \mathcal{P}(\tau, \varphi)]\,, \quad \mathcal{P}(0, \varphi) = \mathcal{P}_0(\varphi) = a(\varphi, x; \omega)\partial_x^m\,. \tag{2.85}$$

We look for an approximate solution of (2.85) of the form

$$\mathcal{P}_N(\tau, \varphi) := \sum_{i=0}^N p_{m-i}(\tau, \varphi, x)\partial_x^{m-i} \tag{2.86}$$

for suitable functions $p_{m-i}(\tau, \varphi, x)$ to be determined. By (2.83)

$$[B(\tau, \varphi), \mathcal{P}_N(\tau, \varphi)] = [B_1(\tau, \varphi), \mathcal{P}_N(\tau, \varphi)] + [B_\infty(\tau, \varphi), \mathcal{P}_N(\tau, \varphi)] \tag{2.87}$$

where $[B_\infty(\tau, \varphi), \mathcal{P}_N(\tau, \varphi)]$ is in $OPS^{-\infty}$, and

$$[B_1(\tau, \varphi), \mathcal{P}_N(\tau, \varphi)] = \sum_{i=0}^N \big[b\partial_x + b_x\,,\, p_{m-i}\partial_x^{m-i} \big]\,.$$

By Lemma 2.11, one has for any $i = 0, \dots, N$,

$$\big[b\partial_x + b_x, p_{m-i}\partial_x^{m-i}\big] = \big(b(p_{m-i})_x - (m-i)b_x p_{m-i}\big)\partial_x^{m-i} + \sum_{j=1}^{N-i} g_j(b, p_{m-i})\partial_x^{m-i-j} + \mathcal{R}_N(b, p_{m-i})$$

where the functions $g_j(b, p_{m-i}) := g_j(b, p_{m-i})(\tau, \varphi, x)$, $j = 0, \dots, N-i$, and the remainders $\mathcal{R}_N(b, p_{m-i})$ can be estimated as follows: there exists $\sigma_N := \sigma_N(m) > 0$ so that for any $s \geq s_0$, (cf. Lemma 2.27-(i))

$$\|g_j(b, p_{m-i})\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{m,N,s} \|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)} \|p_{m-i}\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} + \|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} \|p_{m-i}\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}\,, \tag{2.88}$$

and for any $s \geq s_0$ and $\alpha \in \mathbb{N}$ (cf. Lemma 2.11-(ii))

$$|\mathcal{R}_N(b, p_{m-i})|_{m-N-1,s,\alpha}^{\mathrm{Lip}(\gamma)} \lesssim_{m,N,s,\alpha} \|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)} \|p_{m-i}\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} + \|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} \|p_{m-i}\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}\,. \tag{2.89}$$

Adding up the expansions for $\big[b\partial_x + b_x, p_{m-i}\partial_x^{m-i}\big]$, $0 \leq i \leq N$, yields

$$\begin{aligned}
\big[B_1(\tau,\varphi), \mathcal{P}_N(\tau,\varphi)\big] &= \sum_{i=0}^{N} \big(b(p_{m-i})_x - (m-i)b_x p_{m-i}\big)\partial_x^{m-i} + \sum_{i=0}^{N}\sum_{j=1}^{N-i} g_j(b, p_{m-i})\partial_x^{m-i-j} + \sum_{i=0}^{N} \mathcal{R}_N(b, p_{m-i}) \\
&= \sum_{i=0}^{N} \big(b(p_{m-i})_x - (m-i)b_x p_{m-i}\big)\partial_x^{m-i} + \sum_{k=1}^{N}\sum_{j=1}^{k} g_j(b, p_{m-k+j})\partial_x^{m-k} + \sum_{i=0}^{N} \mathcal{R}_N(b, p_{m-i}) \\
&= \big(b(p_m)_x - m b_x p_m\big)\partial_x^m + \sum_{i=1}^{N} \big(b(p_{m-i})_x - (m-i)b_x p_{m-i} + \widetilde{g}_i\big)\partial_x^{m-i} + \mathcal{Q}_N
\end{aligned} \tag{2.90}$$

where, for any $i = 1, \dots, N$, $\widetilde{g}_i := \sum_{j=1}^{i} g_j(b, p_{m-i+j})$ and $\mathcal{Q}_N := \sum_{i=0}^{N} \mathcal{R}_N(b, p_{m-i}) \in OPS^{m-N-1}$. Defining for any $s \geq 0$,

$$\mathtt{M}_{<i}(s) := \max\{\|p_{m-k}\|_s^{\mathrm{Lip}(\gamma)}, k = 0, \dots, i-1\}\,, \ \mathtt{M}(s) := \max\{\|p_{m-i}\|_s^{\mathrm{Lip}(\gamma)}, i = 0, \dots, N\}\,, \tag{2.91}$$

we deduce from (2.88) and (2.89) that for any $s \geq s_0$, $\alpha \in \mathbb{N}$, $i = 0, \dots, N$,

$$\begin{aligned}
&\|\widetilde{g}_i\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,N} \mathtt{M}_{<i}(s+\sigma_N)\|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} + \mathtt{M}_{<i}(s_0+\sigma_N)\|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)} \\
&|\mathcal{Q}_N|_{m-N-1,s,\alpha}^{\mathrm{Lip}(\gamma)} \lesssim_{s,N} \mathtt{M}(s+\sigma_N)\|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} + \mathtt{M}(s_0+\sigma_N)\|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}\,.
\end{aligned} \tag{2.92}$$

By (2.86), (2.87), and (2.90) the operator $\mathcal{P}_N(\tau,\varphi)$ solves the *approximated* Heisenberg equation

$$\partial_\tau \mathcal{P}_N(\tau,\varphi) = [B(\tau,\varphi), \mathcal{P}_N(\tau,\varphi)] + OPS^{m-N-1}\,,$$

if the functions p_{m-i} solve the transport equations

$$\begin{aligned}
&\partial_\tau p_m = b(p_m)_x - m b_x p_m\,, \\
&\partial_\tau p_{m-i} = b(p_{m-i})_x - (m-i)b_x p_{m-i} + \widetilde{g}_i\,, \quad i = 1, \dots, N\,.
\end{aligned} \tag{2.93}$$

Note that, since $\widetilde{g}_i$ only depends on $p_{m-i+1}, \dots, p_m$, we can solve (2.93) inductively.
Determination of p_m. We solve the first equation in (2.93),

$$\partial_\tau p_m(\tau,\varphi,x) = b(\tau,\varphi,x)\partial_x p_m(\tau,\varphi,x) - m b_x(\tau,\varphi,x)p_m(\tau,\varphi,x)\,, \quad p_m(0,\varphi,x) = a(\varphi,x)\,.$$

By the method of characteristics we deduce that

$$p_m(\tau,\varphi,\gamma^{0,\tau}(\varphi,x)) = \exp\Big(-m\int_0^\tau b_x(t,\varphi,\gamma^{0,t}(\varphi,x))\,dt\Big)a(\varphi,x) \tag{2.94}$$

where $\gamma^{0,\tau}(\varphi, x)$ is given by (2.77). Differentiating the equation (2.76) with respect to the initial datum x, we get

$$\partial_\tau(\partial_x \gamma^{\tau_0,\tau}(x)) = -b_x(\tau, \varphi, \gamma^{\tau_0,\tau}(x))\partial_x \gamma^{\tau_0,\tau}(x)\,, \quad \partial_x \gamma^{\tau_0,\tau_0}(x) = 1\,,$$

implying that

$$\partial_x \gamma^{\tau_0,\tau}(\varphi, x) = \exp\Big(- \int_{\tau_0}^{\tau} b_x(t, \varphi, \gamma^{\tau_0,t}(\varphi, x))\, dt \Big)\,. \tag{2.95}$$

From (2.94) and (2.95) we infer that

$$p_m(\tau, \varphi, y) = \Big([\partial_x \gamma^{0,\tau}(\varphi, x)]^m a(\varphi, x) \Big)|_{x=\gamma^{\tau,0}(\varphi,y)}\,. \tag{2.96}$$

Evaluating the latter identity at $\tau = 1$ and using (2.77), we obtain (2.80).

Inductive determination of p_{m-i}. For $i = 1, \ldots, N$, we solve the inhomogeneous transport equation,

$$\partial_\tau p_{m-i} = b\partial_x p_{m-i} - (m-i)b_x p_{m-i} + \widetilde{g}_i\,, \qquad p_{m-i}(0, \varphi, x) = 0\,.$$

By the method of characteristics one has

$$p_{m-i}(\tau, \varphi, y) = \int_0^\tau \exp\Big(- (m-i) \int_t^\tau b_x(s, \varphi, \gamma^{\tau,s}(\varphi, y))\, ds \Big) \widetilde{g}_i(t, \varphi, \gamma^{\tau,t}(\varphi, y))\, dt\,. \tag{2.97}$$

The functions $p_{m-i}(\varphi, y)$ in the expansion (2.79) are then given by $p_{m-i}(\varphi, y) := p_{m-i}(1, \varphi, y)$.

Lemma 2.29. *There are $\sigma_N^{(N)} > \sigma_N^{(N-1)} > \ldots > \sigma_N^{(0)} > 0$ such that, for any $i \in \{1, \ldots, N\}$, $\tau \in [0,1]$, $s \geq s_0$,*

$$\begin{aligned} \|p_m(\tau, \cdot) - a\|_s^{\mathrm{Lip}(\gamma)} &\lesssim_s \|\beta\|_{s+\sigma_N^{(0)}}^{\mathrm{Lip}(\gamma)} + \|a\|_{s+\sigma_N^{(0)}}^{\mathrm{Lip}(\gamma)} \|\beta\|_{s_0+\sigma_N^{(0)}}^{\mathrm{Lip}(\gamma)}\,, \\ \|p_{m-i}(\tau, \cdot)\|_s^{\mathrm{Lip}(\gamma)} &\lesssim_s \|\beta\|_{s+\sigma_N^{(i)}}^{\mathrm{Lip}(\gamma)} + \|a\|_{s+\sigma_N^{(i)}}^{\mathrm{Lip}(\gamma)} \|\beta\|_{s_0+\sigma_N^{(i)}}\,. \end{aligned} \tag{2.98}$$

Proof. We argue by induction. First we prove the claimed estimate for $p_m - a$ with p_m given by (2.96). Recall that $\gamma^{0,\tau}(\varphi, x) = x + \breve{\beta}(\tau, \varphi, x)$ and $\gamma^{\tau,0}(\varphi, y) = y + \tau\beta(\varphi, y)$ (cf. (2.77)). Since $a(\varphi, y + \tau\beta(\varphi, y)) - a(\varphi, y) = \int_0^\tau a_x(\varphi, y + t\beta(\varphi, y))\beta(\varphi, y) dt$, the claimed estimate for p_m then follows by Lemmata 2.1, 2.27 and assumption (2.78). Now assume that for any $k \in \{1, \ldots, i-1\}$, $1 \leq i \leq N$, the function p_{m-k}, given by (2.97), satisfies the estimates (2.98). The ones for p_{m-i} then follow by Lemmata 2.1, 2.2, 2.27, (2.92), (2.91), and (2.78). ☐

Lemma 2.29 proves (2.81). Furthermore, in view of the definition (2.86) of $\mathcal{P}_N(\tau, \varphi)$, it follows from (2.98), Lemma 2.9, (2.22) and (2.21) that for any $s \geq s_0$, $\alpha \in \mathbb{N}$,

$$|\mathcal{P}_N(\tau, \varphi)|_{m,s,\alpha}^{\mathrm{Lip}(\gamma)} \lesssim_{m,s,N,\alpha} \|a\|_s^{\mathrm{Lip}(\gamma)} + \|\beta\|_{s+\sigma_N^{(N)}}^{\mathrm{Lip}(\gamma)} + \|a\|_{s+\sigma_N^{(N)}}^{\mathrm{Lip}(\gamma)} \|\beta\|_{s_0+\sigma_N^{(N)}}^{\mathrm{Lip}(\gamma)}\,. \tag{2.99}$$

By (2.87), (2.90), and (2.93) we deduce that $\mathcal{P}_N(\tau, \varphi)$ solves

$$\begin{aligned} &\partial_\tau \mathcal{P}_N(\tau, \varphi) = [B(\tau, \varphi), \mathcal{P}_N(\tau, \varphi)] - \mathcal{Q}_N^{(1)}(\tau, \varphi)\,, \quad \mathcal{P}_N(0, \varphi) = a\partial_x^m\,, \\ &\mathcal{Q}_N^{(1)}(\tau, \varphi) := \mathcal{Q}_N(\tau, \varphi) + [B_\infty(\tau, \varphi), \mathcal{P}_N(\tau, \varphi)] \in OPS^{m-N-1}\,. \end{aligned} \tag{2.100}$$

We now estimate the difference between $\mathcal{P}_N(\tau)$ and $\mathcal{P}(\tau)$.

Lemma 2.30. *The operator $\mathcal{R}_N(\tau, \varphi) := \mathcal{P}(\tau, \varphi) - \mathcal{P}_N(\tau, \varphi)$ is given by*

$$\mathcal{R}_N(\tau, \varphi) = \int_0^\tau \Phi(\eta, \tau, \varphi)\mathcal{Q}_N^{(1)}(\eta, \varphi)\Phi(\tau, \eta, \varphi)\, d\eta\,. \tag{2.101}$$

Proof. One writes

$$\mathcal{P}_N(\tau,\varphi)-\mathcal{P}(\tau,\varphi)=\mathcal{V}_N(\tau,\varphi)\Phi(\tau,\varphi)^{-1}\,,\quad \mathcal{V}_N(\tau,\varphi):=\mathcal{P}_N(\tau,\varphi)\Phi(\tau,\varphi)-\Phi(\tau,\varphi)\mathcal{P}_0(\varphi)\,, \tag{2.102}$$

and a direct calculation shows that $\mathcal{V}_N(\tau)$ solves

$$\partial_\tau \mathcal{V}_N(\tau,\varphi)=B(\tau,\varphi)\mathcal{V}_N(\tau,\varphi)-\mathcal{Q}_N^{(1)}(\tau,\varphi)\Phi(\tau,\varphi)\,,\quad \mathcal{V}_N(0,\varphi)=0\,.$$

Hence, by variation of the constants, $\mathcal{V}_N(\tau,\varphi)=-\int_0^\tau \Phi(\tau,\varphi)\Phi(\eta,\varphi)^{-1}\mathcal{Q}_N^{(1)}(\eta,\varphi)\Phi(\eta,\varphi)\,d\eta$ and, by (2.102) and (2.70), we deduce (2.101). □

Next we prove the estimate (2.82) of Proposition 2.28 of $\mathcal{R}_N(\tau,\varphi)$, given by (2.101). First we estimate $\mathcal{Q}_N^{(1)}\in OPS^{m-N-1}$, defined in (2.100). The estimate of $\mathcal{Q}_N$, obtained from (2.92), (2.91), (2.98) , and the one of $[B_\infty(\tau,\varphi),\mathcal{P}_N(\tau,\varphi)]$, obtained from (2.84), (2.99), Lemma 2.10, yield that there exists a constant $\aleph_N>0$ so that for any $s\geq s_0$, $\alpha\in\mathbb{N}$,

$$|\mathcal{Q}_N^{(1)}(\eta,\varphi)|_{m-N-1,s,\alpha}^{\mathrm{Lip}(\gamma)}\lesssim_{m,s,\alpha,N}\|\beta\|_{s+\aleph_N}^{\mathrm{Lip}(\gamma)}+\|a\|_{s+\aleph_N}^{\mathrm{Lip}(\gamma)}\|\beta\|_{s_0+\aleph_N}^{\mathrm{Lip}(\gamma)}\,. \tag{2.103}$$

Let $\lambda_0,n_1,n_2\in\mathbb{N}$ with $\lambda\leq\lambda_0$ and $n_1+n_2+\lambda_0+m\leq N-1$, $k\in\mathbb{S}_+$. In view of the definition (2.101) of $\mathcal{R}_N(\tau,\varphi)$, the claimed estimate of $\langle D\rangle^{n_1}\partial_{\varphi_k}^{\lambda}\mathcal{R}_N(\tau,\varphi)\langle D\rangle^{n_2}$ will follow from corresponding ones of $\langle D\rangle^{n_1}\partial_{\varphi_k}^{\lambda_1}\Phi(\eta,\tau,\varphi)\partial_{\varphi_k}^{\lambda_2}\mathcal{Q}_N^{(1)}(\eta,\varphi)\partial_{\varphi_k}^{\lambda_3}\Phi(\tau,\eta,\varphi)\langle D\rangle^{n_2}$ ($\tau,\eta\in[0,1]$ and $\lambda_1+\lambda_2+\lambda_3=\lambda$) which we write as

$$\Big(\langle D\rangle^{n_1}\partial_{\varphi_k}^{\lambda_1}\Phi(\eta,\tau,\varphi)\langle D\rangle^{-n_1-\lambda_1-1}\Big)\Big(\langle D\rangle^{n_1+\lambda_1+1}\partial_{\varphi_k}^{\lambda_2}\mathcal{Q}_N^{(1)}(\eta,\varphi)\langle D\rangle^{n_2+\lambda_3+1}\Big)\Big(\langle D\rangle^{-n_2-\lambda_3-1}\partial_{\varphi_k}^{\lambda_3}\Phi(\tau,\eta,\varphi)\langle D\rangle^{n_2}\Big).$$

Then, we use Lemma 2.25 to estimate the tame constants of the operators $\langle D\rangle^{n_1}\partial_{\varphi_k}^{\lambda_1}\Phi(\eta,\tau,\varphi)\langle D\rangle^{-n_1-\lambda_1-1}$, $\langle D\rangle^{-n_2-\lambda_3-1}\partial_{\varphi_k}^{\lambda_3}\Phi(\tau,\eta,\varphi)\langle D\rangle^{n_2}$, the estimates (2.103), (2.21) and Lemmata 2.9, 2.16 to estimate the tame constant of $\langle D\rangle^{n_1+\lambda_1+1}\partial_{\varphi_k}^{\lambda_2}\mathcal{Q}_N^{(1)}(\eta,\varphi)\langle D\rangle^{n_2+\lambda_3+1}$ and Lemma 2.14 together with the assumption (2.78), to estimate the tame constant of the composition. The bound (2.82) is finally proved.

Item 4 of Proposition 2.28 can be shown by similar arguments. This completes the proof of the latter. □

In the sequel we also need to study the operator obtained by conjugating $\omega\cdot\partial_\varphi$ with the time one flow $\Phi(\varphi)=\Phi(0,1,\varphi)$ of the transport equation (2.68). Here we analyze the operator $\Phi(\varphi)\circ\omega\cdot\partial_\varphi(\Phi(\varphi)^{-1})$, which turns out to be a pseudo-differential operator of order one with an expansion in decreasing symbols.

Proposition 2.31. (Conjugation of $\omega\cdot\partial_\varphi$) *Let $N,\lambda_0\in\mathbb{N}$, $S>s_0$ and assume that $\beta(\cdot;\omega)$ is in $\mathcal{C}^\infty(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)$ and Lipschitz continuous with respect to $\omega\in\Omega$. Then there exist constants $\sigma_N(\lambda_0),\sigma_N>0$, $\delta(S,N,\lambda_0)\in(0,1)$, $C_0>0$ so that, if*

$$\|\beta\|_{s_0+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)}\leq\delta\,, \tag{2.104}$$

then $\mathcal{P}(\varphi):=\Phi(\varphi)\circ\omega\cdot\partial_\varphi(\Phi(\varphi)^{-1})$ is a pseudo-differential operator of order 1 with an expansion of the form

$$\mathcal{P}(\varphi)=\sum_{i=0}^{N}p_{1-i}(\varphi,x;\omega)\partial_x^{1-i}+\mathcal{R}_N(\varphi)$$

with the following properties:

1. *For any $i=0,\dots,N$ and $s\geq s_0$, $\|p_{1-i}\|_s^{\mathrm{Lip}(\gamma)}\lesssim_{s,N}\|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}$.*
2. *For any $\lambda\in\mathbb{N}$ with $\lambda\leq\lambda_0$, for any $n_1,n_2\in\mathbb{N}$ with $n_1+n_2+\lambda_0\leq N-2$, and for any $k\in\mathbb{S}_+$, the pseudo-differential operator $\langle D\rangle^{n_1}\partial_{\varphi_k}^{\lambda}\mathcal{R}_N(\varphi)\langle D\rangle^{n_2}$ is* Lip(γ)*-tame with a tame constant satisfying, for any $s_0\leq s\leq S$,*

$$\mathfrak{M}_{\langle D\rangle^{n_1}\partial_{\varphi_k}^{\lambda}\mathcal{R}_N(\varphi)\langle D\rangle^{n_2}}(s)\lesssim_{S,N,\lambda_0}\|\beta\|_{s+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)}\,.$$

3. *Let $s_0 < s_1 < S$ and assume that $\|\beta_i\|_{s_1+\sigma_N(\lambda_0)} \le \delta$, $i = 1, 2$. Then*
$$\|\Delta_{12} p_{1-i}\|_{s_1} \lesssim_{s_1,N} \|\Delta_{12}\beta\|_{s_1+\sigma_N}, \quad i = 0, \dots, N,$$
and, for any $\lambda \le \lambda_0$, $n_1, n_2 \in \mathbb{N}$ with $n_1 + n_2 + \lambda_0 \le N - 2$, and $k \in \mathbb{S}_+$
$$\|\langle D\rangle^{n_1} \partial_{\varphi_k}^{\lambda} \Delta_{12} \mathcal{R}_N(\varphi) \langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,N,n_1,n_2} \|\Delta_{12}\beta\|_{s_1+\sigma_N(\lambda_0)}$$
where we refer to Lemma 2.25 for the meaning of Δ_{12}.

Proof. The operator $\Psi(\tau, \varphi) := \Phi(\tau, \varphi) \circ \omega \cdot \partial_\varphi(\Phi(\tau, \varphi)^{-1})$ solves the inhomogeneous Heisenberg equation
$$\partial_\tau \Psi(\tau, \varphi) = [B(\tau, \varphi), \Psi(\tau, \varphi))] - \omega \cdot \partial_\varphi(B(\tau, \varphi)), \quad \Psi(0, \varphi) = 0.$$
The latter equation can be solved in a similar way as (2.85) by looking for approximate solutions of the form of a pseudo-differential operator of order 1, admitting an expansion in homogeneous components (cf. (2.86)). The proof then proceeds in the same way as the one for Proposition 2.28 and hence is omitted. □

We finish this section by the following application of Proposition 2.28 to Fourier multipliers.

Lemma 2.32. *Let $N, \lambda_0 \in \mathbb{N}$, $S > s_0$ and assume that $\mathcal{Q}$ is a Lipschitz family of Fourier multipliers with an expansion of the form*
$$\mathcal{Q} = \sum_{n=0}^{N} c_{m-n}(\omega) \partial_x^{m-n} + \mathcal{Q}_N(\omega), \quad \mathcal{Q}_N(\omega) \in \mathcal{B}(H^s, H^{s+N+1-m}), \quad \forall s \ge 0. \tag{2.105}$$
Then there exist $\sigma_N(\lambda_0)$, $\sigma_N > 0$, and $\delta(S, N, \lambda_0) \in (0, 1)$ so that, if
$$\|\beta\|_{s_0+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)} \le \delta(S, N, \lambda_0), \tag{2.106}$$
then $\Phi(\varphi)\mathcal{Q}\Phi(\varphi)^{-1}$ is an operator of the form $\mathcal{Q} + \mathcal{Q}_\Phi(\varphi) + \mathcal{R}_N(\varphi)$ with the following properties:

1. *$\mathcal{Q}_\Phi(\varphi) = \sum_{n=0}^{N} \alpha_{m-n}(\varphi, x; \omega) \partial_x^{m-n}$ where for any $s \ge s_0$,*
$$\|\alpha_{m-n}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,N} \|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}, \quad n = 0, \dots, N. \tag{2.107}$$
2. *For any $\lambda \in \mathbb{N}$ with $\lambda \le \lambda_0$, $n_1, n_2 \in \mathbb{N}$ with $n_1 + n_2 + \lambda_0 \le N - m - 2$, and $k \in \mathbb{S}_+$, the operator $\langle D\rangle^{n_1} \partial_{\varphi_k}^{\lambda} \mathcal{R}_N \langle D\rangle^{n_2}$ is* Lip(γ)*-tame with a tame constant satisfying*
$$\mathfrak{M}_{\langle D\rangle^{n_1} \partial_{\varphi_k}^{\lambda} \mathcal{R}_N \langle D\rangle^{n_2}}(s) \lesssim_{S,N,\lambda_0} \|\beta\|_{s+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)}, \quad \forall s_0 \le s \le S. \tag{2.108}$$
3. *Let $s_0 < s_1 < S$ and assume that $\|\beta_i\|_{s_1+\sigma_N(\lambda_0)} \le \delta$, $i = 1, 2$. Then*
$$\|\Delta_{12}\alpha_{m-n}\|_{s_1} \lesssim_{s_1,N} \|\Delta_{12}\beta\|_{s_1+\sigma_N}, \quad n = 0, \dots, N,$$
and, for any $\lambda \le \lambda_0$, $n_1, n_2 \in \mathbb{N}$ with $n_1 + n_2 + \lambda_0 \le N - m - 2$, and $k \in \mathbb{S}_+$,
$$\|\langle D\rangle^{n_1} \partial_{\varphi_k}^{\lambda} \Delta_{12} \mathcal{R}_N(\varphi) \langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,N,n_1,n_2} \|\Delta_{12}\beta\|_{s_1+\sigma_N(\lambda_0)}$$
where we refer to Lemma 2.25 for the meaning of Δ_{12}.

Proof. Applying Proposition 2.28 to $\Phi(\varphi)\partial_x^{m-n}\Phi(\varphi)^{-1}$ for $n = 0, \dots, N$, we get
$$\Phi(\varphi)\Big(\sum_{n=0}^{N} c_{m-n}(\omega)\partial_x^{m-n}\Big)\Phi(\varphi)^{-1} = \sum_{n=0}^{N} c_{m-n}(\omega)\partial_x^{m-n} + \mathcal{Q}_\Phi(\varphi) + \mathcal{R}_N^{(1)}(\varphi)$$
where $\mathcal{Q}_\Phi(\varphi) = \sum_{n=0}^{N} \alpha_{m-n}(\varphi, x; \omega)\partial_x^{m-n}$ with α_{m-n} satisfying (2.107) and the remainder $\mathcal{R}_N^{(1)}(\varphi)$ satisfying (2.108). Next we write $\Phi(\varphi)\mathcal{Q}_N\Phi(\varphi)^{-1} = \mathcal{Q}_N + \mathcal{R}_N^{(2)}(\varphi)$ where
$$\mathcal{R}_N^{(2)}(\varphi) := \big(\Phi(\varphi) - \mathrm{Id}\big)\mathcal{Q}_N\Phi(\varphi)^{-1} + \mathcal{Q}_N\big(\Phi(\varphi)^{-1} - \mathrm{Id}\big).$$
We then argue as in the proof of the estimate of the remainder $\mathcal{R}_N(\tau, \varphi)$ in Proposition 2.28. Using Lemma 2.25 and the assumption that $\mathcal{Q}_N$ is a Fourier multiplier in $\mathcal{B}(H^s, H^{s+N+1-m})$ we get that $\mathcal{R}_N^{(2)}(\varphi)$ satisfies (2.108), and $\mathcal{R}_N(\varphi) = \mathcal{R}_N^{(1)}(\varphi) + \mathcal{R}_N^{(2)}(\varphi)$ satisfies (2.108) as well. Item 3 follows by similar arguments. □

3 Integrable features of KdV

3.1 Normal form coordinates for the KdV equation

In this section we rephrase Theorem 1.1 in [15] adapted to our purposes and prove some corollaries.

We consider an open bounded set $\Xi \subset \mathbb{R}^{\mathbb{S}_+}_{>0}$ so that (1.13) holds for some $\delta > 0$. Recall that $\mathcal{V}^s(\delta) \subset \mathcal{E}_s$, $\mathcal{V}(\delta) = \mathcal{V}^0(\delta)$ are defined in (1.24) and that we denote by $\mathfrak{x} = (\theta, y, w)$ its elements. The space $\mathcal{V}(\delta) \cap \mathcal{E}_s$ is endowed with the symplectic form

$$\mathcal{W} := \Big(\sum\nolimits_{j\in\mathbb{S}_+} dy_j \wedge d\theta_j\Big) \oplus \mathcal{W}_\perp \tag{3.1}$$

where $\mathcal{W}_\perp$ is the restriction to $L^2_\perp(\mathbb{T}_1)$ of the symplectic form $\mathcal{W}_{L^2_0}$ defined in (1.7). The Poisson structure $\mathcal{J}$ corresponding to $\mathcal{W}$, defined by the identity $\{F, G\} = \mathcal{W}(X_F, X_G) = \langle \nabla F\, , \, \mathcal{J}\nabla G\rangle$, is the unbounded operator

$$\mathcal{J} : E_s \to E_s\,, \quad (\widehat{\theta}, \widehat{y}, \widehat{w}) \mapsto (-\widehat{y}, \widehat{\theta}, \partial_x \widehat{w}) \tag{3.2}$$

where $\langle\ ,\ \rangle$ is the bilinear form in (1.23).

Theorem 3.1. (Normal KdV coordinates with pseudo-differential expansion, [15]). *Let $\mathbb{S}_+ \subseteq \mathbb{N}$ be finite, Ξ an open bounded subset of $\mathbb{R}^{\mathbb{S}_+}_{>0}$ so that* (1.13) *holds, for some $\delta > 0$. Then, for $\delta > 0$ sufficiently small, there exists a canonical $\mathcal{C}^\infty$ family of diffeomorphisms $\Psi_\nu : \mathcal{V}(\delta) \to \Psi_\nu(\mathcal{V}(\delta)) \subseteq L^2_0(\mathbb{T}_1)\,, (\theta, y, w) \mapsto q$, $\nu \in \Xi$, with the property that Ψ_ν satisfies*

$$\Psi_\nu(\theta, y, 0) = \Psi^{kdv}(\theta, \nu + y, 0)\,, \quad \forall (\theta, y, 0) \in \mathcal{V}(\delta)\,, \quad \forall \nu \in \Xi\,,$$

and is compatible with the scale of Sobolev spaces $H^s_0(\mathbb{T}_1), s \in \mathbb{N}$, in the sense that $\Psi_\nu\big(\mathcal{V}(\delta) \cap \mathcal{E}_s\big) \subseteq H^s_0(\mathbb{T}_1)$ and $\Psi_\nu : \mathcal{V}(\delta) \cap \mathcal{E}_s \to H^s_0(\mathbb{T}_1)$ is a $\mathcal{C}^\infty-$diffeomorphism onto its image, so that the following holds:

(AE1) *For any integer $M \geq 1$, $\nu \in \Xi$, $\mathfrak{x} = (\theta, y, w) \in \mathcal{V}(\delta)$, $\Psi_\nu(\mathfrak{x})$ admits an asymptotic expansion of the form*

$$\Psi_\nu(\theta, y, w) = \Psi^{kdv}(\theta, \nu + y, 0) + w + \sum_{k=1}^{M} a^\Psi_{-k}(\mathfrak{x}; \nu)\, \partial_x^{-k} w + \mathcal{R}^\Psi_M(\mathfrak{x}; \nu) \tag{3.3}$$

where $\mathcal{R}^\Psi_M(\theta, y, 0; \nu) = 0$ and, for any $s \in \mathbb{N}$ and $1 \leq k \leq M$, the functions

$$\mathcal{V}(\delta) \times \Xi \to H^s(\mathbb{T}_1),\, (\mathfrak{x}, \nu) \mapsto a^\Psi_{-k}(\mathfrak{x}; \nu)\,, \quad (\mathcal{V}(\delta) \cap \mathcal{E}_s) \times \Xi \to H^{s+M+1}(\mathbb{T}_1),\, (\mathfrak{x}, \nu) \mapsto \mathcal{R}^\Psi_M(\mathfrak{x}; \nu)\,,$$

are $\mathcal{C}^\infty$.

(AE2) *For any $\mathfrak{x} \in \mathcal{V}^1(\delta)$, $\nu \in \Xi$, the transpose $d\Psi_\nu(\mathfrak{x})^\top$ of the differential $d\Psi_\nu(\mathfrak{x}) : E_1 \to H^1_0(\mathbb{T}_1)$ is a bounded linear operator $d\Psi_\nu(\mathfrak{x})^\top : H^1_0(\mathbb{T}_1) \to E_1$, and, for any $\widehat{q} \in H^1_0(\mathbb{T}_1)$ and integer $M \geq 1$, $d\Psi_\nu(\mathfrak{x})^\top[\widehat{q}]$ admits an expansion of the form*

$$d\Psi_\nu(\mathfrak{x})^\top[\widehat{q}] = \Big(0, 0, \Pi_\perp \widehat{q} + \Pi_\perp \sum_{k=1}^{M} a^{d\Psi^\top}_{-k}(\mathfrak{x}; \nu) \partial_x^{-k} \widehat{q} + \Pi_\perp \sum_{k=1}^{M} (\partial_x^{-k} w)\, \mathcal{A}^{d\Psi^\top}_{-k}(\mathfrak{x}; \nu)[\widehat{q}]\Big) + \mathcal{R}^{d\Psi^\top}_M(\mathfrak{x}; \nu)[\widehat{q}] \tag{3.4}$$

where, for any $s \geq 1$ and $1 \leq k \leq M$,

$$\begin{aligned}
&\mathcal{V}^1(\delta) \times \Xi \to H^s(\mathbb{T}_1)\,,\ (\mathfrak{x}, \nu) \mapsto a^{d\Psi^\top}_{-k}(\mathfrak{x}; \nu)\,,\\
&\mathcal{V}^1(\delta) \times \Xi \to \mathcal{B}(H^1_0(\mathbb{T}_1), H^s(\mathbb{T}_1))\,,\ (\mathfrak{x}, \nu) \mapsto \mathcal{A}^{d\Psi^\top}_{-k}(\mathfrak{x}; \nu)\,,\\
&(\mathcal{V}^1(\delta) \cap \mathcal{E}_s) \times \Xi \to \mathcal{B}(H^s_0(\mathbb{T}_1), E_{s+M+1})\,,\ (\mathfrak{x}, \nu) \mapsto \mathcal{R}^{d\Psi^\top}_M(\mathfrak{x}; \nu)\,,
\end{aligned}$$

are $\mathcal{C}^\infty$. Furthermore,

$$a^{d\Psi^\top}_{-1}(\mathfrak{x}; \nu) = -a^\Psi_{-1}(\mathfrak{x}; \nu)\,. \tag{3.5}$$

(AE3) *For any $\nu \in \Xi$, the Hamiltonian $\mathcal{H}^{kdv}(\cdot\,;\nu) := H^{kdv} \circ \Psi_\nu : \mathcal{V}^1(\delta) \to \mathbb{R}$ is in normal form up to order three, meaning that*

$$\mathcal{H}^{kdv}(\theta, y, w; \nu) = \omega^{kdv}(\nu) \cdot y + \frac{1}{2}\big(\Omega^{kdv}(D;\nu)w, w\big)_{L^2_x} + \frac{1}{2}\Omega^{kdv}_{\mathbb{S}_+}(\nu)[y] \cdot y + \mathcal{R}^{kdv}(\theta, y, w; \nu) \tag{3.6}$$

where $\omega^{kdv}(\nu) := (\omega^{kdv}_n(\nu)_{n\in\mathbb{S}_+}$,

$$\begin{aligned} &\Omega^{kdv}(D;\nu)w := \sum_{n\in\mathbb{S}^\perp} \Omega^{kdv}_n(\nu) w_n e^{\mathrm{i}2\pi n x}\,, \qquad \Omega^{kdv}_{\mathbb{S}_+}(\nu) := (\partial_{I_j}\omega^{kdv}_k(\nu))_{j,k\in\mathbb{S}_+}\,, \\ &\Omega^{kdv}_n(\nu) := \frac{1}{2\pi n}\omega^{kdv}_n(\nu, 0)\,, \quad \forall n \in \mathbb{S}^\perp, \qquad w = \sum_{n\in\mathbb{S}^\perp} w_n e^{\mathrm{i}2\pi n x} \end{aligned} \tag{3.7}$$

and $\mathcal{R}^{kdv} : \mathcal{V}^1(\delta) \times \Xi \to \mathbb{R}$ is a $\mathcal{C}^\infty$ map satisfying

$$\mathcal{R}^{kdv}(\theta, y, w; \nu) = O\big((\|y\| + \|w\|_{H^1_x})^3\big)\,, \tag{3.8}$$

and has the property that, for any $s \geq 1$, its L^2-gradient

$$(\mathcal{V}^1(\delta) \cap \mathcal{E}_s) \times \Xi \to E_s,\ (\mathfrak{x}, \nu) \mapsto \nabla\mathcal{R}^{kdv}(\mathfrak{x};\nu) = \big(\nabla_\theta \mathcal{R}^{kdv}(\mathfrak{x};\nu), \nabla_y \mathcal{R}^{kdv}(\mathfrak{x};\nu), \nabla_w \mathcal{R}^{kdv}(\mathfrak{x};\nu)\big)$$

is a $\mathcal{C}^\infty$ map as well. As a consequence

$$\nabla\mathcal{R}^{kdv}(\theta, 0, 0; \nu) = 0\,,\ d_\perp \nabla\mathcal{R}^{kdv}(\theta, 0, 0; \nu) = 0\,,\ \partial_y \nabla\mathcal{R}^{kdv}(\theta, 0, 0; \nu) = 0\,. \tag{3.9}$$

(Est1) *For any $\nu \in \Xi$, $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\mathfrak{x} \in \mathcal{V}(\delta)$, $1 \leq k \leq M$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_0$, $s \in \mathbb{N}$,*

$$\|\partial^\alpha_\nu a^\Psi_{-k}(\mathfrak{x};\nu)\|_{H^s_x} \lesssim_{s,k,\alpha} 1\,, \qquad \|d^l \partial^\alpha_\nu a^\Psi_{-k}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H^s_x} \lesssim_{s,k,l,\alpha} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_0}\,.$$

Similarly, for any $\nu \in \Xi$, $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\mathfrak{x} \in \mathcal{V}(\delta) \cap \mathcal{E}_s$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_s$, $s \in \mathbb{N}$,

$$\begin{aligned} &\|\partial^\alpha_\nu \mathcal{R}^\Psi_M(\mathfrak{x};\nu)\|_{H^{s+M+1}_x} \lesssim_{s,M,\alpha} \|w\|_{H^s_x}\,, \\ &\|d^l \partial^\alpha_\nu \mathcal{R}^\Psi_M(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H^{s+M+1}_x} \lesssim_{s,M,l,\alpha} \sum_{j=1}^l \Big(\|\widehat{\mathfrak{x}}_j\|_{E_s} \prod_{i\neq j} \|\widehat{\mathfrak{x}}_i\|_{E_0}\Big) + \|w\|_{H^s_x} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_0}\,. \end{aligned}$$

(Est2) *For any $\nu \in \Xi$, $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\mathfrak{x} \in \mathcal{V}^1(\delta)$, $1 \leq k \leq M$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_1$, $s \geq 1$,*

$$\begin{aligned} &\|\partial^\alpha_\nu a^{d\Psi^\top}_{-k}(\mathfrak{x};\nu)\|_{H^s_x} \lesssim_{s,k,\alpha} 1\,, \qquad \|d^l \partial^\alpha_\nu a^{d\Psi^\top}_{-k}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H^s_x} \lesssim_{s,k,l,\alpha} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_1}\,, \\ &\|\partial^\alpha_\nu \mathcal{A}^{d\Psi^\top}_{-k}(\mathfrak{x};\nu)\|_{\mathcal{B}(H^1_0, H^s_x)} \lesssim_{s,k,\alpha} 1\,, \quad \|d^l \partial^\alpha_\nu \mathcal{A}^{d\Psi^\top}_{-k}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{\mathcal{B}(H^1_0, H^s_x)} \lesssim_{s,k,l,\alpha} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_1}\,. \end{aligned}$$

Similarly, for any $\nu \in \Xi$, $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\mathfrak{x} \in \mathcal{V}^1(\delta) \cap \mathcal{E}_s$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_s$, $\widehat{q} \in H^s_0$, $s \geq 1$,

$$\begin{aligned} &\|\partial^\alpha_\nu \mathcal{R}^{d\Psi^\top}_M(\mathfrak{x};\nu)[\widehat{q}]\|_{E_{s+M+1}} \lesssim_{s,M,\alpha} \|\widehat{q}\|_{H^s_x} + \|w\|_{H^s_x}\|\widehat{q}\|_{H^1_x}\,, \\ &\|d^l\big(\partial^\alpha_\nu \mathcal{R}^{d\Psi^\top}_M(\mathfrak{x};\nu)[\widehat{q}]\big)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{E_{s+M+1}} \lesssim_{s,M,l,\alpha} \|\widehat{q}\|_{H^s_x} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_1} + \|\widehat{q}\|_{H^1_x} \sum_{j=1}^l \Big(\|\widehat{\mathfrak{x}}_j\|_{E_s} \prod_{i\neq j} \|\widehat{\mathfrak{x}}_i\|_{E_1}\Big) \\ &\qquad + \|\widehat{q}\|_{H^1_x}\|w\|_{H^s_x} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_1}\,. \end{aligned}$$

We now apply Theorem 3.1 to prove results concerning the extensions of $d\Psi_\nu(\mathfrak{x})^\top$ and $d\Psi_\nu(\mathfrak{x})$ to Sobolev spaces of negative order. We refer to the paragraph after (1.22) for the definitions of $\mathcal{E}_s$, E_s for negative s.

Corollary 3.2. (Extension of $d\Psi_\nu(\mathfrak{x})^\top$ and its asymptotic expansion) *Let $M \geq 1$. There exists $\sigma_M > 0$ so that for any $\mathfrak{x} \in \mathcal{V}^{\sigma_M}(\delta)$ and $\nu \in \Xi$, the operator $d\Psi_\nu(\mathfrak{x})^\top$ extends to a bounded linear operator $d\Psi_\nu(\mathfrak{x})^\top : H_0^{-M-1}(\mathbb{T}_1) \to E_{-M-1}$ and for any $\widehat{q} \in H_0^{-M-1}(\mathbb{T}_1)$, $d\Psi_\nu(\mathfrak{x})^\top[\widehat{q}]$ admits an expansion of the form*

$$d\Psi_\nu(\mathfrak{x})^\top[\widehat{q}] = \Big(0, 0, \Pi_\perp \widehat{q} + \Pi_\perp \sum_{k=1}^{M} a_{-k}^{ext}(\mathfrak{x}; \nu; d\Psi^\top)\partial_x^{-k}\widehat{q}\Big) + \mathcal{R}_M^{ext}(\mathfrak{x}; \nu; d\Psi^\top)[\widehat{q}] \tag{3.10}$$

with the following properties:

(i) For any $s \geq 0$, the maps

$$\mathcal{V}^{\sigma_M}(\delta) \times \Xi \to H^s(\mathbb{T}_1)\,, \quad (\mathfrak{x}, \nu) \mapsto a_{-k}^{ext}(\mathfrak{x}; \nu; d\Psi^\top)\,, \qquad 1 \leq k \leq M\,,$$

*are $\mathcal{C}^\infty$. They satisfy $a_{-1}^{ext}(\mathfrak{x}; \nu; d\Psi^\top) = a_{-1}^{d\Psi^\top}(\mathfrak{x}; \nu)$ (cf. Theorem 3.1-(**AE2**)) and for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, and $(\mathfrak{x}, \nu) \in \mathcal{V}^{\sigma_M}(\delta) \times \Xi$,*

$$\begin{aligned} &\|\partial_\nu^\alpha a_{-k}^{ext}(\mathfrak{x}; \nu; d\Psi^\top)\|_{H_x^s} \lesssim_{s,M,\alpha} 1\,, \\ &\|\partial_\nu^\alpha d^l a_{-k}^{ext}(\mathfrak{x}; \nu; d\Psi^\top)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^s} \lesssim_{s,M,l,\alpha} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,. \end{aligned} \tag{3.11}$$

(ii) For any $-1 \leq s \leq M+1$, the map

$$\mathcal{R}_M^{ext}(\cdot; \cdot; d\Psi^\top) : \mathcal{V}^{\sigma_M}(\delta) \times \Xi \to \mathcal{B}(H_0^{-s}(\mathbb{T}_1), E_{M+1-s})$$

is $\mathcal{C}^\infty$ and satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, $\widehat{q} \in H_0^{-s}(\mathbb{T}_1)$, and $(\mathfrak{x}, \nu) \in \mathcal{V}^{\sigma_M}(\delta) \times \Xi$,

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x}; \nu; d\Psi^\top)[\widehat{q}]\|_{E_{M+1-s}} \lesssim_{M,\alpha} \|\widehat{q}\|_{H_x^{-s}}\,, \\ &\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x}; \nu; d\Psi^\top)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l][\widehat{q}]\|_{E_{M+1-s}} \lesssim_{s,M,l,\alpha} \|\widehat{q}\|_{H_x^{-s}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,. \end{aligned} \tag{3.12}$$

(iii) For any $s \geq 1$, the map

$$\mathcal{R}_M^{ext}(\cdot; \cdot; d\Psi^\top) : \big(\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}\big) \times \Xi \to \mathcal{B}(H_0^s(\mathbb{T}_1), E_{s+M+1})$$

is $\mathcal{C}^\infty$ and satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{s+\sigma_M}$, $\widehat{q} \in H_0^s(\mathbb{T}_1)$, and $(\mathfrak{x}, \nu) \in \big(\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}\big) \times \Xi$,

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x}; \nu; d\Psi^\top)[\widehat{q}]\|_{E_{M+1+s}} \lesssim_{s,M,\alpha} \|\widehat{q}\|_{H_x^s} + \|\mathfrak{x}\|_{s+\sigma_M}\|\widehat{q}\|_{H_x^1}\,, \\ &\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x}; \nu; d\Psi^\top)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l][\widehat{q}]\|_{E_{M+1+s}} \lesssim_{s,M,l,\alpha} \|\widehat{q}\|_{H_x^s} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} \\ &\qquad + \|\widehat{q}\|_{H_x^1}\Big(\sum_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{i\neq j} \|\widehat{\mathfrak{x}}_i\|_{E_{\sigma_M}} + \|\mathfrak{x}\|_{E_{s+\sigma_M}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\Big)\,. \end{aligned} \tag{3.13}$$

Proof. By Theorem 3.1, for any $(\mathfrak{x}, \nu) \in \mathcal{V}(\delta) \times \Xi$, the differential $d\Psi_\nu(\mathfrak{x}) : E_0 \to L_0^2(\mathbb{T}_1)$ is bounded and, for any $M \geq 1$, differentiating (3.3), $d\Psi_\nu(\mathfrak{x})[\widehat{\mathfrak{x}}]$ admits the expansion for any $\widehat{\mathfrak{x}} = (\widehat{\theta}, \widehat{y}, \widehat{w}) \in E_0$ of the form

$$\begin{aligned} &d\Psi_\nu(\mathfrak{x})[\widehat{\mathfrak{x}}] = \widehat{w} + \sum_{k=1}^{M} a_{-k}^{\Psi}(\mathfrak{x}; \nu)\partial_x^{-k}\widehat{w} + \mathcal{R}_M^{(1)}(\mathfrak{x}; \nu)[\widehat{\mathfrak{x}}]\,, \\ &\mathcal{R}_M^{(1)}(\mathfrak{x}; \nu)[\widehat{\mathfrak{x}}] := \sum_{k=1}^{M} (\partial_x^{-k} w) d a_{-k}^{\Psi}(\mathfrak{x}; \nu)[\widehat{\mathfrak{x}}] + d\mathcal{R}_M^{\Psi}(\mathfrak{x}; \nu)[\widehat{\mathfrak{x}}] + d_{\theta,y}\Psi^{kdv}(\theta, \nu + y, 0)[\widehat{\theta}, \widehat{y}]\,. \end{aligned} \tag{3.14}$$

For $\sigma_M \geq M$, the map $\mathcal{R}_M^{(1)} : \mathcal{V}^{\sigma_M}(\delta) \times \Xi \to \mathcal{B}(E_0, H^{M+1}(\mathbb{T}_1))$ is $\mathcal{C}^\infty$ and satisfies, by Theorem 3.1-(**Est1**), for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $l \geq 1$,

$$\begin{aligned}
&\|\partial_\nu^\alpha \mathcal{R}_M^{(1)}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}]\|_{H_x^{M+1}} \lesssim_{M,\alpha} \|\widehat{\mathfrak{x}}\|_{E_0}\,, \\
&\|\partial_\nu^\alpha d^l \mathcal{R}_M^{(1)}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l][\widehat{\mathfrak{x}}]\|_{H_x^{M+1}} \lesssim_{M,l,\alpha} \|\widehat{\mathfrak{x}}\|_{E_0} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,.
\end{aligned} \tag{3.15}$$

Now consider the transpose operator $d\Psi_\nu(\mathfrak{x})^\top : L_0^2(\mathbb{T}_1) \to E_0$. By (3.14), for any $\widehat{q} \in L_0^2(\mathbb{T}_1)$, one has

$$d\Psi_\nu(\mathfrak{x})^\top[\widehat{q}] = \Big(0, 0, \Pi_\perp \widehat{q} + \Pi_\perp \sum_{k=1}^{M} (-1)^k \partial_x^{-k}\big(a_{-k}^{\Psi}(\mathfrak{x};\nu)\ \widehat{q}\big)\Big) + \mathcal{R}_M^{(1)}(\mathfrak{x};\nu)^\top[\widehat{q}]\,. \tag{3.16}$$

Since each function $a_{-k}^{\Psi}(\mathfrak{x};\nu)$ is $\mathcal{C}^\infty$ and $\mathcal{R}_M^{(1)}(\mathfrak{x};\nu)^\top : H^{-M-1}(\mathbb{T}_1) \to E_0$ is bounded, the right hand side of (3.16) defines a linear operator in $\mathcal{B}(H_0^{-M-1}(\mathbb{T}_1), E_{-M-1})$, which we also denote by $d\Psi_\nu(\mathfrak{x})^\top$. By (2.11), the expansion (3.16) yields one of the form (3.10) where by (3.15) and Theorem 3.1-(**Est1**), the remainder $\mathcal{R}_M^{ext}(\mathfrak{x};\nu;d\Psi^\top)$ satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, and $\widehat{q} \in H_0^{-M-1}(\mathbb{T}_1)$

$$\begin{aligned}
&\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d\Psi^\top)[\widehat{q}]\|_{E_0} \lesssim_{M,\alpha} \|\widehat{q}\|_{H_x^{-M-1}}\,, \\
&\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d\Psi^\top)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l][\widehat{q}]\|_{E_0} \lesssim_{M,l,\alpha} \|\widehat{q}\|_{H_x^{-M-1}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,.
\end{aligned} \tag{3.17}$$

The restriction of the operator $d\Psi_\nu(\mathfrak{x})^\top : H_0^{-M-1}(\mathbb{T}_1) \to E_{-M-1}$ to $H_0^1(\mathbb{T}_1)$ coincides with (3.4) and, by the uniqueness of an expansion of this form,

$$\begin{aligned}
&a_{-k}^{ext}(\mathfrak{x};\nu;d\Psi^\top) = a_{-k}^{d\Psi^\top}(\mathfrak{x};\nu)\,, \quad k = 1, \ldots, M\,, \\
&\mathcal{R}_M^{ext}(\mathfrak{x};\nu;d\Psi^\top)[\widehat{q}] = \sum_{k=1}^{M} (\partial_x^{-k} w)\mathcal{A}_{-k}^{d\Psi^\top}(\mathfrak{x};\nu)[\widehat{q}]\ + \mathcal{R}_M^{d\Psi^\top}(\mathfrak{x};\nu)[\widehat{q}]\,, \quad \forall \widehat{q} \in H_0^1(\mathbb{T}_1)\,.
\end{aligned}$$

The claimed estimates (3.11) and (3.13) then follow by Theorem 3.1-(**Est2**). In particular we have, for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, $\widehat{q} \in H_0^1(\mathbb{T}_1)$,

$$\begin{aligned}
&\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d\Psi^\top)[\widehat{q}]\|_{E_{M+2}} \lesssim_{M,\alpha} \|\widehat{q}\|_{H_x^1}\,, \\
&\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d\Psi^\top)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l][\widehat{q}]\|_{E_{M+2}} \lesssim_{M,l,\alpha} \|\widehat{q}\|_{H_x^1} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,.
\end{aligned} \tag{3.18}$$

Finally the estimates (3.12) follow by interpolation between (3.17) and (3.18). □

Corollary 3.3. (Extension of $d_\perp\Psi_\nu(\mathfrak{x})$ and its asymptotic expansion) *Let $M \geq 1$. There exists $\sigma_M > 0$ so that for any $\mathfrak{x} \in \mathcal{V}^{\sigma_M}(\delta)$ and $\nu \in \Xi$, the operator $d_\perp\Psi_\nu(\mathfrak{x})$ extends to a bounded linear operator, $d_\perp\Psi_\nu(\mathfrak{x}) : H_\perp^{-M-2}(\mathbb{T}_1) \to H_0^{-M-2}(\mathbb{T}_1)$, and for any $\widehat{w} \in H_\perp^{-M-2}(\mathbb{T}_1)$, $d_\perp\Psi_\nu(\mathfrak{x})[\widehat{w}]$ admits an expansion*

$$d_\perp\Psi_\nu(\mathfrak{x})[\widehat{w}] = \widehat{w} + \sum_{k=1}^{M} a_{-k}^{ext}(\mathfrak{x};\nu;d_\perp\Psi)\partial_x^{-k}\widehat{w} + \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{w}] \tag{3.19}$$

with the following properties:

(i) For any $s \geq 0$, the maps

$$\mathcal{V}^{\sigma_M}(\delta) \times \Xi \to H^s(\mathbb{T}_1)\,, \quad (\mathfrak{x},\nu) \mapsto a_{-k}^{ext}(\mathfrak{x};\nu;d_\perp\Psi)\,, \qquad 1 \leq k \leq M\,,$$

*are $\mathcal{C}^\infty$. They satisfy $a_{-1}^{ext}(\mathfrak{x};\nu;d_\perp\Psi) = a_{-1}^{\Psi}(\mathfrak{x};\nu)$ (cf. Theorem 3.1-(**AE1**)) and for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, and $(\mathfrak{x},\nu) \in \mathcal{V}^{\sigma_M}(\delta)\times\Xi$,*

$$\begin{aligned} &\|\partial_\nu^\alpha a_{-k}^{ext}(\mathfrak{x};\nu;d_\perp\Psi)\|_{H^s_x} \lesssim_{s,M,\alpha} 1\,,\\ &\|\partial_\nu^\alpha d^l a_{-k}^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l]\|_{H^s_x} \lesssim_{s,M,l,\alpha} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,. \end{aligned} \tag{3.20}$$

(ii) For any $0\le s\le M+2$, the map

$$\mathcal{R}_M^{ext}(\cdot,\cdot;d_\perp\Psi) : \mathcal{V}^{\sigma_M}(\delta)\times\Xi \to \mathcal{B}(H_\perp^{-s}(\mathbb{T}_1), H^{M+1-s}(\mathbb{T}_1))$$

is $\mathcal{C}^\infty$ and satisfies, for any $\alpha\in\mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, $\widehat{w}\in H_\perp^{-s}(\mathbb{T}_1)$, and $(\mathfrak{x},\nu)\in\mathcal{V}^{\sigma_M}(\delta)\times\Xi$,

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{w}]\|_{H_x^{M+1-s}} \lesssim_{M,\alpha} \|\widehat{w}\|_{H_x^{-s}}\,,\\ &\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l][\widehat{w}]\|_{H_x^{M+1-s}} \lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H_x^{-s}} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,. \end{aligned} \tag{3.21}$$

(iii) For any $s\ge 0$, the map

$$\mathcal{R}_M^{ext}(\cdot,\cdot;d_\perp\Psi) : \big(\mathcal{V}^{\sigma_M}(\delta)\cap\mathcal{E}_{s+\sigma_M}\big)\times\Xi \to \mathcal{B}(H_\perp^{s}(\mathbb{T}_1), H^{M+1+s}(\mathbb{T}_1))$$

is $\mathcal{C}^\infty$ and satisfies for any $\alpha\in\mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l \in E_{s+\sigma_M}$, $\widehat{w}\in H_\perp^{s}(\mathbb{T}_1)$, and $(\mathfrak{x},\nu)\in\big(\mathcal{V}^{\sigma_M}(\delta)\cap\mathcal{E}_{s+\sigma_M}\big)\times\Xi$,

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{w}]\|_{H_x^{M+1+s}} \lesssim_{s,M,\alpha} \|\widehat{w}\|_{H^s_x} + \|\mathfrak{x}\|_{E_{s+\sigma_M}}\|\widehat{w}\|_{L^2_x}\,,\\ &\|\partial_\nu^\alpha d^l\big(\mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{w}]\big)[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l]\|_{H_x^{M+1+s}} \lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H^s_x}\prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\\ &\quad + \|\widehat{w}\|_{L^2_x}\Big(\sum_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}}\prod_{i\ne j}\|\widehat{\mathfrak{x}}_i\|_{E_{\sigma_M}} + \|\mathfrak{x}\|_{E_{s+\sigma_M}}\prod_{j=1}^l\|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\Big)\,. \end{aligned} \tag{3.22}$$

Proof. By Theorem 3.1-(**AE2**), for any $(\mathfrak{x},\nu)\in\mathcal{V}^1(\delta)\times\Xi$, the operator $d_\perp\Psi_\nu(\mathfrak{x})^\top : H^1_0(\mathbb{T}_1)\to H^1_\perp(\mathbb{T}_1)$ is bounded and for any $M\ge 1$ and $\widehat{q}\in H^1_0(\mathbb{T}_1)$, $d_\perp\Psi_\nu(\mathfrak{x})^\top[\widehat{q}]$ admits the expansion of the form

$$\begin{aligned} d_\perp\Psi_\nu(\mathfrak{x})^\top[\widehat{q}] &= \Pi_\perp\widehat{q} + \Pi_\perp\sum_{k=1}^M a_{-k}^{d\Psi^\top}(\mathfrak{x};\nu)\partial_x^{-k}\widehat{q} + \mathcal{R}_M^{(2)}(\mathfrak{x};\nu)[\widehat{q}]\,,\\ \mathcal{R}_M^{(2)}(\mathfrak{x};\nu)[\widehat{q}] &:= \Pi_\perp\sum_{k=1}^M(\partial_x^{-k}w)\,\mathcal{A}_{-k}^{d\Psi^\top}(\mathfrak{x};\nu)[\widehat{q}] + \mathcal{R}_M^{d\Psi^\top}(\mathfrak{x};\nu)[\widehat{q}]\,. \end{aligned} \tag{3.23}$$

For $\sigma_M\ge M+1$, the map $\mathcal{R}_M^{(2)} : \mathcal{V}^{\sigma_M}(\delta)\times\Xi\to\mathcal{B}(H^1_0(\mathbb{T}_1), H_\perp^{M+2}(\mathbb{T}_1))$ is $\mathcal{C}^\infty$ and by Theorem 3.1-(**Est2**), satisfies for any $\alpha\in\mathbb{N}^{\mathbb{S}_+}$ and $\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l\in E_{\sigma_M}$

$$\begin{aligned} &\|\partial_\nu^\alpha\mathcal{R}_M^{(2)}(\mathfrak{x};\nu)[\widehat{q}]\|_{H_x^{M+2}} \lesssim_{M,\alpha} \|\widehat{q}\|_{H^1_x}\,,\\ &\|\partial_\nu^\alpha d^l\mathcal{R}_M^{(2)}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l][\widehat{q}]\|_{H_x^{M+2}} \lesssim_{M,l,\alpha} \|\widehat{q}\|_{H^1_x}\prod_{j=1}^l\|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,. \end{aligned} \tag{3.24}$$

Now consider the transpose operator $\big(d_\perp\Psi_\nu(\mathfrak{x})^\top\big)^\top : H_\perp^{-1}(\mathbb{T}_1)\to H_0^{-1}(\mathbb{T}_1)$. It defines an extension of $d_\perp\Psi_\nu(\mathfrak{x})$ to $H_\perp^{-1}(\mathbb{T}_1)$, which we denote again by $d_\perp\Psi_\nu(\mathfrak{x})$. By (3.23), for any $\widehat{w}\in H_\perp^{-1}(\mathbb{T}_1)$, one has

$$d_\perp\Psi_\nu(\mathfrak{x})[\widehat{w}] = \widehat{w} + \sum_{k=1}^M(-1)^k\partial_x^{-k}\big(a_{-k}^{d\Psi^\top}(\mathfrak{x};\nu)\widehat{w}\big) + \mathcal{R}_M^{(2)}(\mathfrak{x};\nu)^\top[\widehat{w}]\,. \tag{3.25}$$

Since each function $a_{-k}^{d\Psi^\top}(\mathfrak{x};\nu)$ is $\mathcal{C}^\infty$ and the operator $\mathcal{R}_M^{(2)}(\mathfrak{x};\nu)^\top : H_\perp^{-M-2}(\mathbb{T}_1) \to H_0^{-1}(\mathbb{T}_1)$ is bounded, the right hand side of (3.25) defines a linear operator in $\mathcal{B}(H_0^{-M-2}(\mathbb{T}_1), E_{-M-2})$, which we also denote by $d\Psi_\nu(\mathfrak{x})$. By (2.11), the expansion (3.25) yields one of the form (3.19) where by (3.24) and Theorem 3.1-(**Est2**), the remainder $\mathcal{R}_M^{ext}(\mathfrak{x};\nu;d\Psi^\top)$ satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, and $\widehat{w} \in H_0^{-M-2}(\mathbb{T}_1)$

$$
\begin{aligned}
&\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{w}]\|_{H_x^{-1}} \lesssim_{M,\alpha} \|\widehat{w}\|_{H_x^{-M-2}}\,,\\
&\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l][\widehat{w}]\|_{H_x^{-1}} \lesssim_{M,l,\alpha} \|\widehat{w}\|_{H_x^{-M-2}} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,.
\end{aligned}
\tag{3.26}
$$

The restriction of the expansion (3.25) to $L_\perp^2(\mathbb{T}_1)$ coincides with the one of $d_\perp\Psi_\nu(\mathfrak{x})[\widehat{w}]$, obtained by differentiating (3.3) (see (3.14)). It then follows from the uniqueness of an expansion of this form that

$$
\begin{aligned}
&a_{-k}^{ext}(\mathfrak{x};\nu;d_\perp\Psi) = a_{-k}^{\Psi}(\mathfrak{x};\nu)\,, \quad k=1,\dots,M\,,\\
&\mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{w}] = \sum_{k=1}^M (\partial_x^{-k}w)\, d_\perp a_{-k}^{\Psi}(\mathfrak{x};\nu)[\widehat{w}] + d_\perp \mathcal{R}_M^{\Psi}(\mathfrak{x};\nu)[\widehat{w}], \quad \forall \widehat{w} \in L_\perp^2(\mathbb{T}_1)\,.
\end{aligned}
$$

The claimed estimates (3.20) and (3.22) thus follow by Theorem 3.1-(**Est1**). In particular, for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, and $\widehat{w} \in L_\perp^2(\mathbb{T}_1)$,

$$
\begin{aligned}
&\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{w}]\|_{H_x^{M+1}} \lesssim_{M,\alpha} \|\widehat{w}\|_{L_x^2}\,,\\
&\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x};\nu;d_\perp\Psi)[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l][\widehat{w}]\|_{H_x^{M+1}} \lesssim_{M,l,\alpha} \|\widehat{w}\|_{L_x^2} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,.
\end{aligned}
\tag{3.27}
$$

The claimed estimates (3.21) are then obtained by interpolating between (3.26) and (3.27). $\square$

3.2 Expansions of linearized Hamiltonian vector fields

For any Hamiltonian of the form $P(u) = \int_{\mathbb{T}_1} f(x,u,u_x)\,dx$ with a $\mathcal{C}^\infty$-smooth density

$$
f : \mathbb{T}_1 \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}\,,\ (x,\zeta_0,\zeta_1) \mapsto f(x,\zeta_0,\zeta_1)\,,
\tag{3.28}
$$

define

$$
\mathcal{P} := P \circ \Psi_\nu\,, \quad \mathcal{P}(\theta,y,w;\nu) := P(\Psi_\nu(\theta,y,w))
\tag{3.29}
$$

where Ψ_ν is the coordinate transformation of Theorem 3.1. As a first result, we provide an expansion of the linearized Hamiltonian vector field $\partial_x d_\perp \nabla_w \mathcal{P}$.

Lemma 3.4. (Expansion of $\partial_x d_\perp \nabla_w \mathcal{P}$) *Let $P(u) = \int_{\mathbb{T}_1} f(x,u,u_x)\,dx$ with $f \in \mathcal{C}^\infty(\mathbb{T}_1 \times \mathbb{R} \times \mathbb{R})$. For any $M \in \mathbb{N}$ there is $\sigma_M > 0$ so that for any $\mathfrak{x} \in \mathcal{V}^{\sigma_M}(\delta)$ and $\nu \in \Xi$, the operator $\partial_x d_\perp \nabla_w \mathcal{P}(\mathfrak{x};\nu)$ admits an expansion of the form*

$$
\partial_x d_\perp \nabla_w \mathcal{P}(\mathfrak{x};\nu)[\cdot] = \Pi_\perp \sum_{k=0}^{M+3} a_{3-k}(\mathfrak{x};\nu;\partial_x d_\perp \nabla_w \mathcal{P})\, \partial_x^{3-k}[\cdot] + \mathcal{R}_M(\mathfrak{x};\nu;\partial_x d_\perp \nabla_w \mathcal{P})[\cdot]
\tag{3.30}
$$

with the following properties:

1. *For any $s \geq 0$, the maps*

$$
(\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}) \times \Xi \to H^s(\mathbb{T}_1)\,, \quad (\mathfrak{x};\nu) \mapsto a_{3-k}(\mathfrak{x};\nu;\partial_x d_\perp \nabla_w \mathcal{P})\,, \qquad 0 \leq k \leq M+3\,,
$$

are $\mathcal{C}^\infty$, and satisfy for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l \in E_{s+\sigma_M}$, and $(\mathfrak{x},\nu) \in \big(\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}\big) \times \Xi$,

$$
\begin{aligned}
&\|\partial_\nu^\alpha a_{3-k}(\mathfrak{x};\nu;\partial_x d_\perp \nabla_w \mathcal{P})\|_{H_x^s} \lesssim_{s,M,\alpha} 1 + \|w\|_{H_x^{s+\sigma_M}}\,,\\
&\|\partial_\nu^\alpha d^l a_{3-k}(\mathfrak{x};\nu;\partial_x d_\perp \nabla_w \mathcal{P})[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l]\|_{H_x^s} \lesssim_{s,M,l,\alpha} \sum_{j=1}^l \big(\|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{n\neq j} \|\widehat{\mathfrak{x}}_n\|_{E_{\sigma_M}}\big) + \|w\|_{H_x^{s+\sigma_M}} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,.
\end{aligned}
\tag{3.31}
$$

2. For any $0 \le s \le M+1$, the map
$$\mathcal{V}^{\sigma_M}(\delta) \times \Xi \to \mathcal{B}(H^{-s}(\mathbb{T}_1), H_\perp^{M+1-s}(\mathbb{T}_1)), \quad (\mathfrak{x}, \nu) \mapsto \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P}),$$
is $\mathcal{C}^\infty$ and satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, $(\mathfrak{x}, \nu) \in \mathcal{V}^{\sigma_M}(\delta) \times \Xi$, and $\widehat{w} \in H_\perp^{-s}(\mathbb{T}_1)$,
$$\begin{aligned}
&\|\partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})[\widehat{w}]\|_{H_x^{M+1-s}} \lesssim_{s,M,\alpha} \|\widehat{w}\|_{H_x^{-s}}, \\
&\|\partial_\nu^\alpha d^l \big(\mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})[\widehat{w}]\big)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{M+1-s}} \lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H_x^{-s}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}.
\end{aligned} \tag{3.32}$$

3. For any $s \ge 0$, the map
$$(\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}) \times \Xi \to \mathcal{B}(H^s(\mathbb{T}_1), H_\perp^{s+M+1}(\mathbb{T}_1)), \quad (\mathfrak{x}, \nu) \mapsto \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P}),$$
is $\mathcal{C}^\infty$ and satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{s+\sigma_M}$, $(\mathfrak{x}, \nu) \in (\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}) \times \Xi$, and $\widehat{w} \in H_\perp^s(\mathbb{T}_1)$,
$$\begin{aligned}
&\|\partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})[\widehat{w}]\|_{H_x^{s+M+1}} \lesssim_{s,M,\alpha} \|\widehat{w}\|_{H_x^s} + \|w\|_{H_x^{s+\sigma_M}} \|\widehat{w}\|_{L_x^2}, \\
&\|\partial_\nu^\alpha d^l \big(\mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})[\widehat{w}]\big)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{s+M+1}} \lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H_x^s} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} \\
&\qquad + \|\widehat{w}\|_{L_x^2} \Big(\|w\|_{H_x^{s+\sigma_M}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} + \sum_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{i \ne j} \|\widehat{\mathfrak{x}}_i\|_{E_{\sigma_M}}\Big).
\end{aligned} \tag{3.33}$$

Remark 3.5. *The coefficient a_3 in (3.30) can be computed as $a_3(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P}) = -(\partial_{\zeta_1}^2 f)(x, u, u_x)\big|_{u=\Psi_\nu(\mathfrak{x})}$.*

Proof. Differentiating (3.29) we have that
$$\nabla \mathcal{P}(\mathfrak{x}; \nu) = (d\Psi_\nu(\mathfrak{x}))^\top \big[\nabla P(\Psi_\nu(\mathfrak{x}))\big], \tag{3.34}$$
where, recalling (3.28),
$$\nabla P(u) = \Pi_0^\perp \big[(\partial_{\zeta_0} f)(x, u, u_x) - \big((\partial_{\zeta_1} f)(x, u, u_x)\big)_x\big] \tag{3.35}$$
and $\Pi_0^\perp$ is the L^2-orthogonal projector of $L^2(\mathbb{T}_1)$ onto $L_0^2(\mathbb{T}_1)$. By (3.34), the $w-$component $\nabla_w \mathcal{P}(\mathfrak{x}; \nu)$ of $\nabla \mathcal{P}(\mathfrak{x}; \nu)$ equals $(d_\perp \Psi_\nu(\mathfrak{x}))^\top \big[\nabla P(\Psi_\nu(\mathfrak{x}))\big]$. Differentiating it with respect to w in direction $\widehat{w}$ then yields
$$d_\perp \nabla_w \mathcal{P}(\mathfrak{x}; \nu)[\widehat{w}] = (d_\perp \Psi_\nu(\mathfrak{x}))^\top \big[d\nabla P(\Psi_\nu(\mathfrak{x}))\big[d_\perp \Psi_\nu(\mathfrak{x})[\widehat{w}]\big]\big] + \big(d_\perp (d_\perp \Psi_\nu(\mathfrak{x}))^\top [\widehat{w}]\big)\big[\nabla P(\Psi_\nu(\mathfrak{x}))\big]. \tag{3.36}$$
Analysis of the first term on the right hand side of (3.36)*:* Evaluating the differential $d\nabla P(u)$ of (3.35) at $u = \Psi_\nu(\mathfrak{x})$, one gets
$$\begin{aligned}
&d(\nabla P)(\Psi_\nu(\mathfrak{x}))[h] = \Pi_0^\perp \big(b_2(\mathfrak{x}; \nu)\partial_x^2 h + b_1(\mathfrak{x}; \nu)\partial_x h + b_0(\mathfrak{x}; \nu)h\big) \\
&b_2(\mathfrak{x}; \nu) := -\partial_{\zeta_1}^2 f(x, u, u_x)\Big|_{u=\Psi_\nu(\mathfrak{x})}, \qquad b_1(\mathfrak{x}; \nu) := (b_2(\mathfrak{x}; \nu))_x, \\
&b_0(\mathfrak{x}; \nu) := \big((\partial_{\zeta_0}^2 f)(x, u, u_x) - \big((\partial_{\zeta_0 \zeta_1}^2 f)(x, u, u_x)\big)_x\big)\big|_{u=\Psi_\nu(\mathfrak{x})}.
\end{aligned} \tag{3.37}$$
By Lemma 2.2 and Theorem 3.1 one infers that for any $s \ge 0$, the maps
$$(\mathcal{V}^3(\delta) \cap \mathcal{E}_{s+3}) \times \Xi \to H_x^s, \quad (\mathfrak{x}; \nu) \mapsto b_i(\mathfrak{x}; \nu), \quad i = 0, 1, 2,$$
are $\mathcal{C}^\infty$ and satisfy for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{s+3}$, and $(\mathfrak{x}, \nu) \in \big(\mathcal{V}^3(\delta) \cap \mathcal{E}_{s+3}\big) \times \Xi$,
$$\begin{aligned}
&\|\partial_\nu^\alpha b_i(\mathfrak{x}; \nu)\|_{H_x^s} \lesssim_{s,\alpha} 1 + \|w\|_{H_x^{s+3}}, \\
&\|\partial_\nu^\alpha d^l b_i(\mathfrak{x}; \nu)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^s} \lesssim_{s,l,\alpha} \sum_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{s+3}} \prod_{i \ne j} \|\widehat{\mathfrak{x}}_i\|_{E_3} + \|w\|_{H_x^{s+3}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_3}.
\end{aligned} \tag{3.38}$$

By Corollary 3.2 (expansion of $(d_\perp\Psi_\nu)^\top$), Corollary 3.3 (expansion of $d_\perp\Psi_\nu$), (3.38) (estimates of b_i), (3.37) (formula for $d(\nabla P)(\Psi_\nu(\mathfrak{x}))$), and Lemma 2.11 (composition), one obtains the expansion

$$\partial_x(d_\perp\Psi_\nu(\mathfrak{x}))^\top\big[d\nabla P(\Psi_\nu(\mathfrak{x}))\big[d_\perp\Psi_\nu(\mathfrak{x})[\cdot]\big]\big] = \Pi_\perp\sum_{k=0}^{M+3} a^{(1)}_{3-k}(\mathfrak{x};\nu)\partial_x^{3-k} + R_1(\mathfrak{x};\nu) \tag{3.39}$$

where $a_3^{(1)}(\mathfrak{x};\nu) = b_2(\mathfrak{x};\nu)$, the functions $a^{(1)}_{3-k}(\mathfrak{x};\nu)$, $k = 0,\dots,M+3$, and the remainder $R_1(\mathfrak{x};\nu)$ satisfy the claimed properties 1-3 of the lemma, in particular (3.31)-(3.33).

Analysis of the second term on the right hand side of (3.36)*:* Since $d\Psi_\nu(\mathfrak{x})$ is symplectic, $d\Psi_\nu(\mathfrak{x})^\top = \mathcal{J}^{-1}d\Psi_\nu(\mathfrak{x})^{-1}\partial_x$ where $\mathcal{J}$ is the Poisson operator defined in (3.2), implying that for any $\widehat{w}\in H^1_\perp(\mathbb{T}_1)$,

$$\begin{aligned} d_\perp\big(d\Psi_\nu(\mathfrak{x})^\top\big)[\widehat{w}] &= -\mathcal{J}^{-1}d\Psi_\nu(\mathfrak{x})^{-1}\big(d_\perp d\Psi_\nu(\mathfrak{x})[\widehat{w}]\big)d\Psi_\nu(\mathfrak{x})^{-1}\partial_x \\ &= -d\Psi_\nu(\mathfrak{x})^\top\partial_x^{-1}d\big(d_\perp\Psi_\nu(\mathfrak{x})[\widehat{w}]\big)\big[\mathcal{J}d\Psi_\nu(\mathfrak{x})^\top\,\cdot\,\big]\,. \end{aligned}$$

By this identity we get

$$\partial_x\big(d_\perp(d_\perp\Psi_\nu(\mathfrak{x}))^\top[\cdot]\big)\big[\nabla P(\Psi_\nu(\mathfrak{x}))\big] = -\partial_x d\Psi_\nu(\mathfrak{x})^\top\partial_x^{-1}d\big(d_\perp\Psi_\nu(\mathfrak{x})[\cdot]\big)\big[\mathcal{J}d\Psi_\nu(\mathfrak{x})^\top\nabla P(\Psi_\nu(\mathfrak{x}))\big]\,. \tag{3.40}$$

Arguing as for the first term on the right hand side of (3.36) (cf. (3.39)) one gets an expansion of the form

$$\partial_x\big(d_\perp(d_\perp\Psi_\nu(\mathfrak{x}))^\top[\cdot]\big)\big[\nabla P(\Psi_\nu(\mathfrak{x}))\big] = \Pi_\perp\sum_{k=3}^{M+3} a^{(2)}_{3-k}(\mathfrak{x};\nu)\partial_x^{3-k} + R_2(\mathfrak{x};\nu) \tag{3.41}$$

where the functions $a^{(2)}_{3-k}(\mathfrak{x};\nu)$, $k = 3,\dots,M+3$, and the remainder $R_2(\mathfrak{x};\nu)$ satisfy the claimed properties 1-3 of the lemma, in particular (3.31)-(3.33).

Conclusion: By (3.36) and the above analysis of the expansions (3.39) and (3.41), the lemma and Remark 3.5 follow. $\square$

As a second result of this section we derive an expansion for the linearized Hamiltonian vector field $\partial_x d_\perp\nabla_w\mathcal{H}^{kdv}$ where $\mathcal{H}^{kdv}(\cdot;\nu) = H^{kdv}\circ\Psi_\nu$ (cf. Theorem 3.1-(**AE3**)).

Lemma 3.6. (Expansion of $\partial_x d_\perp\nabla_w\mathcal{H}^{kdv}$) *For any $M\in\mathbb{N}$ there is $\sigma_M\geq M+1$ so that, for any $(\mathfrak{x},\nu)\in\mathcal{V}^{\sigma_M}(\delta)\times\Xi$, the operator $\partial_x d_\perp\nabla_w\mathcal{H}^{kdv}(\mathfrak{x};\nu)$ admits an expansion of the form*

$$\begin{aligned} &\partial_x d_\perp\nabla_w\mathcal{H}^{kdv}(\mathfrak{x};\nu)[\cdot] = \partial_x\Omega^{kdv}(D;\nu)[\cdot] + \partial_x d_\perp\nabla_w\mathcal{R}^{kdv}(\mathfrak{x};\nu)[\cdot]\,, \\ &\partial_x d_\perp\nabla_w\mathcal{R}^{kdv}(\mathfrak{x};\nu)[\cdot] = \Pi_\perp\sum_{k=0}^{M+1} a_{1-k}(\mathfrak{x};\nu;\partial_x d_\perp\nabla_w\mathcal{R}^{kdv})\,\partial_x^{1-k}[\cdot] + \mathcal{R}_M(\mathfrak{x};\nu;\partial_x d_\perp\nabla_w\mathcal{R}^{kdv})[\cdot]\,, \end{aligned} \tag{3.42}$$

with the following properties:

1. *For any $s\geq 0$, the maps*

$$(\mathcal{V}^{\sigma_M}(\delta)\cap\mathcal{E}_{s+\sigma_M})\times\Xi\to H^s(\mathbb{T}_1),\ (\mathfrak{x},\nu)\mapsto a_{1-k}(\mathfrak{x};\nu;\partial_x d_\perp\nabla_w\mathcal{R}^{kdv})\,,\qquad 0\leq k\leq M+1\,,$$

are $\mathcal{C}^\infty$ and satisfy for any $\alpha\in\mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l\in E_{s+\sigma_M}$, and $(\mathfrak{x},\nu)\in(\mathcal{V}^{\sigma_M}(\delta)\cap\mathcal{E}_{s+\sigma_M})\times\Xi$,

$$\begin{aligned} &\|\partial_\nu^\alpha a_{1-k}(\mathfrak{x};\nu;\partial_x d_\perp\nabla_w\mathcal{R}^{kdv})\|_{H^s_x}\lesssim_{s,k,\alpha}\|y\| + \|w\|_{H_x^{s+\sigma_M}}\,, \\ &\|d^l\partial_\nu^\alpha a_{1-k}(\mathfrak{x};\nu;\partial_x d_\perp\nabla_w\mathcal{R}^{kdv})[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l]\|_{H^s_x}\lesssim_{s,k,l,\alpha}\sum_{j=1}^{l}\big(\|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}}\prod_{n\neq j}\|\widehat{\mathfrak{x}}_n\|_{E_{\sigma_M}}\big) \\ &\qquad\qquad + (\|y\| + \|w\|_{H_x^{s+\sigma_M}})\prod_{j=1}^{l}\|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,. \end{aligned} \tag{3.43}$$

2. For any $0 \leq s \leq M+1$, the map

$$\mathcal{R}_M(\cdot;\cdot;\partial_x d_\perp \nabla_w \mathcal{R}^{kdv}) : \mathcal{V}^{\sigma_M}(\delta) \times \Xi \to \mathcal{B}(H_\perp^{-s}(\mathbb{T}_1), H_\perp^{M+1-s}(\mathbb{T}_1))$$

is $\mathcal{C}^\infty$ and satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, $(\mathfrak{x}, \nu) \in \mathcal{V}^{\sigma_M}(\delta) \times \Xi$, and $\widehat{w} \in H_\perp^{-s}(\mathbb{T}_1)$,

$$\|\partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x};\nu;\partial_x d_\perp \nabla \mathcal{R}^{kdv})[\widehat{w}]\|_{H_x^{M+1-s}} \lesssim_{s,M,\alpha} (\|y\| + \|w\|_{H_x^{\sigma_M}})\|\widehat{w}\|_{H_x^{-s}}\,, \tag{3.44}$$

$$\|d^l \partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x};\nu;\partial_x d_\perp \nabla \mathcal{R}^{kdv})[\widehat{w}][\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{M+1-s}} \lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H_x^{-s}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,. \tag{3.45}$$

3. For any $s \geq 0$, the map

$$\mathcal{R}_M(\cdot;\cdot;\partial_x d_\perp \nabla_w \mathcal{R}^{kdv}) : (\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}) \times \Xi \to \mathcal{B}(H_\perp^{s}(\mathbb{T}_1), H_\perp^{s+M+1}(\mathbb{T}_1))\,,$$

is $\mathcal{C}^\infty$ and satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{s+\sigma_M}$, $(\mathfrak{x}, \nu) \in (\mathcal{E}_{s+\sigma_M} \cap \mathcal{V}^{\sigma_M}(\delta)) \times \Xi$, and $\widehat{w} \in H_\perp^{s}(\mathbb{T}_1)$,

$$\begin{aligned}&\|\partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x};\nu;\partial_x d_\perp \nabla \mathcal{R}^{kdv})[\widehat{w}]\|_{H_x^{s+M+1}} \\ &\lesssim_{s,M,\alpha} (\|y\| + \|w\|_{H_x^{s+\sigma_M}})\|\widehat{w}\|_{L_x^2} + (\|y\| + \|w\|_{H_x^{\sigma_M}})\|\widehat{w}\|_{H_x^{s}}\,,\end{aligned} \tag{3.46}$$

$$\begin{aligned}\|d^l \partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x};\nu;\partial_x d_\perp \nabla \mathcal{R}^{kdv})[\widehat{w}][\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{s+M+1}} &\lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H_x^{s}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} \\ + \|\widehat{w}\|_{L_x^2} \sum_{j=1}^{l} \Big(\|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{n \neq j} \|\widehat{\mathfrak{x}}_n\|_{E_{\sigma_M}} \Big) &+ \|\widehat{w}\|_{L_x^2} \|w\|_{H_x^{s+\sigma_M}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\,.\end{aligned} \tag{3.47}$$

Proof. Differentiating $\mathcal{H}^{kdv}(\mathfrak{x};\nu) = H^{kdv}(\Psi_\nu(\mathfrak{x}))$, we get

$$\nabla_w \mathcal{H}^{kdv}(\mathfrak{x};\nu) = (d_\perp \Psi_\nu(\mathfrak{x}))^\top \big[\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big] \tag{3.48}$$

where, recalling (1.4),

$$\nabla H^{kdv}(u) = \Pi_0^\perp\big(3u^2 - u_{xx}\big) \tag{3.49}$$

and $\Pi_0^\perp$ is the L^2-orthogonal projector onto $L_0^2(\mathbb{T}_1)$. Differentiating (3.48) with respect to w in direction $\widehat{w}$ we get

$$\begin{aligned}&d_\perp \nabla_w \mathcal{H}^{kdv}(\mathfrak{x};\nu)[\widehat{w}] = \\ &(d_\perp \Psi_\nu(\mathfrak{x}))^\top \big[d\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))[d_\perp \Psi_\nu(\mathfrak{x})[\widehat{w}]]\big] + \big(d_\perp (d_\perp \Psi_\nu(\mathfrak{x}))^\top[\widehat{w}]\big)\big[\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big]\,.\end{aligned} \tag{3.50}$$

On the other hand, by (3.6)

$$d_\perp \nabla_w \mathcal{H}^{kdv}(\mathfrak{x};\nu) = \Omega^{kdv}(D;\nu) + d_\perp \nabla_w \mathcal{R}^{kdv}(\mathfrak{x};\nu)$$

and by (3.9) $d_\perp \nabla_w \mathcal{R}^{kdv}(\theta, 0, 0;\nu) = 0$, implying that

$$\begin{aligned}&d_\perp \nabla_w \mathcal{H}^{kdv}(\theta, 0, 0;\nu) = \Omega^{kdv}(D;\nu)\,, \\ &d_\perp \nabla_w \mathcal{R}^{kdv}(\mathfrak{x};\nu) = d_\perp \nabla_w \mathcal{H}^{kdv}(\theta, y, w;\nu) - d_\perp \nabla_w \mathcal{H}^{kdv}(\theta, 0, 0;\nu)\,.\end{aligned} \tag{3.51}$$

In order to obtain the expansion (3.42) it thus suffices to expand $d_\perp \nabla_w \mathcal{H}^{kdv}(\theta, y, w;\nu))[\widehat{w}]$ and then subtract from it the expansion of $d_\perp \nabla_w \mathcal{H}^{kdv}(\theta, 0, 0;\nu))[\widehat{w}]$. We analyze separately the two terms in (3.50).
Analysis of the first term on the right hand side of (3.50)*:* Evaluating the differential $d\nabla H^{kdv}(u)$ at $u = \Psi_\nu(\mathfrak{x})$, one gets

$$d(\nabla H^{kdv})(\Psi_\nu(\mathfrak{x}))[h] = \Pi_0^\perp\big(-\partial_x^2 h + b_0(\mathfrak{x};\nu)h\big)\,, \quad b_0(\mathfrak{x};\nu) := 6\Psi_\nu(\mathfrak{x})\,. \tag{3.52}$$

By Theorem 3.1-(**AE1**) and the estimates (**Est1**), the function $b_0(\mathfrak{x};\nu)$ satisfies, for any $s \geq 0$,

$$\begin{aligned}
&\|\partial_\nu^\alpha b_0(\mathfrak{x};\nu)\|_{H^s_x} \lesssim_{s,\alpha} 1 + \|w\|_{H^{s+1}_x} \,, \\
&\|\partial_\nu^\alpha d^l b_0(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}_1,\ldots,\widehat{\mathfrak{x}}_l]\|_{H^s_x} \lesssim_{s,l,\alpha} \sum_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{s+1}} \prod_{i\neq j} \|\widehat{\mathfrak{x}}_i\|_{E_1} + \|w\|_{H^{s+1}_x} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_1} \,.
\end{aligned} \tag{3.53}$$

By Corollary 3.2 (expansion of $(d_\perp \Psi_\nu)^\top$), Corollary 3.3 (expansion of $d_\perp\Psi_\nu$), (3.53) (estimates of b_0), (3.52) $\big($formula for $d(\nabla H^{kdv})(\Psi_\nu(\mathfrak{x}))\big)$, and Lemma 2.11 (composition), one obtains the expansion

$$\begin{aligned}
&\partial_x (d_\perp\Psi_\nu(\mathfrak{x}))^\top \big[d\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big[d_\perp\Psi_\nu(\mathfrak{x})[\cdot]\big]\big] \\
&= \Pi_\perp\big(-\partial_x^3 - (a_{-1}^{\Psi}(\mathfrak{x};\nu) + a_{-1}^{d\Psi^\top}(\mathfrak{x};\nu))\partial_x^2 + \sum_{k=0}^{M+1} a^{(1)}_{1-k}(\mathfrak{x};\nu)\partial_x^{1-k}\big) + R_1(\mathfrak{x};\nu) \\
&\overset{(3.5)}{=} \Pi_\perp\big(-\partial_x^3 + \sum_{k=0}^{M+1} a^{(1)}_{1-k}(\mathfrak{x};\nu)\partial_x^{1-k}\big) + R_1(\mathfrak{x};\nu)
\end{aligned} \tag{3.54}$$

where the functions $a^{(1)}_{1-k}(\mathfrak{x};\nu)$, $k = 0,\ldots,M+1$ and the remainder $R_1(\mathfrak{x};\nu)$ satisfy the properties 1–3 stated in Lemma 3.4, in particular (3.31)-(3.33).

Analysis of the second term on the right hand side of (3.50)*:* By (3.40) one has

$$\partial_x\big(d_\perp(d_\perp\Psi_\nu(\mathfrak{x}))^\top[\cdot]\big)\big[\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big] = -\partial_x d\Psi_\nu(\mathfrak{x})^\top\partial_x^{-1} d\big(d_\perp\Psi_\nu(\mathfrak{x})[\cdot]\big)\big[\mathcal{J} d\Psi_\nu(\mathfrak{x})^\top \nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big] \,.$$

Arguing as for the first term on the right hand side of (3.50) one obtains an expansion of the form

$$\partial_x\big(d_\perp(d_\perp\Psi_\nu(\mathfrak{x}))^\top[\cdot]\big)\big[\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big] = \Pi_\perp \sum_{k=0}^{M+1} a^{(2)}_{1-k}(\mathfrak{x};\nu)\partial_x^{1-k} + R_2(\mathfrak{x};\nu) \tag{3.55}$$

where $a^{(2)}_1(\mathfrak{x};\nu) = 0$ (cf. (3.10)) and where the functions $a^{(2)}_{1-k}(\mathfrak{x};\nu)$, $k = 1,\ldots,M+1$ and the remainder $R_2(\mathfrak{x};\nu)$ satisfy the properties 1-3 of Lemma 3.4, in particular (3.31)-(3.33).

Conclusion: Combining (3.50), (3.51), (3.54), and (3.55) one obtains the claimed expansion (3.42) with

$$\begin{aligned}
&a_{1-k}(\mathfrak{x};\nu;\partial_x d_\perp\nabla_w\mathcal{R}^{kdv}) := a^{(1)}_{1-k}(\mathfrak{x};\nu) - a^{(1)}_{1-k}(\theta,0,0;\nu) + a^{(2)}_{1-k}(\mathfrak{x};\nu) - a^{(2)}_{1-k}(\theta,0,0;\nu) \\
&\mathcal{R}_M(\mathfrak{x};\nu;\partial_x d_\perp\nabla\mathcal{R}^{kdv}) := R_1(\mathfrak{x};\nu) - R_1(\theta,0,0;\nu) + R_2(\mathfrak{x};\nu) - R_2(\theta,0,0;\nu)\,.
\end{aligned}$$

Since $a^{(1)}_{1-k}(\mathfrak{x};\nu)$, $R_1(\mathfrak{x};\nu)$, and $a^{(2)}_{1-k}(\mathfrak{x};\nu)$, $R_2(\mathfrak{x};\nu)$ satisfy properties 1-3 of Lemma 3.4, in particular (3.31)-(3.33), the claimed estimates (3.43)-(3.47) then follow by the mean value theorem. □

3.3 Frequencies of KdV

In this section we record properties of the KdV frequencies ω_n^{kdv} used in this paper. In Section 6 we need to analyze $\partial_x\Omega^{kdv}(D;I)$. Recall that by (3.7), $\Omega^{kdv}(D;I)$ is defined for $I \in \Xi \subset \mathbb{R}^{\mathbb{S}_+}_{>0}$. Actually, it is defined on all of $\mathbb{R}^{\mathbb{S}_+}_{>0}$ (cf. (1.10)) and according to [15, Lemma 4.1] $\partial_x\Omega^{kdv}(D;I)$ can be written as

$$\partial_x\Omega^{kdv}(D;I) = -\partial_x^3 + Q^{kdv}_{-1}(D;I) \tag{3.56}$$

where $Q^{kdv}_{-1}(D;I)$ is a family of Fourier multiplier operators of order -1 with an expansion in homogeneous components up to any order.

Lemma 3.7. *For any $M \in \mathbb{N}$ and $I \in \mathbb{R}^{\mathbb{S}_+}_{>0}$, $Q^{kdv}_{-1}(D;I)$ admits an expansion of the form*

$$Q^{kdv}_{-1}(D;I) = \Omega^{kdv}_{-1}(D;I) + \mathcal{R}_M(D;I;Q^{kdv}_{-1})\,, \quad \Omega^{kdv}_{-1}(\xi;I) = \sum_{k=1}^{M} a_{-k}(I;\Omega^{kdv}_{-1})\chi_0(\xi)(\mathrm{i}2\pi\xi)^{-k}\,, \tag{3.57}$$

where the functions $a_{-k}(I;\Omega^{kdv}_{-1})$ are real analytic and bounded on compact subsets of $\mathbb{R}^{\mathbb{S}_+}_{>0}$, $a_{-k}(I;\Omega^{kdv}_{-1})$ vanishes identially for k even, and $\mathcal{R}_M(D;I;Q^{kdv}_{-1})$ is a Fourier multiplier operator with multipliers

$$\mathcal{R}_M(n;I;Q^{kdv}_{-1}) = \frac{\mathcal{R}^{\omega_n}_M(I)}{(2\pi n)^{M+1}}\,, \qquad \mathcal{R}_M(-n;I;Q^{kdv}_{-1}) = -\mathcal{R}_M(n;I;Q^{kdv}_{-1})\,, \quad \forall n \in \mathbb{S}^\perp_+\,, \tag{3.58}$$

where the functions $I \mapsto \mathcal{R}^{\omega_n}_M(I)$ are real analytic and satisfy, for any $j \in \mathbb{S}_+$, $\beta \in \mathbb{N}$,

$$\sup_{n\in\mathbb{S}^\perp} |\mathcal{R}^{\omega_n}_M(I)| \le C_M\,, \quad \sup_{n\in\mathbb{S}^\perp} |\partial^\beta_{I_j}\mathcal{R}^{\omega_n}_M(I)| \le C_{M,\beta}\,,$$

uniformly on compact subsets of $\mathbb{R}^{\mathbb{S}_+}_{>0}$.

Proof. The result follows by [15, Lemma C.7]. □

In Section 7, we shall use the following asymptotics of the KdV frequencies

$$\omega^{kdv}_n(I,0) - (2\pi n)^3 = O(n^{-1})\,, \quad n\,\partial_I\omega^{kdv}_n(I,0) = O(1)\,, \tag{3.59}$$

uniformly on compact sets of actions $I \in \mathbb{R}^{\mathbb{S}_+}_{>0}$.

Lemma 3.8. *([16, Proposition 15.5])* **(Non-degeneracy of KdV frequencies)** *For any finite subset $\mathbb{S}_+ \subset \mathbb{N}$ the following holds on $\mathbb{R}^{\mathbb{S}_+}_{>0}$:*
(i) The map $I \mapsto \det\big((\partial_{I_k}\omega^{kdv}_n(I,0))_{k,n\in\mathbb{S}_+}\big)$ is real analytic and does not vanish identically.
(ii) For any $\ell \in \mathbb{Z}^{\mathbb{S}_+}$and $j,k \in \mathbb{S}^\perp$ with $(\ell,j,k) \neq (0,j,j)$, the following functions are real analytic and do not vanish identically,

$$\sum_{n\in\mathbb{S}_+} \ell_n\omega^{kdv}_n + \omega^{kdv}_j \neq 0\,, \qquad \sum_{n\in\mathbb{S}_+} \ell_n\omega^{kdv}_n + \omega^{kdv}_j - \omega^{kdv}_k \neq 0\,. \tag{3.60}$$

Remark 3.9. *It was shown in [10] that for any $I \in \mathbb{R}^{\mathbb{S}_+}_{>0}$, $\det\big((\partial_{I_k}\omega^{kdv}_n(I,0))_{k,n\in\mathbb{S}_+}\big) \neq 0$.*

4 Nash-Moser theorem

In the symplectic variables $(\theta,y,w) \in \mathcal{V}(\delta)\cap\mathcal{E}_s$ defined by Theorem 3.1, with symplectic 2-form given by (3.1), the Hamiltonian equation (1.1) reads

$$\partial_t\theta = -\nabla_y\mathcal{H}_\varepsilon\,, \qquad \partial_t y = \nabla_\theta\mathcal{H}_\varepsilon\,, \qquad \partial_t w = \partial_x\nabla_w\mathcal{H}_\varepsilon\,, \tag{4.1}$$

where $\mathcal{H}_\varepsilon := H_\varepsilon \circ \Psi_\nu$ and H_ε given by (1.4). More explicitly,

$$\begin{aligned} &\mathcal{H}_\varepsilon(\theta,y,w;\nu) = \mathcal{H}^{kdv}(\theta,y,w;\nu) + \varepsilon\mathcal{P}(\theta,y,w;\nu)\,, \\ &\mathcal{H}^{kdv} = H^{kdv}\circ\Psi_\nu\,, \quad \mathcal{P} = P\circ\Psi_\nu\,, \quad \nu \in \Xi\,, \end{aligned} \tag{4.2}$$

where $\mathcal{H}^{kdv}(\theta,y,w;\nu)$ has the normal form expansion (3.6). We denote by $X_{\mathcal{H}_\varepsilon}$ the Hamiltonian vector field associated to $\mathcal{H}_\varepsilon$. For $\varepsilon = 0$, the Hamiltonian system (4.1) possesses, for any value of the parameter $\nu \in \Xi$, the invariant torus $\mathbb{T}^{\mathbb{S}_+} \times \{0\} \times \{0\}$, filled by quasi-periodic finite gap solutions of the KdV equation with frequency vector $\omega^{kdv}(\nu) := (\omega^{kdv}_n(\nu,0))_{n\in\mathbb{S}_+}$ introduced in (1.11).

By our choice of Ξ, the map $-\omega^{kdv} : \Xi \to \Omega := -\omega^{kdv}(\Xi)$ is a real analytic diffeomorphism. In the sequel, we consider ν as a function of the parameter $\omega \in \Omega$, namely

$$\nu \equiv \nu(\omega) := (\omega^{kdv})^{-1}(-\omega)\ . \tag{4.3}$$

For simplicity we often will not record the dependence of the Hamiltonian $\mathcal{H}_\varepsilon$ on $\nu = (\omega^{kdv})^{-1}(-\omega)$.

Consider the set of diophantine frequencies in Ω,

$$\mathtt{DC}(\gamma,\tau) := \Big\{\omega \in \Omega : |\omega \cdot \ell| \geq \frac{\gamma}{\langle \ell \rangle^{\tau}}, \quad \forall \ell \in \mathbb{Z}^{\mathbb{S}_+} \setminus \{0\}\Big\}. \tag{4.4}$$

For any torus embedding $\mathbb{T}^{\mathbb{S}_+} \to \mathcal{V}(\delta) \cap \mathcal{E}_s$, $\varphi \mapsto (\theta(\varphi), y(\varphi), w(\varphi))$, close to the identity, consider its lift

$$\breve{\iota} : \mathbb{R}^{\mathbb{S}_+} \to \mathbb{R}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times H^s_\perp(\mathbb{T}_1), \quad \breve{\iota}(\varphi) = (\varphi, 0, 0) + \iota(\varphi), \tag{4.5}$$

where $\iota(\varphi) = (\Theta(\varphi), y(\varphi), w(\varphi))$, with $\Theta(\varphi) := \theta(\varphi) - \varphi$, is $(2\pi\mathbb{Z})^{\mathbb{S}_+}$ periodic.

We look for a torus embedding $\breve{\iota}$ such that $\mathcal{F}_\omega(\iota, \zeta) = 0$ where

$$\mathcal{F}_\omega(\iota,\zeta) := \begin{pmatrix} \omega \cdot \partial_\varphi \theta(\varphi) + (\nabla_y \mathcal{H}_\varepsilon)(\breve{\iota}(\varphi)) \\ \omega \cdot \partial_\varphi y(\varphi) - (\nabla_\theta \mathcal{H}_\varepsilon)(\breve{\iota}(\varphi)) - \zeta \\ \omega \cdot \partial_\varphi w(\varphi) - \partial_x (\nabla_w \mathcal{H}_\varepsilon)(\breve{\iota}(\varphi)) \end{pmatrix}. \tag{4.6}$$

The additional variable $\zeta \in \mathbb{R}^{\mathbb{S}_+}$ is introduced in order to control the average of the y-component of the linearized Hamiltonian equations – see Section 5. Actually any invariant torus for $X_{\mathcal{H}_{\varepsilon,\zeta}} = X_{\mathcal{H}_\varepsilon} + (0,\zeta,0)$ with modified Hamiltonian

$$\mathcal{H}_{\varepsilon,\zeta}(\theta, y, w) := \mathcal{H}_\varepsilon(\theta, y, w) + \zeta \cdot \theta, \quad \zeta \in \mathbb{R}^{\mathbb{S}_+}, \tag{4.7}$$

is invariant for $X_{\mathcal{H}_\varepsilon}$, see (5.5). Notice that $\mathcal{H}_{\varepsilon,\zeta}$ is not periodic in θ, but that its Hamiltonian vector field is. The Lipschitz Sobolev norm of the periodic part $\iota(\varphi) = (\Theta(\varphi), y(\varphi), w(\varphi))$ of the embedded torus (4.5) is

$$\|\iota\|_s^{\mathrm{Lip}(\gamma)} := \|\Theta\|_s^{\mathrm{Lip}(\gamma)} + \|y\|_s^{\mathrm{Lip}(\gamma)} + \|w\|_s^{\mathrm{Lip}(\gamma)}$$

where $\|w\|_s^{\mathrm{Lip}(\gamma)}$ is the Lipschitz Sobolev norm introduced in (2.1) and

$$\|\Theta\|_s^{\mathrm{Lip}(\gamma)} \equiv \|\Theta\|_{H^s_\varphi}^{\mathrm{Lip}(\gamma)} := \|\Theta\|_{H^s(\mathbb{T}^{\mathbb{S}_+}, \mathbb{R}^{\mathbb{S}_+})}^{\mathrm{Lip}(\gamma)}, \qquad \|y\|_s^{\mathrm{Lip}(\gamma)} \equiv \|y\|_{H^s_\varphi}^{\mathrm{Lip}(\gamma)} := \|y\|_{H^s(\mathbb{T}^{\mathbb{S}_+}, \mathbb{R}^{\mathbb{S}_+})}^{\mathrm{Lip}(\gamma)}. \tag{4.8}$$

Theorem 4.1. (Nash-Moser) *There exist $\bar{s} > (|\mathbb{S}_+| + 1)/2$ and $\varepsilon_0 > 0$ so that for any $0 < \varepsilon \leq \varepsilon_0$, there is a measurable subset $\Omega_\varepsilon \subseteq \Omega$ satisfying*

$$\lim_{\varepsilon \to 0} \frac{\mathrm{meas}(\Omega_\varepsilon)}{\mathrm{meas}(\Omega)} = 1 \tag{4.9}$$

and for any $\omega \in \Omega_\varepsilon$, there exists a torus embedding $\breve{\iota}_\omega$ as in (4.5) which satisfies the estimate

$$\|\breve{\iota}_\omega - (\varphi, 0, 0)\|_{\bar{s}}^{\mathrm{Lip}(\gamma)} = O(\varepsilon \gamma^{-2}), \qquad \gamma = \varepsilon^{\mathfrak{a}}, \ 0 < \mathfrak{a} \ll 1,$$

and solves

$$\omega \cdot \partial_\varphi \breve{\iota}_\omega(\varphi) - X_{\mathcal{H}_\varepsilon}(\breve{\iota}_\omega(\varphi)) = 0.$$

As a consequence the embedded torus $\breve{\iota}_\omega(\mathbb{T}^{\mathbb{S}_+})$ is invariant for the Hamiltonian vector field $X_{\mathcal{H}_\varepsilon(\cdot;\nu)}$ with $\nu = (\omega^{kdv})^{-1}(-\omega)$, and it is filled by quasi-periodic solutions of (4.1) with frequency vector $\omega \in \Omega_\varepsilon$. Furthermore, the quasi-periodic solution $\breve{\iota}_\omega(\omega t) = \omega t + \iota_\omega(\omega t)$ is linearly stable.

Theorem 4.1 is proved in Section 8. The main issue concerns the construction of an approximate right inverse of the linearized operator $d_{\iota,\zeta}\mathcal{F}_\omega(\iota,\zeta)$ at an approximate solution. This construction is carried out in Sections 5, 6 and 7.

Along the proof we shall use the following tame estimates of the Hamiltonian vector field $X_{\mathcal{H}_\varepsilon}$ with respect to the norm $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$. Recalling the expansion (3.6) provided in Theorem 3.1, and the definition of $\mathcal{P}$ in (3.29), we decompose the Hamiltonian $\mathcal{H}_\varepsilon$ defined in (4.2) as

$$\begin{aligned} &\mathcal{H}_\varepsilon = \mathcal{N} + \mathcal{P}_\varepsilon \quad \text{where} \\ &\mathcal{N}(y, w; \nu) := \omega^{kdv}(\nu) \cdot y + \frac{1}{2}\Omega^{kdv}_{\mathbb{S}_+}(\nu)[y] \cdot y + \frac{1}{2}\big(\Omega^{kdv}(D;\nu) w\,,\, w\big)_{L^2_x}, \quad \mathcal{P}_\varepsilon := \mathcal{R}^{kdv} + \varepsilon \mathcal{P}. \end{aligned} \tag{4.10}$$

Lemma 4.2. *There exists $\sigma_1 = \sigma_1(\mathbb{S}_+) > 0$ so that for any $s \geq 0$, any torus embedding $\breve{\iota}$ of the form* (4.5) *with $\|\iota\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)} \leq \delta$, and any maps $\widehat{\iota}, \widehat{\iota}_1, \widehat{\iota}_2 : \mathbb{T}^{\mathbb{S}_+} \to E_s$, the following tame estimates hold:*

$$\begin{aligned}
\|X_{\mathcal{P}_\varepsilon}(\breve{\iota})\|_s^{\mathrm{Lip}(\gamma)} &\lesssim_s \varepsilon(1 + \|\iota\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)}) + \|\iota\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)}\|\iota\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)} , \\
\|dX_{\mathcal{P}_\varepsilon}(\breve{\iota})[\widehat{\iota}]\|_s^{\mathrm{Lip}(\gamma)} &\lesssim_s (\varepsilon + \|\iota\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)})\|\widehat{\iota}\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)}\|\widehat{\iota}\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)} , \\
\|d^2X_{\mathcal{H}_\varepsilon}(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|_s^{\mathrm{Lip}(\gamma)} &\lesssim_s \|\widehat{\iota}_1\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)}\|\widehat{\iota}_2\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)} + \|\widehat{\iota}_1\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)}\|\widehat{\iota}_2\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)}(\|\widehat{\iota}_1\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)}\|\widehat{\iota}_2\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)} .
\end{aligned}$$

Proof. Note that $X_{\mathcal{P}_\varepsilon} = \varepsilon X_{\mathcal{P}} + X_{\mathcal{R}^{kdv}}$ and $d^2X_{\mathcal{H}_\varepsilon} = d^2X_{\mathcal{N}} + d^2X_{\mathcal{P}_\varepsilon}$. The claimed estimates then follow from estimates of $\varepsilon X_{\mathcal{P}}$, obtained from Lemmata 3.4, 2.23, 2.24, and from estimates of $X_{\mathcal{R}^{kdv}}$ obtained from Lemmata 3.6, 2.23, 2.24, and the mean value theorem. □

5 Approximate inverse

In order to implement a convergent Nash-Moser scheme that leads to a solution of $\mathcal{F}_\omega(\iota, \zeta) = 0$ (cf. (4.6)) we construct an *almost-approximate right inverse* (see Theorem 5.6) of the linearized operator

$$d_{\iota,\zeta}\mathcal{F}_\omega(\iota, \zeta)[\widehat{\iota}, \widehat{\zeta}] = \omega \cdot \partial_\varphi \widehat{\iota} - d_\iota X_{\mathcal{H}_\varepsilon}(\breve{\iota})[\widehat{\iota}] - (0, \widehat{\zeta}, 0) \tag{5.1}$$

where $\mathcal{H}_\varepsilon = \mathcal{N} + \mathcal{P}_\varepsilon$ is the Hamiltonian in (4.10). Note that the perturbation $\mathcal{P}_\varepsilon$ and the differential $d_{\iota,\zeta}\mathcal{F}_\omega(\iota, \zeta)$ are independent of ζ. In the sequel, we will often write $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$ instead of $d_{\iota,\zeta}\mathcal{F}_\omega(\iota, \zeta)$.

Since the θ, y, and w components of $d_\iota X_{\mathcal{H}_\varepsilon}(\breve{\iota}(\varphi))[\widehat{\iota}]$ are all coupled, inverting the linear operator $d_{\iota,\zeta}\mathcal{F}_\omega(\iota, \zeta)$ in (5.1) is intricate. As a first step, we implement the approach developed in [3], [7], [9], to approximately reduce $d_{\iota,\zeta}\mathcal{F}_\omega(\iota, \zeta)$ to a triangular form – see (5.29) below.

Along this section we assume the following hypothesis, which is verified by the approximate solutions obtained at each step of the Nash-Moser Theorem 8.1.

- **Ansatz.** *The map $\omega \mapsto \iota(\omega) := \breve{\iota}(\varphi; \omega) - (\varphi, 0, 0)$ is Lipschitz continuous with respect to $\omega \in \Omega$, and, for $\gamma \in (0, 1)$, $\mu_0 := \mu_0(\tau, \mathbb{S}_+) > 0$ (with τ being specified later (cf. Section 8))*

  $$\|\iota\|_{\mu_0}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2} , \quad \|Z\|_{s_0}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon , \tag{5.2}$$

  where Z is the "error function" defined by

  $$Z(\varphi) := (Z_1, Z_2, Z_3)(\varphi) := \mathcal{F}_\omega(\iota, \zeta)(\varphi) = \omega \cdot \partial_\varphi \breve{\iota}(\varphi) - X_{\mathcal{H}_\varepsilon}(\breve{\iota}(\varphi)) - (0, \zeta, 0) . \tag{5.3}$$

We first notice that the 2-form $\mathcal{W}$ given in (3.1) is

$$\mathcal{W} := \Big(\sum\nolimits_{j \in \mathbb{S}_+} dy_j \wedge d\theta_j\Big) \oplus \mathcal{W}_\perp = d\Lambda$$

where Λ is the Liouville 1-form

$$\Lambda_{(\theta,y,w)}[\widehat{\theta}, \widehat{y}, \widehat{w}] := \sum\nolimits_{j \in \mathbb{S}_+} y_j \widehat{\theta}_j + \frac{1}{2}\big(\partial_x^{-1} w \, , \, \widehat{w}\big)_{L_x^2} . \tag{5.4}$$

Arguing as in [3, Lemma 6.1], one obtains

$$|\zeta|^{\mathrm{Lip}(\gamma)} \lesssim \|Z\|_{s_0}^{\mathrm{Lip}(\gamma)} . \tag{5.5}$$

An invariant torus $\breve{\iota}$ with Diophantine flow is isotropic, meaning that the pull-back $\breve{\iota}^*\Lambda$ of the 1-form Λ is closed, or equivalently that the pull back $\breve{\iota}^*\mathcal{W}$ satisfies $\breve{\iota}^*\mathcal{W} = \breve{\iota}^* d\Lambda = d\breve{\iota}^*\Lambda = 0$ (cf. [7]). For an approximately invariant torus embedding $\breve{\iota}$, the 1-form

$$\breve{\iota}^*\Lambda = \sum\nolimits_{k \in \mathbb{S}_+} a_k(\varphi) d\varphi_k , \quad a_k(\varphi) := \big([\partial_\varphi \theta(\varphi)]^\top y(\varphi)\big)_k + \frac{1}{2}(\partial_x^{-1} w(\varphi), \partial_{\varphi_k} w(\varphi))_{L_x^2} , \tag{5.6}$$

is only "approximately closed", in the sense that

$$i_0^*\mathcal{W} = d\, i_0^*\Lambda = \sum\nolimits_{\substack{k,j \in \mathbb{S}_+ \\ k<j}} A_{kj}(\varphi) d\varphi_k \wedge d\varphi_j , \quad A_{kj}(\varphi) := \partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi) , \tag{5.7}$$

is of order $O(Z)$. More precisely, the following lemma holds.

Lemma 5.1. *Let $\omega \in \mathtt{DC}(\gamma, \tau)$ (cf. (4.4)). Then the coefficients A_{kj} in (5.7) satisfy*

$$\|A_{kj}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \gamma^{-1}\big(\|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\big) \tag{5.8}$$

for some $\sigma = \sigma(\tau, \mathbb{S}_+) > 0$.

Proof. The A_{kj} satisfy the identity $\omega \cdot \partial_\varphi A_{kj} = \mathcal{W}\big(\partial_\varphi Z(\varphi)\underline{e}_k, \partial_\varphi \breve{\iota}(\varphi)\underline{e}_j\big) + \mathcal{W}\big(\partial_\varphi \breve{\iota}_0(\varphi)\underline{e}_k, \partial_\varphi Z(\varphi)\underline{e}_j\big)$ where $\underline{e}_k$, $k \in \mathbb{S}_+$, denotes the standard basis of $\mathbb{R}^{\mathbb{S}_+}$ (cf. [7, Lemma 5]). Then (5.8) follows by (5.2) and (2.10). □

As in [7], [3] we first modify the approximate torus $\breve{\iota}$ to obtain an isotropic torus $\breve{\iota}_\delta$ which is still approximately invariant. Let $\Delta_\varphi := \sum_{k\in\mathbb{S}_+} \partial^2_{\varphi_k}$.

Lemma 5.2. (Isotropic torus) *Let $\omega \in \mathtt{DC}(\gamma, \tau)$. The torus $\breve{\iota}_\delta(\varphi) := (\theta(\varphi), y_\delta(\varphi), w(\varphi))$ defined by*

$$y_\delta(\varphi) := y(\varphi) - [\partial_\varphi \theta(\varphi)]^{-\top}\rho(\varphi)\,, \qquad \rho_j(\varphi) := \Delta_\varphi^{-1}\sum_{k\in\mathbb{S}_+} \partial_{\varphi_k} A_{kj}(\varphi)\,, \tag{5.9}$$

is isotropic and there is $\sigma = \sigma(\tau, \mathbb{S}_+) > 0$ so that, for any $s \geq s_0$

$$\|y_\delta - y\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \tag{5.10}$$

$$\|y_\delta - y\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \gamma^{-1}\big(\|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\big)\,, \tag{5.11}$$

$$\|\mathcal{F}_\omega(\iota_\delta, \zeta)\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} \tag{5.12}$$

$$\|d_\iota \iota_\delta[\widehat{\iota}]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\widehat{\iota}\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|\widehat{\iota}\|_{s_0}^{\mathrm{Lip}(\gamma)}\,. \tag{5.13}$$

Remark 5.3. *In the sequel, ω will always be assumed to be in $\mathtt{DC}(\gamma, \tau)$. Furthermore, $\sigma := \sigma(\tau, \mathbb{S}_+)$ will denote different, possibly larger "loss of derivatives" constants.*

Proof. The Lemma follows as in [3, Lemma 6.3] by Lemma 4.2, (5.6)-(5.8) and the ansatz (5.2). □

In order to find an approximate inverse of the linearized operator $d_{\iota,\zeta}\mathcal{F}_\omega(\iota_\delta)$, we introduce the symplectic diffeomorpshim $G_\delta : (\phi, \eta, v) \mapsto (\theta, y, w)$ of the phase space $\mathbb{T}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times L^2_\perp(\mathbb{T}_1)$, defined by

$$\begin{pmatrix} \theta \\ y \\ w \end{pmatrix} := G_\delta \begin{pmatrix} \phi \\ \eta \\ v \end{pmatrix} := \begin{pmatrix} \theta(\phi) \\ y_\delta(\phi) + [\partial_\phi \theta(\phi)]^{-T}\eta - \big[(\partial_\theta \tilde{w})(\theta(\phi))\big]^\top \partial_x^{-1} v \\ w(\phi) + v \end{pmatrix} \tag{5.14}$$

where $\tilde{w} := w \circ \theta^{-1}$. It is proved in [7, Lemma 2] that G_δ is symplectic, since by Lemma 5.2, $\breve{\iota}_\delta$ is an isotropic torus embedding. In the new coordinates, $\breve{\iota}_\delta$ is the trivial embedded torus $(\phi, \eta, v) = (\phi, 0, 0)$ and the Hamiltonian vector field $X_{\mathcal{H}_{\varepsilon,\zeta}}$ (with $\mathcal{H}_{\varepsilon,\zeta}$ defined in (4.7)) is given by

$$X_{\mathcal{K}} = (dG_\delta)^{-1} X_{\mathcal{H}_{\varepsilon,\zeta}} \circ G_\delta \qquad \text{where} \qquad \mathcal{K} \equiv \mathcal{K}_{\varepsilon,\zeta} := \mathcal{H}_{\varepsilon,\zeta} \circ G_\delta\,. \tag{5.15}$$

The Taylor expansion of $\mathcal{K}$ in η, v at the trivial torus $(\phi, 0, 0)$ is of the form

$$\begin{aligned} \mathcal{K}(\phi, \eta, v, \zeta) = \theta(\phi)\cdot\zeta + \mathcal{K}_{00}(\phi) + \mathcal{K}_{10}(\phi)\cdot\eta + (\mathcal{K}_{01}(\phi), v)_{L^2_x} + \frac{1}{2}\mathcal{K}_{20}(\phi)\eta\cdot\eta \\ + \big(\mathcal{K}_{11}(\phi)\eta, v\big)_{L^2_x} + \frac{1}{2}\big(\mathcal{K}_{02}(\phi)v, v\big)_{L^2_x} + \mathcal{K}_{\geq 3}(\phi, \eta, v) \end{aligned} \tag{5.16}$$

where $\mathcal{K}_{\geq 3}$ collects the terms which are at least cubic in the variables (η, v), $\mathcal{K}_{00}(\phi) \in \mathbb{R}$, $\mathcal{K}_{10}(\phi) \in \mathbb{R}^{\mathbb{S}_+}$, $\mathcal{K}_{01}(\phi) \in L^2_\perp(\mathbb{T}_1)$, $\mathcal{K}_{20}(\phi)$ is a $|\mathbb{S}_+| \times |\mathbb{S}_+|$ real matrix, $\mathcal{K}_{02}(\phi) : L^2_\perp(\mathbb{T}_1) \to L^2_\perp(\mathbb{T}_1)$ is a linear self-adjoint operator and $\mathcal{K}_{11}(\phi) : \mathbb{R}^{\mathbb{S}_+} \to L^2_\perp(\mathbb{T}_1)$ is a linear operator of finite rank. At an exact solution of $\mathcal{F}_\omega(\iota, \zeta) = 0$ one has $Z = 0$ and the coefficients in the Taylor expansion (5.16) satisfy $\mathcal{K}_{00} = \text{const}$, $\mathcal{K}_{10} = -\omega$, $\mathcal{K}_{01} = 0$.

Lemma 5.4. *There exists $\sigma := \sigma(\tau, \mathbb{S}_+)$ so that*

$$\begin{aligned} & \|\partial_\phi \mathcal{K}_{00}\|_s^{\mathrm{Lip}(\gamma)} + \|\mathcal{K}_{10} + \omega\|_s^{\mathrm{Lip}(\gamma)} + \|\mathcal{K}_{01}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} . \\ & \|\mathcal{K}_{20} - \Omega_{\mathbb{S}_+}^{kdv}(\nu)\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \varepsilon + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} , \\ & \|\mathcal{K}_{11}\eta\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \varepsilon\gamma^{-2} \|\eta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \|\eta\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} , \\ & \|\mathcal{K}_{11}^\top v\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \varepsilon\gamma^{-2} \|v\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \|v\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} . \end{aligned} \tag{5.17}$$

Proof. The lemma follows as in [7], [3], by applying Lemma 4.2 and (5.2), (5.10), (5.11), (5.12). □

Denote by $\mathrm{Id}_\perp$ the identity operator on $L^2_\perp(\mathbb{T}_1)$. The linear transformation $dG_\delta|_{(\varphi,0,0)} \equiv dG_\delta(\varphi, 0, 0)$ then reads

$$dG_\delta|_{(\varphi,0,0)} \begin{pmatrix} \widehat{\phi} \\ \widehat{\eta} \\ \widehat{v} \end{pmatrix} := \begin{pmatrix} \partial_\phi \theta(\varphi) & 0 & 0 \\ \partial_\phi y_\delta(\varphi) & [\partial_\phi \theta(\varphi)]^{-\top} & -[(\partial_\theta \tilde{w})(\theta(\varphi))]^\top \partial_x^{-1} \\ \partial_\phi w(\varphi) & 0 & \mathrm{Id}_\perp \end{pmatrix} \begin{pmatrix} \widehat{\phi} \\ \widehat{\eta} \\ \widehat{v} \end{pmatrix} . \tag{5.18}$$

It approximately transforms the linearized operator $d_{\iota,\zeta}\mathcal{F}_\omega(\iota_\delta)$ (see the proof of Theorem 5.6) into the one obtained when the Hamiltonian system with Hamiltonian $\mathcal{K}$ (cf. (5.15)) is linearized at $(\phi, \eta, v) = (\varphi, 0, 0)$, differentiated also with respect to ζ, and when ∂_t is exchanged by $\omega \cdot \partial_\varphi$,

$$\begin{pmatrix} \widehat{\phi} \\ \widehat{\eta} \\ \widehat{v} \\ \widehat{\zeta} \end{pmatrix} \mapsto \begin{pmatrix} \omega \cdot \partial_\varphi \widehat{\phi} + \partial_\phi \mathcal{K}_{10}(\varphi)[\widehat{\phi}] + \mathcal{K}_{20}(\varphi)\widehat{\eta} + \mathcal{K}_{11}^\top(\varphi)\widehat{v} \\ \omega \cdot \partial_\varphi \widehat{\eta} - \big(\partial_\phi \theta(\varphi)\big)^\top [\widehat{\zeta}] - \partial_\phi\big(\partial_\phi \theta(\varphi)^\top [\zeta]\big)[\widehat{\phi}] - \partial_{\phi\phi}\mathcal{K}_{00}(\varphi)[\widehat{\phi}] - [\partial_\phi \mathcal{K}_{10}(\varphi)]^\top \widehat{\eta} - [\partial_\phi \mathcal{K}_{01}(\varphi)]^\top \widehat{v} \\ \omega \cdot \partial_\varphi \widehat{v} - \partial_x\{\partial_\phi \mathcal{K}_{01}(\varphi)[\widehat{\phi}] + \mathcal{K}_{11}(\varphi)\widehat{\eta} + \mathcal{K}_{02}(\varphi)\widehat{v}\} \end{pmatrix} . \tag{5.19}$$

Using (5.2) and (5.10), one shows as in [3] that the induced operator $\widehat{\iota} := (\widehat{\phi}, \widehat{\eta}, \widehat{v}) \mapsto dG_\delta[\widehat{\iota}]$ satisfies

$$\|dG_\delta(\varphi, 0, 0)[\widehat{\iota}]\|_s^{\mathrm{Lip}(\gamma)} , \|dG_\delta(\varphi, 0, 0)^{-1}[\widehat{\iota}]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\widehat{\iota}\|_s^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}\|_{s_0}^{\mathrm{Lip}(\gamma)} , \tag{5.20}$$

$$\|d^2 G_\delta(\varphi, 0, 0)[\widehat{\iota}_1, \widehat{\iota}_2]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\widehat{\iota}_1\|_s^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}_2\|_{s_0}^{\mathrm{Lip}(\gamma)} + \|\widehat{\iota}_1\|_{s_0}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}_2\|_s^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}_1\|_{s_0}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}_2\|_{s_0}^{\mathrm{Lip}(\gamma)} . \tag{5.21}$$

In order to construct an "almost-approximate" inverse of (5.19) we need that

$$\mathcal{L}_\omega := \Pi_\perp \big(\omega \cdot \partial_\varphi - \partial_x \mathcal{K}_{02}(\varphi)\big)_{|L^2_\perp} \tag{5.22}$$

is "almost-invertible" up to remainders of size $O(N_{n-1}^{-\mathtt{a}})$ (see precisely (5.26)) where

$$N_n := K_n^p , \quad \forall n \geq 0 , \tag{5.23}$$

and

$$K_n := K_0^{\chi^n} , \quad \chi := 3/2 , \tag{5.24}$$

are the scales used in the nonlinear Nash-Moser iteration in Section 8. Based on results obtained in Sections 6-7, the almost invertibility of $\mathcal{L}_\omega$ is proved in Theorem 7.11, but here it is stated as an assumption to avoid the involved definition of the set Ω_o. Recall that $\mathtt{DC}(\gamma, \tau)$ is the set of diophantine frequencies in Ω (cf. (4.4)).

- **Almost-invertibility of $\mathcal{L}_\omega$.** *There exists a subset $\Omega_o \subset \mathtt{DC}(\gamma, \tau)$ such that, for all $\omega \in \Omega_o$, the operator $\mathcal{L}_\omega$ in (5.22) admits a decomposition*

$$\mathcal{L}_\omega = \mathcal{L}_\omega^< + \mathcal{R}_\omega + \mathcal{R}_\omega^\perp \tag{5.25}$$

with the following properties: there exist constants $K_0, N_0, \sigma, \tau_1, \mu(\mathtt{b}), \mathtt{a}, p, \mathfrak{s}_M > 0$ so that for any $\mathfrak{s}_M \leq s \leq S$ and $\omega \in \Omega_o$ one has:

(i) The operators $\mathcal{R}_\omega$, $\mathcal{R}_\omega^\perp$ satisfy the estimates

$$\|\mathcal{R}_\omega h\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \varepsilon\gamma^{-2} N_{n-1}^{-\mathtt{a}} \big(\|h\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1} \|\iota\|_{s+\mu(\mathtt{b})+\sigma}^{\mathrm{Lip}(\gamma)} \|h\|_{\mathfrak{s}_M+\sigma}^{\mathrm{Lip}(\gamma)} \big) \,, \tag{5.26}$$

$$\|\mathcal{R}_\omega^\perp h\|_{\mathfrak{s}_M}^{\mathrm{Lip}(\gamma)} \lesssim_{S,b} K_n^{-b} \big(\|h\|_{\mathfrak{s}_M+b+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1} \|\iota\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\sigma+b}^{\mathrm{Lip}(\gamma)} \|h\|_{\mathfrak{s}_M+\sigma}^{\mathrm{Lip}(\gamma)} \big) \,, \qquad \forall b > 0 \,, \tag{5.27}$$

(ii) For any $g \in H_\perp^{s+\sigma}(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1)$, there is a solution $h := (\mathcal{L}_\omega^<)^{-1} g \in H_\perp^s(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1)$ of the linear equation $\mathcal{L}_\omega^< h = g$, satisfying the tame estimates

$$\|(\mathcal{L}_\omega^<)^{-1} g\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \gamma^{-1} \big(\|g\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1} \|\iota\|_{s+\mu(\mathtt{b})+\sigma}^{\mathrm{Lip}(\gamma)} \|g\|_{\mathfrak{s}_M+\sigma}^{\mathrm{Lip}(\gamma)} \big) \,. \tag{5.28}$$

In order to find an almost-approximate inverse of the linear operator (5.19) and hence of $d_{\iota,\zeta}\mathcal{F}_\omega(\iota_\delta)$, it is sufficient to invert the operator

$$\mathbb{D}[\widehat{\phi}, \widehat{\eta}, \widehat{v}, \widehat{\zeta}] := \begin{pmatrix} \omega\cdot\partial_\varphi \widehat{\phi} + \mathcal{K}_{20}(\varphi)\widehat{\eta} + \mathcal{K}_{11}(\varphi)^\top \widehat{v} \\ \omega\cdot\partial_\varphi\widehat{\eta} - \partial_\phi\theta(\varphi)^\top \widehat{\zeta} \\ \mathcal{L}_\omega^< \widehat{v} - \partial_x \mathcal{K}_{11}(\varphi)\widehat{\eta} \end{pmatrix} \tag{5.29}$$

obtained by neglecting in (5.19) the terms $\partial_\phi\mathcal{K}_{10}$, $\partial_{\phi\phi}\mathcal{K}_{00}$, $\partial_\phi\mathcal{K}_{00}$, $\partial_\phi\mathcal{K}_{01}$, $\partial_\phi\big(\partial_\phi\theta(\varphi)^\top[\zeta]\big)$ and by replacing $\mathcal{L}_\omega$ by $\mathcal{L}_\omega^<$ (cf. (5.25)). Note that the remainder $\mathcal{L}_\omega - \mathcal{L}_\omega^< = \mathcal{R}_\omega + \mathcal{R}_\omega^\perp$ is small and that by Lemma 5.4 and (5.5), $\partial_\phi\mathcal{K}_{10}$, $\partial_{\phi\phi}\mathcal{K}_{00}$, $\partial_\phi\mathcal{K}_{00}$, $\partial_\phi\mathcal{K}_{01}$ and $\partial_\phi\big(\partial_\phi\theta(\varphi)^\top[\zeta]\big)$ are $O(Z)$.

We look for an inverse of $\mathbb{D}$ by solving the system

$$\mathbb{D}[\widehat{\phi}, \widehat{\eta}, \widehat{v}, \widehat{\zeta}] = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \,. \tag{5.30}$$

We first consider the second equation in (5.30), $\omega\cdot\partial_\varphi\widehat{\eta} = g_2 + \partial_\phi\theta(\varphi)^\top\widehat{\zeta}$. Since $\partial_\varphi\theta(\varphi) = \mathrm{Id} + \partial_\varphi\Theta(\varphi)$, the average $\langle \partial_\varphi\theta^\top\rangle_\varphi = \frac{1}{(2\pi)^{|\mathbb{S}_+|}}\int_{\mathbb{T}^{\mathbb{S}_+}} \partial_\varphi\theta^\top(\varphi) d\varphi$ equals the identity matrix Id of $\mathbb{R}^{\mathbb{S}_+}$. We then define

$$\widehat{\zeta} := -\langle g_2\rangle_\varphi \tag{5.31}$$

so that $\langle g_2 + \partial_\phi\theta(\varphi)^\top\widehat{\zeta}\rangle_\varphi$ vanishes and choose

$$\widehat{\eta} := \widehat{\eta}_0 + \widehat{\eta}_1, \quad \widehat{\eta}_1 := (\omega\cdot\partial_\varphi)^{-1}\big(g_2 + \partial_\phi\theta(\varphi)^\top\widehat{\zeta}\big) \tag{5.32}$$

where the constant vector $\widehat{\eta}_0 \in \mathbb{R}^{\mathbb{S}_+}$ will be determined in order to control the average of the first equation in (5.30). Next we consider the third equation in (5.30), $(\mathcal{L}_\omega^<)\widehat{v} = g_3 + \partial_x\mathcal{K}_{11}(\varphi)\widehat{\eta}$, which, by assumption (5.28) on the inveritibility of $\mathcal{L}_\omega^<$, has the solution

$$\widehat{v} := (\mathcal{L}_\omega^<)^{-1}\big(g_3 + \partial_x\mathcal{K}_{11}(\varphi)\widehat{\eta}_1\big) + (\mathcal{L}_\omega^<)^{-1}\partial_x\mathcal{K}_{11}(\varphi)\widehat{\eta}_0 \,. \tag{5.33}$$

Finally, we solve the first equation in (5.30). After substituting the solutions $\widehat{\zeta}$, $\widehat{\eta}$, defined in (5.32), and $\widehat{v}$, defined in (5.33), this equation becomes

$$\omega\cdot\partial_\varphi\widehat{\phi} = g_1 + M_1\widehat{\eta}_0 + M_2 g_2 + M_3 g_3 + M_4\widehat{\zeta} \tag{5.34}$$

where $M_j : \varphi \mapsto M_j(\varphi)$, $1 \le j \le 4$, are defined as

$$M_1(\varphi) := -\mathcal{K}_{20}(\varphi) - \mathcal{K}_{11}(\varphi)^\top(\mathcal{L}_\omega^<)^{-1}\partial_x\mathcal{K}_{11}(\varphi) \,, \tag{5.35}$$

$$M_2(\varphi) := M_1(\varphi)[\omega\cdot\partial_\varphi]^{-1} \,, \tag{5.36}$$

$$M_3(\varphi) := -\mathcal{K}_{11}(\varphi)^\top(\mathcal{L}_\omega^<)^{-1} \,, \tag{5.37}$$

$$M_4(\varphi) := M_2(\varphi)\partial_\phi\theta(\varphi)^\top \,. \tag{5.38}$$

In order to solve equation (5.34) we have to choose $\widehat{\eta}_0$ such that the right hand side of it has zero average. By Lemma 5.4, by the ansatz (5.2) and (5.28), the φ-averaged matrix is $\langle M_1\rangle_\varphi = \Omega^{kdv}_{\mathbb{S}_+}(\nu) + O(\varepsilon\gamma^{-2})$. Since the matrix $\Omega^{kdv}_{\mathbb{S}_+}(\nu) = (\partial_{I_k}\omega^{kdv}_n(\nu))_{k,n\in\mathbb{S}_+}$ is invertible (cf. Lemma 3.8-(i), Remark 3.9), $\langle M_1\rangle_\varphi$ is invertible for $\varepsilon\gamma^{-2}$ small enough and $\langle M_1\rangle_\varphi^{-1} = \Omega^{kdv}_{\mathbb{S}_+}(\nu)^{-1} + O(\varepsilon\gamma^{-2})$. We then define

$$\widehat{\eta}_0 := -\langle M_1\rangle_\varphi^{-1}\Big(\langle g_1\rangle_\varphi + \langle M_2 g_2\rangle_\varphi + \langle M_3 g_3\rangle_\varphi + \langle M_4\widehat{\zeta}\rangle_\varphi\Big)\,. \tag{5.39}$$

With this choice of $\widehat{\eta}_0$, the equation (5.34) has the solution

$$\widehat{\phi} := (\omega\cdot\partial_\varphi)^{-1}\big(g_1 + M_1[\widehat{\eta}_0] + M_2 g_2 + M_3 g_3 + M_4\widehat{\zeta}\big)\,. \tag{5.40}$$

Altogether we have obtained a solution $(\widehat{\phi},\widehat{\eta},\widehat{v},\widehat{\zeta})$ of the linear system (5.30).

Proposition 5.5. *Assume the ansatz* (5.2) *with $\mu_0 = \mu(\mathtt{b})+\sigma$ and the estimates* (5.28) *hold. Then, for any $\omega\in\Omega_o$ and any $g := (g_1,g_2,g_3)$ with $g_1,g_2\in H^{s+\sigma}(\mathbb{T}^{\mathbb{S}_+},\mathbb{R}^{\mathbb{S}_+})$, $g_3\in H^{s+\sigma}_\perp(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)$, and $\mathfrak{s}_M\le s\le S$, the system* (5.30) *has a solution $(\widehat{\phi},\widehat{\eta},\widehat{v},\widehat{\zeta}) := \mathbb{D}^{-1}g$, where $\widehat{\phi}$, $\widehat{\eta}$, $\widehat{v}$, $\widehat{\zeta}$ are defined in* (5.31)-(5.33), (5.39)-(5.40) *and satisfy*

$$\|\mathbb{D}^{-1}g\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \gamma^{-2}\big(\|g\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1}\|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\mu(\mathtt{b})+\sigma}\|g\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\sigma}\big). \tag{5.41}$$

Proof. The proposition follows by the definitions of $\widehat{\zeta}$ (cf. (5.31)), $\widehat{\eta}_1$ (cf. (5.32)), $\widehat{v}$ (cf. (5.33)), $\widehat{\eta}_0$ (cf. (5.39)), $\widehat{\phi}$ (cf. (5.40)), the definitions of M_j, $1\le j\le 4$, in (5.35)-(5.38), the estimates of Lemma 5.4, and the assumptions (5.2) and (5.28). ☐

Let $\widetilde{G}_\delta : (\phi,\eta,v,\zeta)\mapsto \big(G_\delta(\phi,\eta,v),\,\zeta\big)$ and notice that its differential $d\widetilde{G}_\delta(\phi,\eta,v,\zeta)$ is independent of ζ. In the sequel, we denote it by $d\widetilde{G}_\delta(\phi,\eta,v)$ or $d\widetilde{G}_\delta|_{(\phi,\eta,v)}$. Finally we prove that the operator

$$\mathbf{T}_0 := \mathbf{T}_0(\iota) := d\widetilde{G}_\delta|_{(\varphi,0,0)}\circ\mathbb{D}^{-1}\circ\big(dG_\delta|_{(\varphi,0,0)}\big)^{-1} \tag{5.42}$$

is an almost-approximate right inverse for $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$. Let $\|(\phi,\eta,v,\zeta)\|_s^{\mathrm{Lip}(\gamma)} := \max\{\|(\phi,\eta,v)\|_s^{\mathrm{Lip}(\gamma)}, |\zeta|^{\mathrm{Lip}(\gamma)}\}$.

Theorem 5.6. (Almost-approximate inverse) *Assume that* (5.25)-(5.28) *hold (Almost-invertibility of $\mathcal{L}_\omega$, $\omega\in\Omega_0$). Then there exists $\sigma_2 := \sigma_2(\tau,\mathbb{S}_+) > 0$ so that, if the ansatz* (5.2) *holds with $\mu_0\ge\mathfrak{s}_M+\mu(\mathtt{b})+\sigma_2$, then for any $\omega\in\Omega_o$ and any $g := (g_1,g_2,g_3)$ with $g_1,g_2\in H^{s+\sigma}(\mathbb{T}^{\mathbb{S}_+},\mathbb{R}^{\mathbb{S}_+})$, $g_3\in H^{s+\sigma}_\perp(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)$, and $\mathfrak{s}_M\le s\le S$, $\mathbf{T}_0(\iota)g$, defined by* (5.42), *satisfies*

$$\|\mathbf{T}_0(\iota)g\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \gamma^{-2}\Big(\|g\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_2} + N_0^{\tau_1}\gamma^{-1}\|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\mu(\mathtt{b})+\sigma_2}\|g\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\sigma_2}\Big)\,. \tag{5.43}$$

Moreover $\mathbf{T}_0(\iota)$ is an almost-approximate inverse of $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$, namely

$$d_{\iota,\zeta}\mathcal{F}_\omega(\iota)\circ\mathbf{T}_0(\iota) - \mathrm{Id} = \mathcal{P} + \mathcal{P}_\omega + \mathcal{P}_\omega^\perp \tag{5.44}$$

where

$$\|\mathcal{P}g\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M} \lesssim_S \gamma^{-3}\|\mathcal{F}_\omega(\iota,\zeta)\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\sigma_2}\Big(1 + N_0^{\tau_1}\gamma^{-1}\|\iota\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\mu(\mathtt{b})+\sigma_2}\Big)\|g\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\sigma_2} \tag{5.45}$$

$$\|\mathcal{P}_\omega g\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M} \lesssim_S \varepsilon\gamma^{-4}N_{n-1}^{-\mathtt{a}}\big(1 + N_0^{\tau_1}\gamma^{-1}\|\iota\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\mu(\mathtt{b})+\sigma_2}\big)\|g\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\sigma_2}\,, \tag{5.46}$$

$$\|\mathcal{P}^\perp_\omega g\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M} \lesssim_{S,b} \gamma^{-2}K_n^{-b}\big(\|g\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\sigma_2+b} + N_0^{\tau_1}\gamma^{-1}\|\iota\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\mu(\mathtt{b})+\sigma_2+b}\|g\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M+\sigma_2}\big)\,,\qquad \forall b>0\,. \tag{5.47}$$

Proof. The bound (5.43) follows from the definition of $\mathbf{T}_0(\iota)$ (cf.(5.42)), the estimate of $\mathbb{D}^{-1}$ (cf. (5.41)), and the estimates of $dG_\delta(\varphi,0,0)$ and of its inverse (cf. (5.20)). By formula (5.1)) for $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$ and since only the $y-$components of $\breve{\iota}_\delta$ and $\breve{\iota}$ differ from each other (cf. (5.9)), the difference $\mathcal{E}_0 := d_{\iota,\zeta}\mathcal{F}_\omega(\iota) - d_{\iota,\zeta}\mathcal{F}_\omega(\iota_\delta)$ can be written as

$$\mathcal{E}_0[\widehat{\iota},\widehat{\zeta}] = \int_0^1 \partial_y d_\iota X_{\mathcal{H}_\varepsilon}(\theta, y_\delta + s(y-y_\delta), w)[y-y_\delta,\widehat{\iota}]ds. \tag{5.48}$$

We introduce the projection $\Pi : (\widehat{\imath}, \widehat{\zeta}) \mapsto \widehat{\imath}$. Denote by $\mathtt{u} := (\phi, \eta, v)$ the symplectic coordinates defined by G_δ (cf. (5.14)). Under the symplectic map G_δ, the nonlinear operator $\mathcal{F}_\omega$ (cf. (4.6)) is transformed into

$$\mathcal{F}_\omega(G_\delta(\mathtt{u}(\varphi)), \zeta) = dG_\delta(\mathtt{u}(\varphi))[\omega \cdot \partial_\varphi \mathtt{u}(\varphi) - X_{\mathcal{K}}(\mathtt{u}(\varphi), \zeta)] \tag{5.49}$$

where $\mathcal{K} = \mathcal{H}_{\varepsilon,\zeta} \circ G_\delta$ (cf. (5.15)). Differentiating (5.49) at the trivial torus $\mathtt{u}_\delta(\varphi) = G_\delta^{-1}(\imath_\delta)(\varphi) = (\varphi, 0, 0)$, we get

$$\begin{aligned} d_{\imath,\zeta}\mathcal{F}_\omega(\imath_\delta) &= dG_\delta(\mathtt{u}_\delta)\big(\omega \cdot \partial_\varphi - d_{\mathtt{u},\zeta} X_{\mathcal{K}}(\mathtt{u}_\delta, \zeta)\big) d\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1} + \mathcal{E}_1 \,, & (5.50)\\ \mathcal{E}_1 &:= d^2 G_\delta(\mathtt{u}_\delta)\big[dG_\delta(\mathtt{u}_\delta)^{-1}\mathcal{F}_\omega(\imath_\delta, \zeta),\, dG_\delta(\mathtt{u}_\delta)^{-1}\Pi[\,\cdot\,]\,\big] \,. & (5.51)\end{aligned}$$

In expanded form $\omega \cdot \partial_\varphi - d_{\mathtt{u},\zeta} X_{\mathcal{K}}(\mathtt{u}_\delta, \zeta)$ is provided by (5.19). Recalling the definition of $\mathbb{D}$ in (5.29) and the discussion following it, we decompose $\omega \cdot \partial_\varphi - d_{\mathtt{u},\zeta} X_{\mathcal{K}}(\mathtt{u}_\delta, \zeta)$ as

$$\omega \cdot \partial_\varphi - d_{\mathtt{u},\zeta} X_{\mathcal{K}}(\mathtt{u}_\delta, \zeta) = \mathbb{D} + R_Z + \mathbb{R}_\omega + \mathbb{R}_\omega^\perp \tag{5.52}$$

where

$$R_Z[\widehat{\phi}, \widehat{\eta}, \widehat{v}, \widehat{\zeta}] := \begin{pmatrix} \partial_\phi \mathcal{K}_{10}(\varphi)[\widehat{\phi}] \\ -\partial_{\phi\phi}\mathcal{K}_{00}(\varphi)[\widehat{\phi}] - \partial_\phi\big(\partial_\phi \theta(\varphi)^\top[\zeta]\big)[\widehat{\phi}] - [\partial_\phi \mathcal{K}_{10}(\varphi)]^\top \widehat{\eta} - [\partial_\phi \mathcal{K}_{01}(\varphi)]^\top \widehat{v} \\ -\partial_x\big(\partial_\phi \mathcal{K}_{01}(\varphi)[\widehat{\phi}]\big) \end{pmatrix}$$

and

$$\mathbb{R}_\omega[\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\zeta}] := \begin{pmatrix} 0 \\ 0 \\ \mathcal{R}_\omega[\widehat{w}] \end{pmatrix}, \qquad \mathbb{R}_\omega^\perp[\widehat{\phi}, \widehat{y}, \widehat{w}, \widehat{\zeta}] := \begin{pmatrix} 0 \\ 0 \\ \mathcal{R}_\omega^\perp[\widehat{w}] \end{pmatrix}.$$

By (5.48) and (5.50)-(5.52) we get the decomposition

$$d_{\imath,\zeta}\mathcal{F}_\omega(\imath) = dG_\delta(\mathtt{u}_\delta) \circ \mathbb{D} \circ \big(d\widetilde{G}_\delta(\mathtt{u}_\delta)\big)^{-1} + \mathcal{E} + \mathcal{E}_\omega + \mathcal{E}_\omega^\perp \tag{5.53}$$

where

$$\mathcal{E} := \mathcal{E}_0 + \mathcal{E}_1 + dG_\delta(\mathtt{u}_\delta) R_Z \big(d\widetilde{G}_\delta(\mathtt{u}_\delta)\big)^{-1}, \tag{5.54}$$

$$\mathcal{E}_\omega := dG_\delta(\mathtt{u}_\delta)\mathbb{R}_\omega\big(d\widetilde{G}_\delta(\mathtt{u}_\delta)\big)^{-1}, \qquad \mathcal{E}_\omega^\perp := dG_\delta(\mathtt{u}_\delta)\mathbb{R}_\omega^\perp\big(d\widetilde{G}_\delta(\mathtt{u}_\delta)\big)^{-1}. \tag{5.55}$$

Letting the operator $\mathbf{T}_0 = \mathbf{T}_0(\imath)$ (cf. (5.42)) act from the right to both sides of the identity (5.53) and recalling that $\mathtt{u}_\delta(\varphi) = (\varphi, 0, 0)$, one obtains

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

To derive the claimed estimate for $\mathcal{P}$ we first need to estimate $\mathcal{E}$. By (5.2), (5.5) (estimate for ζ), (5.17) (estimates related to $\imath_\delta$), (5.10)–(5.12) (estimates of the components of R_Z), and (5.20)-(5.21) (estimates of $dG_\delta(u_\delta)$ and its inverse) one infers that

$$\|\mathcal{E}[\widehat{\imath}, \widehat{\zeta}]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \gamma^{-1}\Big(\|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|\widehat{\imath}\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|\widehat{\imath}\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|\imath\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|\widehat{\imath}\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\Big), \tag{5.56}$$

for some $\sigma > 0$, where Z is the error function, $Z = \mathcal{F}_\omega(\imath, \zeta)$ (cf. (5.3)). The claimed estimate (5.45) for $\mathcal{P}$ then follows from (5.56), the estimate (5.43) of $\mathbf{T}_0$, and the ansatz (5.2). The claimed estimates (5.46), (5.47) for $\mathcal{P}_\omega$ and, respectively, $\mathcal{P}_\omega^\perp$ follow by the assumed estimates (5.26)-(5.27) of $\mathcal{R}_\omega$ and $\mathcal{R}_\omega^\perp$, the estimate (5.43) of $\mathbf{T}_0$, the estimate (5.20) of $dG_\delta(u_\delta)$ and its inverse, and the ansatz (5.2). □

The goal of Sections 6 and 7 below is to prove that the Hamiltonian operator $\mathcal{L}_\omega$, defined in (5.22), satisfies the almost-invertibility property stated in (5.25)-(5.28).

6 Reduction of $\mathcal{L}_\omega$ up to order zero

The goal of this section is to reduce the Hamiltonian operator $\mathcal{L}_\omega$, defined in (5.22), to a differential operator of order three with constant coefficients, up to order zero – see (6.67) below for a precise statement. In the sequel, we consider torus embeddings $\breve{\iota}(\varphi) = (\varphi, 0, 0) + \iota(\varphi)$ with $\iota(\cdot) \equiv \iota(\cdot\,;\omega)$, $\omega \in \mathtt{DC}(\gamma, \tau)$ (cf. (4.4)), satisfying

$$\|\iota\|_{\mu_0}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}\ , \qquad \varepsilon\gamma^{-2} \le \delta(S) \tag{6.1}$$

where $\mu_0 := \mu_0(\tau, \mathbb{S}_+) > s_0$, $S > s_0$ are sufficiently large, $0 < \delta(S) < 1$ is sufficiently small, and $0 < \gamma < 1$. The Sobolev index S will be fixed in (8.2). In the course of the Nash-Moser iteration we will verify that (6.1) is satisfied by each approximate solution – see the bounds (8.8). For a quantity $g(\iota) \equiv g(\breve{\iota})$ such as an operator, a map, or a scalar function, depending on $\breve{\iota}(\varphi) = (\varphi, 0, 0) + \iota(\varphi)$, we denote for any two such tori embeddings $\breve{\iota}_1$, $\breve{\iota}_2$ by $\Delta_{12} g$ the difference

$$\Delta_{12} g := g(\iota_2) - g(\iota_1)\,.$$

6.1 Expansion of $\mathcal{L}_\omega$

As a first step, we derive an expansion of the operator $\mathcal{L}_\omega = \Pi_\perp\big(\omega \cdot \partial_\varphi - \partial_x \mathcal{K}_{02}(\varphi)\big)_{|L^2_\perp}$, defined in (5.22).

Lemma 6.1. *The Hamiltonian operator $\partial_x \mathcal{K}_{02}(\varphi)$ acting on $L^2_\perp(\mathbb{T}_1)$ is of the form*

$$\partial_x \mathcal{K}_{02}(\varphi) = \Pi_\perp \partial_x (d_\perp \nabla_w \mathcal{H}_\varepsilon)(\breve{\iota}_\delta(\varphi)) + R(\varphi) \tag{6.2}$$

where $\mathcal{H}_\varepsilon$ is the Hamiltonian defined in (4.2) *and the remainder $R(\varphi)$ is given by*

$$R(\varphi)[h] = \sum\nolimits_{j \in \mathbb{S}_+} \big(h\,,\, g_j\big)_{L^2_x} \chi_j\,, \quad \forall h \in L^2_\perp(\mathbb{T}_1)\,, \tag{6.3}$$

with functions $g_j, \chi_j \in H^s_\perp$, $j \in \mathbb{S}_+$, satisfying, for some $\sigma := \sigma(\tau, \mathbb{S}_+) > 0$ and any $s \ge s_0$

$$\|g_j\|_s^{\mathrm{Lip}(\gamma)} + \|\chi_j\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \varepsilon + \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\,. \tag{6.4}$$

Let $s_1 \ge s_0$ and let $\breve{\iota}_1, \breve{\iota}_2$ be two tori satisfying (6.1) *with $\mu_0 \ge s_1 + \sigma$. Then, for any $j \in \mathbb{S}_+$,*

$$\|\Delta_{12} g_j\|_{s_1} + \|\Delta_{12}\chi_j\|_{s_1} \lesssim_{s_1} \|\iota_2 - \iota_1\|_{s_1+\sigma}\,.$$

Proof. The lemma follows as in [9, Lemma 6.1], using Lemma 4.2 and the ansatz (6.1). □

By Lemma 6.1 the linear Hamiltonian operator $\mathcal{L}_\omega$ has the form

$$\mathcal{L}_\omega = \mathcal{L}_\omega^{(0)} - R\,, \qquad \mathcal{L}_\omega^{(0)} := \omega \cdot \partial_\varphi - \Pi_\perp \partial_x (d_\perp \nabla_w \mathcal{H}_\varepsilon)(\breve{\iota}_\delta(\varphi)) \tag{6.5}$$

where here and in the sequel, we write $\omega \cdot \partial_\varphi$ instead of $\Pi_\perp\ \omega \cdot \partial_\varphi|_{L^2_\perp}$ in order to simplify notation. We now prove that the Hamiltonian operator $\mathcal{L}_\omega^{(0)}$, acting on $L^2_\perp(\mathbb{T}_1)$, is a sum of a pseudo-differential operator of order three, a Fourier multiplier with $\varphi-$independent coefficients and a small smoothing remainder. Since $\mathcal{H}_\varepsilon = \mathcal{H}^{kdv} + \varepsilon\mathcal{P}$ (cf. (4.2)) and $\partial_x d_\perp \nabla_w \mathcal{H}^{kdv} = \partial_x \Omega^{kdv} + \partial_x d_\perp \nabla_w \mathcal{R}^{kdv}$ (cf. (3.6)) we have

$$\mathcal{L}_\omega^{(0)} = \omega \cdot \partial_\varphi + \partial_x^3 - \Pi_\perp Q_{-1}^{kdv}(D;\omega) - \Pi_\perp \partial_x d_\perp \nabla_w \mathcal{R}^{kdv}(\breve{\iota}_\delta) - \varepsilon \Pi_\perp \partial_x d_\perp \nabla_w \mathcal{P}(\breve{\iota}_\delta) \tag{6.6}$$

where we write ∂_x^3 instead of $\partial_x^3|_{L^2_\perp}$ and where $Q_{-1}^{kdv}(D;\omega)$ is given by (cf. (3.56))

$$Q_{-1}^{kdv}(D;\omega) \equiv Q_{-1}^{kdv}(D;\nu(\omega)) = \partial_x \Omega^{kdv}(D;\nu(\omega)) + \partial_x^3\,, \tag{6.7}$$

with $\nu(\omega)$ defined in (4.3). The operator $Q_{-1}^{kdv}(D;\omega)$ is a Fourier multiplier with $\varphi-$independent coefficients. It admits an expansion as described in the following lemma.

Lemma 6.2. *For any $M \in \mathbb{N}$,*

$$Q_{-1}^{kdv}(D;\omega) = \sum_{k=1}^{M} c_{-k}^{kdv}(\omega)\partial_x^{-k} + \mathcal{R}_M(Q_{-1}^{kdv};\omega) \tag{6.8}$$

where for any $1 \leq k \leq M$, the function $\Omega \to \mathbb{R}$, $\omega \mapsto c_{-k}^{kdv}(\omega)$ is Lipschitz and where $\mathcal{R}_M(Q_{-1}^{kdv};\omega) : L^2_\perp(\mathbb{T}_1) \to L^2_\perp(\mathbb{T}_1)$ is a Lipschitz family of diagonal operators of order $-M-1$. Furthermore, for any $n_1, n_2 \in \mathbb{N}$, $n_1 + n_2 \leq M+1$, the operator $\langle D\rangle^{n_1}\mathcal{R}_M(Q_{-1}^{kdv};\omega)\langle D\rangle^{n_2}$ is $\mathrm{Lip}(\gamma)$-tame with a tame constant satisfying $\mathfrak{M}_{\langle D\rangle^{n_1}\mathcal{R}_M(Q_{-1}^{kdv};\omega)\langle D\rangle^{n_2}}(s) \leq C(s,M)$ for any $s \geq s_0$ and $C(s,M) > 0$.

Proof. The claimed statements follow by Lemma 3.7. $\square$

Lemma 6.3. *For any $M \in \mathbb{N}$, the Hamiltonian operator $\mathcal{L}_\omega^{(0)}$, acting on $L^2_\perp(\mathbb{T}_1)$, defined in (6.5), admits an expansion of the form*

$$\mathcal{L}_\omega^{(0)} := \omega \cdot \partial_\varphi - \Pi_\perp\Big(a_3^{(0)}\partial_x^3 + 2(a_3^{(0)})_x\partial_x^2 + a_1^{(0)}\partial_x + \mathrm{Op}(r_0^{(0)}) + Q_{-1}^{kdv}(D;\omega)\Big) + \mathcal{R}_M^{(0)}(\breve{\iota}_\delta(\varphi);\omega) \tag{6.9}$$

where $a_3^{(0)} := a_3^{(0)}(\varphi,x;\omega)$, $a_1^{(0)} := a_1^{(0)}(\varphi,x;\omega)$ are real valued functions satisfying for any $s \geq s_0$

$$\|a_3^{(0)} + 1\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon(1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}), \quad \|a_1^{(0)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)} \tag{6.10}$$

for some $\sigma_M > 0$. The pseudo-differential symbol $r_0^{(0)} := r_0^{(0)}(\varphi,x,\xi;\omega)$ has an expansion in homogeneous components

$$r_0^{(0)}(\varphi,x,\xi;\omega) = \sum_{k=0}^{M} a_{-k}^{(0)}(\varphi,x;\omega)(\mathrm{i}2\pi\xi)^{-k}\chi_0(\xi) \tag{6.11}$$

(with χ_0 defined in (2.18)) where the coefficients $a_{-k}^{(0)} := a_{-k}^{(0)}(\varphi,x;\omega)$ satisfy

$$\sup_{k=0,\dots,M}\|a_{-k}^{(0)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}, \quad \forall s \geq s_0\,, \tag{6.12}$$

the remainder is defined by

$$\mathcal{R}_M^{(0)}(\breve{\iota}_\delta(\varphi);\omega) := -\mathcal{R}_M(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{R}^{kdv}) - \varepsilon\mathcal{R}_M(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{P}) \tag{6.13}$$

and the latter two remainder terms are given by (3.42) and (3.30) with $\nu(\omega) = (\omega^{kdv})^{-1}(-\omega)$.

Let $s_1 \geq s_0$ and let $\breve{\iota}_1, \breve{\iota}_2$ be two tori satisfying (6.1) for $\mu_0 \geq s_1 + \sigma_M$. Then, for any $0 \leq k \leq M+1$,

$$\|\Delta_{12}a_3^{(0)}\|_{s_1} \lesssim_{s_1,M} \varepsilon\|\iota_1 - \iota_2\|_{s_1+\sigma_M}, \quad \|\Delta_{12}a_{1-k}^{(0)}\|_{s_1} \lesssim_{s_1,M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}\,. \tag{6.14}$$

Proof. By the definition (6.6) of $\mathcal{L}_\omega^{(0)}$, the expansion (3.42) of $\partial_x d_\perp\nabla_w\mathcal{R}^{kdv}$, the expansion (3.30) of $\partial_x d_\perp\nabla_w\mathcal{P}$, and the formula for the coefficient of ∂_x^2, described in Lemma 2.6, one obtains (6.9) with

$$\begin{aligned}
&a_3^{(0)}(\varphi,x;\omega) := -1 + \varepsilon a_3(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{P})\,,\\
&a_1^{(0)}(\varphi,x;\omega) := a_1(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{R}^{kdv}) + \varepsilon a_1(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{P})\,,\\
&a_{-k}^{(0)}(\varphi,x;\omega) := a_{-k}(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{R}^{kdv}) + \varepsilon a_{-k}(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{P})\,, \quad k = 0,\dots,M\,,
\end{aligned}$$

and $\nu(\omega) = (\omega^{kdv})^{-1}(-\omega)$. By Lemma 3.6-1, the functions $a_{1-k}(\mathfrak{x};\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{R}^{kdv})$, $0 \leq k \leq M+1$, satisfy the hypothesis of Lemma 2.23-(*ii*). In view of (5.10) one then infers that for any $s \geq s_0$

$$\|a_{1-k}(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{R}^{kdv})\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}$$

for some $\sigma_M > 0$. Similarly, by the first item of Lemma 3.4, the functions $a_{3-k}(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{P})$, $0 \leq k \leq M+3$, satisfy the hypothesis of Lemma 2.23-(*i*), implying that for any $s \geq s_0$,

$$\|a_{3-k}(\breve{\iota}_\delta(\varphi);\nu(\omega);\partial_x d_\perp\nabla_w\mathcal{P})\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} 1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}$$

for some $\sigma_M > 0$, proving (6.10), (6.12). The estimates (6.14) follow by similar arguments. $\square$

We remark that in the finitely many steps of our reduction procedure, described in this section, the *loss of derivatives* $\sigma_M = \sigma_M(\tau, \mathbb{S}_+) > 0$ might have to be increased, but the notation will not be changed.

6.2 Quasi-periodic reparametrization of time

We conjugate the operator $\mathcal{L}_\omega$ (cf. (6.5)) by the change of variable induced by the quasi-periodic reparametrization of time

$$\vartheta = \varphi + \alpha^{(1)}(\varphi)\omega \qquad \text{or equivalently} \qquad \varphi = \vartheta + \breve{\alpha}^{(1)}(\vartheta)\omega$$

where $\alpha^{(1)} : \mathbb{T}^{\mathbb{S}_+} \to \mathbb{R}$, is a small, real valued function chosen below (cf. (6.17)). Denote by

$$(\Phi^{(1)}h)(\varphi, x) := h(\varphi + \alpha^{(1)}(\varphi)\omega, x)\,, \quad ((\Phi^{(1)})^{-1}h)(\vartheta, x) := h(\vartheta + \breve{\alpha}^{(1)}(\vartheta)\omega, x)\,, \tag{6.15}$$

the induced diffeomorphisms on functions. The goal is to achieve that the operator $\mathcal{L}_\omega^{(1)}$, defined in (6.20), is of the form (6.21), so that its highest order coefficient $a_3^{(1)}$ satisfies (6.23). The latter property will allow us in Section 6.3 to conjugate $\mathcal{L}_\omega^{(1)}$ to an operator with constant highest order coefficient (cf. (6.40)).

Since by (6.10), the coefficient $a_3^{(0)}$ satisfies $a_3^{(0)} = -1 + O(\varepsilon)$, the expression $(a_3^{(0)}(\varphi, x))^{\frac{1}{3}}$ is well defined where $(x)^{\frac{1}{3}}$ denotes the branch of the third root of $x \in (-\infty, 0)$, determined by $(-1)^{\frac{1}{3}} = -1$.

Lemma 6.4. *Let m_3 be the constant*

$$m_3(\omega) := \frac{1}{(2\pi)^{|\mathbb{S}_+|}} \int_{\mathbb{T}^{\mathbb{S}_+}} \Big(\int_{\mathbb{T}_1} \frac{dx}{(a_3^{(0)}(\vartheta, x; \omega))^{\frac{1}{3}}} \Big)^{-3} d\vartheta\,, \tag{6.16}$$

and define, for $\omega \in \mathtt{DC}(\gamma, \tau)$, the function

$$\breve{\alpha}^{(1)}(\vartheta; \omega) := (\omega \cdot \partial_\varphi)^{-1} \Big[\frac{1}{m_3} \Big(\int_{\mathbb{T}_1} \frac{dx}{(a_3^{(0)}(\vartheta, x; \omega))^{\frac{1}{3}}} \Big)^{-3} - 1 \Big]\,. \tag{6.17}$$

Then for any $M \in \mathbb{N}$, there exists a constant $\sigma_M > 0$ so that the following holds:
(i) The constant m_3 satisfies

$$|m_3 + 1|^{\mathrm{Lip}(\gamma)} \lesssim_M \varepsilon \tag{6.18}$$

and for any $s \geq s_0$, $\alpha^{(1)}, \breve{\alpha}^{(1)}$ satisfy

$$\|\alpha^{(1)}\|_s^{\mathrm{Lip}(\gamma)}, \|\breve{\alpha}^{(1)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon\gamma^{-1}(1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)})\,. \tag{6.19}$$

(ii) The Hamiltonian operator

$$\mathcal{L}_\omega^{(1)} := \frac{1}{\rho} \Phi^{(1)} \mathcal{L}_\omega \, (\Phi^{(1)})^{-1}\,, \quad \rho(\vartheta) := \Phi^{(1)}(1 + \omega \cdot \partial_\vartheta \breve{\alpha}^{(1)}) = 1 + \Phi^{(1)}(\omega \cdot \partial_\vartheta \breve{\alpha}^{(1)})\,, \tag{6.20}$$

admits an expansion of the form

$$\mathcal{L}_\omega^{(1)} = \omega \cdot \partial_\vartheta - \Big(a_3^{(1)} \partial_x^3 + 2(a_3^{(1)})_x \partial_x^2 + a_1^{(1)} \partial_x + \mathrm{Op}(r_0^{(1)}) + Q_{-1}^{kdv}(D; \omega) \Big) + \mathcal{R}_M^{(1)} \tag{6.21}$$

where the coefficients $a_3^{(1)} := a_3^{(1)}(\vartheta, x; \omega)$, $a_1^{(1)} := a_1^{(1)}(\vartheta, x; \omega)$ are real valued and satisfy

$$\|a_3^{(1)} + 1\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon(1 + \|\iota\|_{s+\sigma_M}), \quad \|a_1^{(1)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}\,, \quad \forall s \geq s_0\,, \tag{6.22}$$

and

$$\int_{\mathbb{T}_1} \frac{dx}{(a_3^{(1)}(\vartheta, x; \omega))^{\frac{1}{3}}} = m_3^{-\frac{1}{3}}\,, \quad \forall \vartheta \in \mathbb{T}^{\mathbb{S}_+}\,. \tag{6.23}$$

The function $r_0^{(1)} \equiv r_0^{(1)}(\vartheta, x, \xi; \omega)$ is a pseudo-differential symbol in S^0 and admits an expansion of the form

$$r_0^{(1)}(\vartheta, x, \xi; \omega) = \sum_{k=0}^{M} a_{-k}^{(1)}(\vartheta, x; \omega)(\mathrm{i}2\pi\xi)^{-k}\chi_0(\xi) \tag{6.24}$$

where for any $0 \le k \le M$, $s \ge s_0$,

$$\|a_{-k}^{(1)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)} . \tag{6.25}$$

Furthermore, the function ρ *appearing in* (6.20) *satisfies*

$$\|\rho - 1\|_s^{\mathrm{Lip}(\gamma)},\ \|\rho^{-1} - 1\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)} . \tag{6.26}$$

Let $s_1 \ge s_0$ *and let* ι_1, ι_2 *be two tori satisfying* (6.1) *with* $\mu_0 \ge s_1 + \sigma_M$. *Then*

$$\begin{aligned} &|\Delta_{12} m_3|, \|\Delta_{12}\alpha^{(1)}\|_{s_1}, \|\Delta_{12}\breve\alpha^{(1)}\|_{s_1}, \|\Delta_{12}a_1^{(1)}\|_{s_1}\,, \|\Delta_{12}\rho^{\pm 1}\|_{s_1} \lesssim_{s_1,M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}\,,\\ &\|\Delta_{12}a_{-k}^{(1)}\|_{s_1} \lesssim_{s_1,M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}\,, \quad \forall k = 0, \dots, M\,. \end{aligned} \tag{6.27}$$

(iii) Let $S > \mathfrak{s}_M$ *where* $\mathfrak{s}_M$ *is defined in* (2.54). *Then the maps* $(\Phi^{(1)})^{\pm 1}$ *are* $\mathrm{Lip}(\gamma)$*-1-tame operators with a tame constant satisfying*

$$\mathfrak{M}_{(\Phi^{(1)})^{\pm 1}}(s) \lesssim_{S,M} 1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}, \quad \forall s_0 + 1 \le s \le S\,. \tag{6.28}$$

For any given $\lambda_0 \in \mathbb{N}$ *there exists a constant* $\sigma_M(\lambda_0) > 0$ *so that for any* $m \in \mathbb{S}_+$, $\lambda, n_1, n_2 \in \mathbb{N}$ *with* $\lambda \le \lambda_0$ *and* $n_1 + n_2 + \lambda_0 \le M + 1$, *the operator* $\partial_{\varphi_m}^\lambda \langle D\rangle^{n_1} \mathcal{R}_M^{(1)} \langle D\rangle^{n_2}$ *is* $\mathrm{Lip}(\gamma)$*-tame with a tame constant satisfying*

$$\mathfrak{M}_{\partial_{\varphi_m}^\lambda \langle D\rangle^{n_1} \mathcal{R}_M^{(1)} \langle D\rangle^{n_2}}(s) \lesssim_{S,M} \varepsilon + \|\iota\|_{s+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)}\,, \quad \forall \mathfrak{s}_M \le s \le S\,. \tag{6.29}$$

If in addition $s_1 \ge \mathfrak{s}_M$ *and* $\breve\iota_1, \breve\iota_2$ *are two tori satisfying* (6.1) *with* $\mu_0 \ge s_1 + \sigma_M(\lambda_0)$, *then*

$$\|\partial_{\varphi_m}^\lambda \langle D\rangle^{n_1} \Delta_{12}\mathcal{R}_M^{(1)} \langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,M,\lambda_0} \|\iota_1 - \iota_2\|_{s_1+\sigma_M(\lambda_0)}\,. \tag{6.30}$$

Proof. Writing $\Pi_\perp$ as $\mathrm{Id} + (\Pi_\perp - \mathrm{Id})$ the expression (6.9) for $\mathcal{L}_\omega^{(0)}$ becomes

$$\mathcal{L}_\omega^{(0)} = \omega\cdot\partial_\varphi - \Big(a_3^{(0)}\partial_x^3 + 2(a_3^{(0)})_x\partial_x^2 + a_1^{(0)}\partial_x + \mathrm{Op}(r_0^{(0)}) + Q_{-1}^{kdv}(D;\omega)\Big) + \mathcal{R}_M^{(I)}(\breve\iota_\delta(\varphi);\omega) + \mathcal{R}_M^{(0)}(\breve\iota_\delta(\varphi);\omega)$$

where using that $(\mathrm{Id} - \Pi_\perp)\partial_x^3 h = 0$ for any $h \in H_\perp^s$, the operator $\mathcal{R}_M^{(I)} \equiv \mathcal{R}_M^{(I)}(\breve\iota_\delta(\varphi);\omega)$ can be written as

$$\mathcal{R}_M^{(I)} = (\mathrm{Id} - \Pi_\perp)\Big((a_3^{(0)} + 1)\partial_x^3 + 2(a_3^{(0)})_x\partial_x^2 + a_1^{(0)}\partial_x + \mathrm{Op}(r_0^{(0)})\Big)\,. \tag{6.31}$$

Since $(\mathrm{Id} - \Pi_\perp)h = \sum_{j\in\mathbb{S}} \big(h, e^{-\mathrm{i}2\pi jx}\big)_{L_x^2} e^{\mathrm{i}2\pi jx}$ for any h in L_x^2, $\mathcal{R}_M^{(I)}$ is a finite rank operator of the form (6.3) with functions $g_j, \chi_j \in H_\perp^s$ satisfying (6.4) (use (6.10), (6.12)).

The estimate (6.28) follows by Lemma 2.1-(*iii*) and (6.19). Notice that

$$\Phi^{(1)} \circ \omega\cdot\partial_\varphi \circ (\Phi^{(1)})^{-1} = \rho(\vartheta)\omega\cdot\partial_\vartheta\,, \quad \rho := \Phi^{(1)}(1 + \omega\cdot\partial_\varphi\breve\alpha^{(1)})\,,$$

and that any Fourier multiplier $g(D)$ is left unchanged under conjugation, i.e. $\Phi^{(1)}g(D)(\Phi^{(1)})^{-1} = g(D)$. Using (6.5) and (6.9), we obtain (6.21) where

$$a_3^{(1)} := \Phi^{(1)}\Big(\frac{a_3^{(0)}}{1 + \omega\cdot\partial_\varphi\breve\alpha^{(1)}}\Big)\,, \tag{6.32}$$

$a_1^{(1)} := \frac{1}{\rho}\Phi^{(1)}(a_1^{(0)})$, $r_0^{(1)}$ is of the form (6.24) with $a_{-k}^{(1)} := \frac{1}{\rho}\Phi^{(1)}(a_{-k}^{(0)})$, and the remainder $\mathcal{R}_M^{(1)}$ is given by

$$\mathcal{R}_M^{(1)} = \frac{1}{\rho}\Phi^{(1)}\mathcal{R}_M^{(I)}(\Phi^{(1)})^{-1} + \frac{1}{\rho}\Phi^{(1)}\mathcal{R}_M^{(0)}(\breve\iota_\delta(\varphi))(\Phi^{(1)})^{-1} - \frac{1}{\rho}\Phi^{(1)}R(\varphi)(\Phi^{(1)})^{-1}\,. \tag{6.33}$$

We choose $\breve\alpha^{(1)}$ such that (6.23) holds, obtaining (6.16), (6.17). We now verify the estimates, stated in items (*i*), (*ii*). Recall that we assume throughout that (6.1) holds. The estimates (6.18)-(6.19) follow by (6.16), (6.17), (6.10), and by using Lemma 2.1-(*iii*), Lemma 2.2. The estimate (6.26) on ρ follows by the definition (6.20), (6.17), and by applying Lemma 2.1-(*iii*), Lemma 2.2. Hence, by Lemma 2.1 and the estimates (6.10), (6.12), and (6.26), we deduce (6.25). The estimates (6.27) are obtained by similar arguments. Let us now prove item (*iii*). The estimate (6.28) follows from Lemma 2.1-(*iii*). Since $(\Phi^{(1)})^{\pm 1}$ commutes with every Fourier multiplier, we get

$$\frac{1}{\rho}\langle D\rangle^{n_1}\Phi^{(1)}\mathcal{R}_M^{(0)}(\breve{\iota}_\delta(\varphi))(\Phi^{(1)})^{-1}\langle D\rangle^{n_2} = \frac{1}{\rho}\langle D\rangle^{n_1}\mathcal{R}_M^{(0)}(\breve{\iota}_{\delta,\alpha}(\varphi))\langle D\rangle^{n_2} \tag{6.34}$$

where $\breve{\iota}_{\delta,\alpha}(\varphi) := \breve{\iota}_\delta(\varphi + \alpha^{(1)}(\varphi)\omega)$. By Lemma 2.1, (5.10), and (6.19) one has $\|\iota_{\delta,\alpha}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}$. Moreover, by (6.3), we have

$$\frac{1}{\rho}\Phi^{(1)}R(\varphi)(\Phi^{(1)})^{-1}h = \sum\nolimits_{j\in\mathbb{S}_+}\big(h\,,\,(\Phi^{(1)}g_j)\big)_{L^2_x}\frac{1}{\rho}(\Phi^{(1)}\chi_j)\,, \quad \forall h \in L^2_\perp\,, \tag{6.35}$$

and by (6.31), the conjugated operator $\frac{1}{\rho}\Phi^{(1)}\mathcal{R}_M^{(I)}(\Phi^{(1)})^{-1}h$ has the same form. The estimates (6.29) then follow by (6.34), (6.13), and Lemmata 3.4, 3.6, 2.24 to estimate the first term on the right hand side of (6.33) and by (6.35), (6.28), (6.4) and Lemma 2.22, to estimate the second and third term in (6.33). The estimates (6.30) are proved by similar arguments. $\square$

6.3 Elimination of the (φ, x)-dependence of the highest order coefficient

The goal of this section is to remove the (φ, x)-dependence of the coefficient $a_3^{(1)}(\varphi, x)$ of the Hamiltonian operator $\mathcal{L}_\omega^{(1)}$, given by (6.20)-(6.21), where we rename ϑ with φ. Actually this step will at the same time also remove the coefficient of ∂_x^2. We achieve these goals by conjugating the operator $\mathcal{L}_\omega^{(1)}$ by the flow $\Phi^{(2)}(\tau, \varphi)$, acting on $L^2_\perp(\mathbb{T}_1)$, defined by the transport equation

$$\partial_\tau\Phi^{(2)}(\tau,\varphi) = \Pi_\perp\partial_x\big(b^{(2)}(\tau,\varphi,x)\Phi^{(2)}(\tau,\varphi)\big)\,, \quad \Phi^{(2)}(0,\varphi) = \mathrm{Id}_\perp\,, \tag{6.36}$$

for a real valued function

$$b^{(2)} \equiv b^{(2)}(\tau,\varphi,x) := \frac{\beta^{(2)}(\varphi,x)}{1+\tau\beta_x^{(2)}(\varphi,x)}\,,$$

where $\beta^{(2)}(\varphi, x)$ is a small, real valued periodic function chosen in (6.38) below. The flow $\Phi^{(2)}(\tau, \varphi)$ is well defined for $0 \le \tau \le 1$ and satisfies the tame estimates provided in Lemma 2.25. Since the vector field $\Pi_\perp\partial_x\big(b^{(2)}h\big)$, $h \in H^s_\perp(\mathbb{T}_1)$, is Hamiltonian (it is generated by the Hamiltonian $\frac{1}{2}\int_{\mathbb{T}_1} b^{(2)}h^2\,dx$), each $\Phi^{(2)}(\tau, \varphi)$, $0 \le \tau \le 1$, $\varphi \in \mathbb{T}^{\mathbb{S}_+}$ is a symplectic linear isomorphism of $H^s_\perp(\mathbb{T}_1)$. Therefore the time one conjugated operator

$$\mathcal{L}_\omega^{(2)} := \Phi^{(2)}\mathcal{L}_\omega^{(1)}\big(\Phi^{(2)}\big)^{-1}\,, \quad \Phi^{(2)} := \Phi^{(2)}(1,\varphi)\,, \tag{6.37}$$

is a Hamiltonian operator acting on $H^s_\perp(\mathbb{T}_1)$.

Given the (τ, φ)-dependent family of diffeomorphisms of the torus $\mathbb{T}_1$, $x \mapsto y = x+\tau\beta^{(2)}(\varphi, x)$, we denote the family of its inverses by $y \mapsto x = y + \breve{\beta}^{(2)}(\tau, \varphi, y)$.

Lemma 6.5. *Let $\breve{\beta}^{(2)}(\varphi, y; \omega) \equiv \breve{\beta}^{(2)}(1, \varphi, y; \omega)$ be the real valued, periodic function*

$$\breve{\beta}^{(2)}(\varphi,y;\omega) := \partial_y^{-1}\Big(\frac{m_3^{1/3}}{(a_3^{(1)}(\varphi,y;\omega))^{1/3}} - 1\Big) \tag{6.38}$$

(which is well defined by (6.23)) and let $M \in \mathbb{N}$. Then there exists $\sigma_M > 0$ so that the following holds:
(i) For any $s \ge s_0$

$$\|\beta^{(2)}\|_s^{\mathrm{Lip}(\gamma)}, \|\breve{\beta}^{(2)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon\big(1+\|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}\big)\,. \tag{6.39}$$

(ii) The Hamiltonian operator $\mathcal{L}_\omega^{(2)}$ in (6.37) admits an expansion of the form

$$\mathcal{L}_\omega^{(2)} = \omega\cdot\partial_\varphi - \big(m_3\partial_x^3 + a_1^{(2)}\partial_x + \mathrm{Op}(r_0^{(2)}) + Q_{-1}^{kdv}(D;\omega)\big) + \mathcal{R}_M^{(2)} \tag{6.40}$$

where $a_1^{(2)} := a_1^{(2)}(\varphi, x; \omega)$ *is a real valued, periodic function, satisfying*

$$\|a_1^{(2)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)} . \tag{6.41}$$

The pseudo-differential symbol $r_0^{(2)} \equiv r_0^{(2)}(\varphi, x, \xi; \omega)$ *is in* S^0 *and satisfies, for any* $s \geq s_0$, *the estimate*

$$|\mathrm{Op}(r_0^{(2)})|_{0,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)} . \tag{6.42}$$

Let $s_1 \geq s_0$ *and let* $\breve{\iota}_1, \breve{\iota}_2$ *be two tori satisfying* (6.1) *for* $\mu_0 \geq s_1 + \sigma_M$. *Then, for any* $k = 0, \ldots, M$,

$$\|\Delta_{12}\beta^{(2)}\|_{s_1}, \|\Delta_{12}\breve{\beta}^{(2)}\|_{s_1}, \|\Delta_{12}a_1^{(2)}\|_{s_1}, |\Delta_{12}\mathrm{Op}(r_0^{(2)})|_{0,s_1,0} \lesssim_{s_1,M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M} . \tag{6.43}$$

(iii) Let $S > \mathfrak{s}_M$. *Then the symplectic maps* $(\Phi^{(2)})^{\pm 1}$ *are* $\mathrm{Lip}(\gamma)$*-1 tame operators with a tame constant satisfying*

$$\mathfrak{M}_{(\Phi^{(2)})^{\pm 1}}(s) \lesssim_{S,M} 1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}, \quad \forall s_0 + 1 \leq s \leq S . \tag{6.44}$$

Let $\lambda_0 \in \mathbb{N}$. *Then there exists a constant* $\sigma_M(\lambda_0) > 0$ *such that, for any* $\lambda, n_1, n_2 \in \mathbb{N}$ *with* $\lambda \leq \lambda_0$ *and* $n_1 + n_2 + \lambda_0 \leq M - 1$, *the operator* $\partial_{\varphi_m}^\lambda \langle D\rangle^{n_1} \mathcal{R}_M^{(2)} \langle D\rangle^{n_2}$, $m \in \mathbb{S}_+$, *is* $\mathrm{Lip}(\gamma)$*-tame with a tame constant satisfying*

$$\mathfrak{M}_{\partial_{\varphi_m}^\lambda \langle D\rangle^{n_1} \mathcal{R}_M^{(2)} \langle D\rangle^{n_2}}(s) \lesssim_{S,M,\lambda_0} \varepsilon + \|\iota\|_{s+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)}, \quad \forall \mathfrak{s}_M \leq s \leq S . \tag{6.45}$$

Let $s_1 \geq \mathfrak{s}_M$ *and* ι_1, ι_2 *be tori satisfying* (6.1) *with* $\mu_0 \geq s_1 + \sigma_M(\lambda_0)$. *Then*

$$\|\partial_{\varphi_m}^\lambda \langle D\rangle^{n_1} \Delta_{12}\mathcal{R}_M^{(2)} \langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,M,\lambda_0} \|\iota_1 - \iota_2\|_{s_1+\sigma_M(\lambda_0)} . \tag{6.46}$$

Proof. The proof of this lemma uses the Egorov type results proved in Section 2.5. According to (6.21), (6.24), the conjugated operator is given by

$$\begin{aligned}
&\mathcal{L}_\omega^{(2)} = \Phi^{(2)}\mathcal{L}_\omega^{(1)}(\Phi^{(2)})^{-1} \qquad (6.47)\\
&= \omega\cdot\partial_\varphi - \Phi^{(2)}a_3^{(1)}\partial_x^3(\Phi^{(2)})^{-1} - 2\Phi^{(2)}(a_3^{(1)})_x\partial_x^2(\Phi^{(2)})^{-1} - \Phi^{(2)}a_1^{(1)}\partial_x(\Phi^{(2)})^{-1}\\
&\quad - \sum_{k=0}^{M}\Phi^{(2)}a_{-k}^{(1)}\partial_x^{-k}(\Phi^{(2)})^{-1} - \Phi^{(2)}Q_{-1}^{kdv}(D;\omega)(\Phi^{(2)})^{-1} + \Phi^{(2)}\mathcal{R}_M^{(1)}(\Phi^{(2)})^{-1} + \Phi^{(2)}\big(\omega\cdot\partial_\varphi\,(\Phi^{(2)})^{-1}\big) .
\end{aligned}$$

By (6.38), (6.18), (6.22) and Lemmata 2.1, 2.2, the estimate (6.39) follows. Using the ansatz (6.1) with $\mu_0 > 0$ large enough, the estimate (6.39) implies that $\|\beta^{(2)}\|_{s_0+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)} \lesssim_{M,\lambda_0} \varepsilon\gamma^{-2}$, where the constant $\sigma_M(\lambda_0)$ is the constant appearing in the smallness conditions (2.78), (2.104), (2.106). Now we apply Proposition 2.28 to expand the terms

$$\Phi^{(2)}a_3^{(1)}\partial_x^3(\Phi^{(2)})^{-1}, \quad 2\Phi^{(2)}(a_3^{(1)})_x\partial_x^2(\Phi^{(2)})^{-1}, \quad \Phi^{(2)}a_{1-k}^{(1)}\partial_x^{1-k}(\Phi^{(2)})^{-1}, \ 0 \leq k \leq M+1 ,$$

Lemma 2.32 to expand the term $\Phi^{(2)}Q_{-1}^{kdv}(D;\omega)(\Phi^{(2)})^{-1}$, and Proposition 2.31 to expand $\Phi^{(2)}\big(\omega\cdot\partial_\varphi\,(\Phi^{(2)})^{-1}\big)$. Using also the estimates (6.10), (6.12), (6.39) one deduces (6.41), (6.42). By the choice of $\breve{\beta}^{(2)}$ in (6.38) and Proposition 2.28, the coefficient of the highest order term of $\Phi^{(2)}a_3^{(1)}\partial_x^3(\Phi^{(2)})^{-1}$ (and of $\mathcal{L}_\omega^{(2)}$) is given by

$$\big([1 + \breve{\beta}_y^{(2)}(\varphi, y)]^3 a_3^{(1)}(\varphi, y)\big)|_{y=x+\beta^{(2)}(\varphi,x)} = m_3$$

which is constant in (φ, x) by (6.23). Since $\Phi^{(2)}$ is symplectic, the operator $\mathcal{L}_\omega^{(2)}$ is Hamiltonian and hence by Lemma 2.6 the second order term equals $2(m_3)_x\partial_x^2$ which vanishes since m_3 is constant. The remainder $\Phi^{(2)}\mathcal{R}_M^{(1)}(\Phi^{(2)})^{-1}$ can be estimated by arguing as at the end of the proof of Proposition 2.28 (estimate of $\mathcal{R}_N(\tau,\varphi)$), using Lemma 2.25 to estimate $\Phi^{(2)}$, $(\Phi^{(2)})^{-1}$, the estimate (6.29) for $\mathcal{R}_M^{(1)}$, the estimate (6.39) of $\beta^{(2)}$, $\breve{\beta}^{(2)}$, and the ansatz (6.1) with μ_0 large enough. The estimates (6.44) follow by (2.72) and (6.39). The estimates (6.43), (6.46) are derived by similar arguments. $\square$

6.4 Elimination of the x-dependence of the first order coefficient

The goal of this section is to remove the x-dependence of the coefficient $a_1^{(2)}(\varphi,x)$ of the Hamiltonian operator $\mathcal{L}_\omega^{(2)}$ in (6.37), (6.40). We conjugate the operator $\mathcal{L}_\omega^{(2)}$ by the change of variable induced by the flow $\Phi^{(3)}(\tau,\varphi)$, acting on $L^2_\perp(\mathbb{T}_1)$, defined by

$$\partial_\tau \Phi^{(3)}(\tau,\varphi) = \Pi_\perp\big(b^{(3)}(\varphi,x)\partial_x^{-1}\Phi^{(3)}(\tau,\varphi)\big)\,, \quad \Phi^{(3)}(0) = \mathrm{Id}_\perp\,, \tag{6.48}$$

where $b^{(3)}(\varphi,x)$ is a small, real valued, periodic function chosen in (6.50) below. Since the vector field $\Pi_\perp\big(b^{(3)}\partial_x^{-1}h\big)$, $h \in H^s_\perp(\mathbb{T}_1)$, is Hamiltonian (it is generated by the Hamiltonian $\frac{1}{2}\int_{\mathbb{T}_1} b^{(3)}(\partial_x^{-1}h)^2\,dx$), each $\Phi^{(3)}(\tau,\varphi)$ is a symplectic linear isomorphism of $H^s_\perp$ for any $0 \le \tau \le 1$ and $\varphi \in \mathbb{T}^{\mathbb{S}_+}$, and the time one conjugated operator

$$\mathcal{L}_\omega^{(3)} := \Phi^{(3)}\mathcal{L}_\omega^{(2)}\big(\Phi^{(3)}\big)^{-1}\,, \quad \Phi^{(3)} := \Phi^{(3)}(1)\,, \tag{6.49}$$

is Hamiltonian.

Lemma 6.6. *Let $b^{(3)}(\varphi,x;\omega)$ be the real valued periodic function*

$$b^{(3)}(\varphi,x;\omega) := \frac{1}{3m_3}\partial_x^{-1}\Big(a_1^{(2)}(\varphi,x;\omega) - \langle a_1^{(2)}\rangle_x(\varphi;\omega)\Big)\,, \quad \langle a_1^{(2)}\rangle_x(\varphi;\omega) := \int_{\mathbb{T}_1} a_1^{(2)}(\varphi,x;\omega)\,dx \tag{6.50}$$

and let $M \in \mathbb{N}$. Then there exists $\sigma_M > 0$ with the following properties:
(i) For any $s \ge s_0$,

$$\|b^{(3)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)} \tag{6.51}$$

and the symplectic maps $(\Phi^{(3)})^{\pm 1}$ are $\mathrm{Lip}(\gamma)$-tame and satisfy

$$\mathfrak{M}_{(\Phi^{(3)})^{\pm 1}}(s) \lesssim_{s,M} 1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}\,. \tag{6.52}$$

(ii) The Hamiltonian operator in (6.49) admits an expansion of the form

$$\mathcal{L}_\omega^{(3)} = \omega\cdot\partial_\varphi - \big(m_3\partial_x^3 + a_1^{(3)}(\varphi)\partial_x + \mathrm{Op}(r_0^{(3)}) + Q_{-1}^{kdv}(D;\omega)\big) + \mathcal{R}_M^{(3)} \tag{6.53}$$

where the real valued, periodic function $a_1^{(3)}(\varphi;\omega) := \langle a_1^{(2)}\rangle_x(\varphi;\omega)$ satisfies

$$\|a_1^{(3)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}\,, \tag{6.54}$$

and $r_0^{(3)} := r_0^{(3)}(\varphi,x,\xi;\omega)$ is a pseudo-differential symbol in S^0 satisfying for any $s \ge s_0$,

$$|\mathrm{Op}(r_0^{(3)})|_{0,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}\,. \tag{6.55}$$

Let $s_1 \ge s_0$ and let $\breve{\iota}_1, \breve{\iota}_2$ be two tori satisfying (6.1) with $\mu_0 \ge s_1 + \sigma_M$. Then

$$\|\Delta_{12}b^{(3)}\|_{s_1}\,, \|\Delta_{12}a_1^{(3)}\|_{s_1} \lesssim_{s_1,M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}\,, \quad |\Delta_{12}\mathrm{Op}(r_0^{(3)})|_{0,s_1,0} \lesssim_{s_1,M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}\,. \tag{6.56}$$

(iii) Let $S > \mathfrak{s}_M$, $\lambda_0 \in \mathbb{N}$. Then there exists a constant $\sigma_M(\lambda_0) > 0$ so that for any $m \in \mathbb{S}_+$ and $\lambda, n_1, n_2 \in \mathbb{N}$ with $\lambda \le \lambda_0$ and $n_1 + n_2 + \lambda_0 \le M - 1$, the operator $\langle D\rangle^{n_1}\partial_{\varphi_m}^\lambda \mathcal{R}_M{}^{(3)}\langle D\rangle^{n_2}$, is $\mathrm{Lip}(\gamma)$-tame with tame constants satisfying

$$\mathfrak{M}_{\partial_{\varphi_m}^\lambda \langle D\rangle^{n_1}\mathcal{R}_M^{(3)}\langle D\rangle^{n_2}}(s) \lesssim_{S,M,\lambda_0} \varepsilon + \|\iota\|_{s+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)}\,, \quad \forall \mathfrak{s}_M \le s \le S\,. \tag{6.57}$$

Let $s_1 \ge \mathfrak{s}_M$ and let $\breve{\iota}_1, \breve{\iota}_2$ be tori satisfying (6.1) with $\mu_0 \ge s_1 + \sigma_M(\lambda_0)$. Then

$$\|\partial_{\varphi_m}^\lambda \langle D\rangle^{n_1}\Delta_{12}\mathcal{R}_M^{(3)}\langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,M,\lambda_0} \|\iota_1 - \iota_2\|_{s_1+\sigma_M(\lambda_0)}\,. \tag{6.58}$$

Proof. The estimate (6.51) follows by the definition (6.50) and (6.41), (6.18). We now provide estimates for the flow

$$\Phi^{(3)}(\tau) = \exp\big(\tau \Pi_\perp b^{(3)}(\varphi, x; \omega)\partial_x^{-1}\big), \quad \forall \tau \in [-1, 1].$$

By (2.20), Lemma 2.9, and (6.51), one infers that for any $s \geq s_0$, $|\Pi_\perp b^{(3)} \partial_x^{-1}|_{-1,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}$. Therefore, by Lemma 2.12, there exists $\sigma_M > 0$ such that, if (6.1) holds with $\mu_0 \geq \sigma_M$, then, for any $s \geq s_0$,

$$\sup_{\tau \in [-1,1]} |\Phi^{(3)}(\tau) - \mathrm{Id}|_{0,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_s \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}. \tag{6.59}$$

The latter estimate, together with Lemma 2.16, imply (6.52).

By (6.40) and using Lemma 6.2 for the operator $Q_{-1}^{kdv}(D; \omega)$, one has that

$$\Phi^{(3)} \mathcal{L}_\omega^{(2)} (\Phi^{(3)})^{-1} = \omega \cdot \partial_\varphi - \Phi^{(3)}\big(m_3 \partial_x^3 + a_1^{(2)} \partial_x\big)(\Phi^{(3)})^{-1} - Q_{-1}^{kdv}(D;\ \omega) + \mathcal{R}_0^{(I)} + \mathcal{R}_M^{(3)}$$

where

$$\begin{aligned} \mathcal{R}_0^{(I)} &:= -\Phi^{(3)} \mathrm{Op}(r_0^{(2)})(\Phi^{(3)})^{-1} + \Phi^{(3)}\big(\omega \cdot \partial_\varphi (\Phi^{(3)})^{-1}\big) - (\Phi^{(3)} - \mathrm{Id}_\perp)\Pi_\perp \Big(\sum_{k=1}^M c_{-k}^{kdv}(\omega)\partial_x^{-k}\Big)(\Phi^{(3)})^{-1} \\ &\quad - \Pi_\perp \Big(\sum_{k=1}^M c_{-k}^{kdv}(\omega)\partial_x^{-k}\Big)\Big((\Phi^{(3)})^{-1} - \mathrm{Id}_\perp\Big), \\ \mathcal{R}_M^{(3)} &:= \Phi^{(3)} \mathcal{R}_M^{(2)} (\Phi^{(3)})^{-1} - (\Phi^{(3)} - \mathrm{Id}_\perp)\mathcal{R}_M(\omega, Q_{-1}^{kdv})(\Phi^{(3)})^{-1} - \mathcal{R}_M(\omega, Q_{-1}^{kdv})\big((\Phi^{(3)})^{-1} - \mathrm{Id}_\perp\big). \end{aligned} \tag{6.60}$$

Note that $\mathcal{R}_0^{(I)}$ is a pseudo-differential operator in OPS^0 (cf. Lemma 2.12). Moreover, by a Lie expansion, recalling (6.48), one has

$$\begin{aligned} \Phi^{(3)}\big(m_3 \partial_x^3 + a_1^{(2)}\partial_x\big)(\Phi^{(3)})^{-1} &= m_3 \partial_x^3 + a_1^{(2)} \partial_x + [\Pi_\perp b^{(3)} \partial_x^{-1},\, m_3 \partial_x^3 + a_1^{(2)} \partial_x] \\ &\quad + \int_0^1 (1-\tau) \Phi^{(3)}(\tau) \Big[\Pi_\perp b^{(3)} \partial_x^{-1}, \Big[\Pi_\perp b^{(3)}\partial_x^{-1}, m_3 \partial_x^3 + a_1^{(2)}\partial_x\Big]\Big] \Phi^{(3)}(\tau)^{-1}\, d\tau \\ &= m_3 \partial_x^3 + \big(a_1^{(2)} - 3 m_3 b_x^{(3)}\big)\partial_x + \mathcal{R}_0^{(II)}, \end{aligned}$$

$$\begin{aligned} \mathcal{R}_0^{(II)} &:= -3 m_3 b_{xx}^{(3)} - m_3 b_{xxx}^{(3)} \partial_x^{-1} + [\Pi_\perp b^{(3)} \partial_x^{-1}, a_1^{(2)} \partial_x] + [(\Pi_\perp - \mathrm{Id}) b^{(3)} \partial_x^{-1}, m_3 \partial_x^3] \\ &\quad + \int_0^1 (1-\tau)\Phi^{(3)}(\tau) \Big[\Pi_\perp b^{(3)} \partial_x^{-1},\ \Big[\Pi_\perp b^{(3)} \partial_x^{-1}, m_3 \partial_x^3 + a_1^{(2)}\partial_x\Big]\Big] \Phi^{(3)}(\tau)^{-1}\, d\tau \in OPS^0. \end{aligned} \tag{6.61}$$

Note that $\mathcal{R}_0^{(II)}$ is a pseudo-differential operator in OPS^0 (cf. Lemma 2.12). Hence, (6.60)-(6.61) and the choice of $b^{(3)}$ in (6.50) lead to the expansion (6.53) with $\mathcal{R}_M^{(3)}$ given by (6.60) and

$$\mathrm{Op}(r_0^{(3)}) := -\mathcal{R}_0^{(I)} + \mathcal{R}_0^{(II)}. \tag{6.62}$$

The estimate (6.54) follows by (6.22).

The estimate (6.55) on the operator $\mathrm{Op}(r_0^{(3)})$ follows by the definitions (6.60), (6.61), (6.62), by applying the estimates (6.18), (6.41), (6.42), (6.51), (6.59), (2.20), (2.21), (2.22), (2.24), (2.26) (using the ansatz (6.1) with μ_0 large enough). Next we estimate the remainder $\mathcal{R}_M^{(3)}$, defined in (6.60). We only consider the second term in the definition of $\mathcal{R}_M^{(3)}$, since the estimates the first and third terms can be obtained similarly. We recall that the operator $\mathcal{R}_M(Q_{-1}^{kdv}; \omega)$ is φ-independent. For $m \in \mathbb{S}_+$ and $\lambda, n_1, n_2 \in \mathbb{N}$ with $\lambda \leq \lambda_0$ and $n_1 + n_2 + \lambda_0 \leq M - 2$, one has

$$\begin{aligned} &\langle D\rangle^{n_1} \partial_{\varphi_m}^\lambda \Big((\Phi^{(3)} - \mathrm{Id}_\perp)\mathcal{R}_M(Q_{-1}^{kdv}; \omega)(\Phi^{(3)})^{-1}\Big)\langle D\rangle^{n_2} \\ &= \sum_{\lambda_1 + \lambda_2 = \lambda} C_{\lambda_1, \lambda_2} \langle D\rangle^{n_1} \partial_{\varphi_m}^{\lambda_1} (\Phi^{(3)} - \mathrm{Id}_\perp) \mathcal{R}_M(Q_{-1}^{kdv}; \omega) \partial_{\varphi_m}^{\lambda_2} (\Phi^{(3)})^{-1} \langle D\rangle^{n_2} \\ &= \sum_{\lambda_1 + \lambda_2 = \lambda} C_{\lambda_1, \lambda_2} \Big(\langle D\rangle^{n_1} \partial_{\varphi_m}^{\lambda_1}(\Phi^{(3)} - \mathrm{Id}_\perp)\langle D\rangle^{-n_1}\Big)\Big(\langle D\rangle^{n_1} \mathcal{R}_M(Q_{-1}^{kdv}; \omega)\langle D\rangle^{n_2}\Big)\Big(\langle D\rangle^{-n_2} \partial_{\varphi_m}^{\lambda_2}(\Phi^{(3)})^{-1}\langle D\rangle^{n_2}\Big). \end{aligned} \tag{6.63}$$

By the estimates (2.21), (2.24), (6.59) and Lemma 2.16, one has

$$
\begin{aligned}
&\mathfrak{M}_{\langle D\rangle^{n_1}\partial_{\varphi_m}^{\lambda_1}(\Phi^{(3)}-\mathrm{Id}_\perp)\langle D\rangle^{-n_1}}(s) \lesssim_s |\langle D\rangle^{n_1}\partial_{\varphi_m}^{\lambda_1}(\Phi^{(3)}-\mathrm{Id}_\perp)\langle D\rangle^{-n_1}|_{0,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)}\,,\\
&\mathfrak{M}_{\langle D\rangle^{-n_2}\partial_{\varphi_m}^{\lambda_2}(\Phi^{(3)})^{-1}\langle D\rangle^{n_2}}(s) \lesssim_s |\langle D\rangle^{-n_2}\partial_{\varphi_m}^{\lambda_2}(\Phi^{(3)})^{-1}\langle D\rangle^{n_2}|_{0,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} 1 + \|\iota\|_{s+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)}\,,
\end{aligned}
$$

and therefore, by Lemmata 2.14, 6.8 and using (6.1), the operator (6.63) satisfies (6.57). The estimates (6.56), (6.58) follow by similar arguments. □

6.5 Elimination of the φ-dependence of the first order term

The goal of this section is to remove the φ-dependence of the coefficient $a_1^{(3)}(\varphi)$ of the Hamiltonian operator $\mathcal{L}_\omega^{(3)}$ in (6.49), (6.53). We conjugate the operator $\mathcal{L}_\omega^{(3)}$ by the variable transformation $\Phi^{(4)} \equiv \Phi^{(4)}(\varphi)$,

$$
(\Phi^{(4)}w)(\varphi,x) = w(\varphi, x+b^{(4)}(\varphi))\,, \quad ((\Phi^{(4)})^{-1}h)(\varphi,x) = h(\varphi, x-b^{(4)}(\varphi))\,,
$$

where $b^{(4)}(\varphi)$ is a small, real valued, periodic function chosen in (6.65) below. Notice that $\Phi^{(4)}$ is the time-one flow of the transport equation $\partial_\tau w = b^{(4)}(\varphi)\partial_x w$. Each $\Phi^{(4)}(\varphi)$ is a symplectic linear isomorphism of $H_\perp^s(\mathbb{T}_1)$, and the conjugated operator

$$
\mathcal{L}_\omega^{(4)} := \Phi^{(4)}\mathcal{L}_\omega^{(3)}\big(\Phi^{(4)}\big)^{-1} \tag{6.64}
$$

is Hamiltonian.

Lemma 6.7. *Assume that $\omega \in \mathtt{DC}(\gamma,\tau)$. Let $b^{(4)}(\varphi)$ be the real valued, periodic function*

$$
b^{(4)}(\varphi;\omega) := -(\omega\cdot\partial_\varphi)^{-1}\big(a_1^{(3)}(\varphi;\omega) - m_1\big)\,, \quad m_1 := \frac{1}{(2\pi)^{|\mathbb{S}_+|}}\int_{\mathbb{T}^{\mathbb{S}_+}} a_1^{(3)}(\varphi;\omega)\,d\varphi \tag{6.65}
$$

and let $M\in\mathbb{N}$. Then there exists $\sigma_M>0$ with the following properties:
(i) The constant m_1 and the function $b^{(4)}$ satisfy

$$
|m_1|^{\mathrm{Lip}(\gamma)} \lesssim_M \varepsilon\gamma^{-2}\,, \quad \|b^{(4)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \gamma^{-1}\big(\varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}\big)\,, \quad \forall s\geq s_0\,. \tag{6.66}
$$

(ii) The Hamiltonian operator in (6.64) *admits an expansion of the form*

$$
\mathcal{L}_\omega^{(4)} = \omega\cdot\partial_\varphi - \big(m_3\partial_x^3 + m_1\partial_x + \mathrm{Op}(r_0^{(4)}) + Q_{-1}^{kdv}(D;\omega)\big) + \mathcal{R}_M^{(4)} \tag{6.67}
$$

where $r_0^{(4)} := r_0^{(4)}(\varphi,x,\xi;\omega)$ is a pseudo-differential symbol in S^0 satisfying for any $s\geq s_0$,

$$
|\mathrm{Op}(r_0^{(4)})|_{0,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}\,, \quad \forall s\geq s_0\,. \tag{6.68}
$$

Let $s_1\geq s_0$ and let $\breve\iota_1,\breve\iota_2$ be two tori satisfying (6.1) *with $\mu_0 \geq s_1+\sigma_M$. Then*

$$
|\Delta_{12}m_1|\,,\ \|\Delta_{12}b^{(4)}\|_{s_1} \lesssim_{s_1,M} \|\iota_1-\iota_2\|_{s_1+\sigma_M}\,, \quad |\Delta_{12}\mathrm{Op}(r_0^{(4)})|_{0,s_1,0} \lesssim_{s_1,M} \|\iota_1-\iota_2\|_{s_1+\sigma_M}\,. \tag{6.69}
$$

(iii) Let $S>\mathfrak{s}_M$. Then the maps $(\Phi^{(4)})^{\pm1}$ are Lip(γ)*-tame operators with a tame constant satisfying*

$$
\mathfrak{M}_{(\Phi^{(4)})^{\pm1}}(s) \lesssim_{S,M} 1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}\,, \quad \forall s_0\leq s\leq S\,. \tag{6.70}
$$

Let $\lambda_0\in\mathbb{N}$. Then there exists a constant $\sigma_M(\lambda_0)>0$ so that for any $\lambda,n_1,n_2\in\mathbb{N}$ with $\lambda\leq\lambda_0$ and $n_1+n_2+2\lambda_0\leq M-3$, the operator $\partial_{\varphi_m}^\lambda\langle D\rangle^{n_1}\mathcal{R}_M^{(4)}\langle D\rangle^{n_2}$, $m\in\mathbb{S}_+$, is Lip(γ)*-tame with a tame constant satisfying*

$$
\mathfrak{M}_{\partial_{\varphi_m}^\lambda\langle D\rangle^{n_1}\mathcal{R}_M^{(4)}\langle D\rangle^{n_2}}(s) \lesssim_{S,M,\lambda_0} \varepsilon + \|\iota\|_{s+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)}\,, \quad \forall \mathfrak{s}_M\leq s\leq S\,. \tag{6.71}
$$

Let $s_1\geq\mathfrak{s}_M$ and let $\breve\iota_1,\breve\iota_2$ be two tori satisfying (6.1) *with $\mu_0\geq s_1+\sigma_M(\lambda_0)$. Then*

$$
\|\partial_{\varphi_m}^\lambda\langle D\rangle^{n_1}\Delta_{12}\mathcal{R}_M^{(4)}\langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,M,\lambda_0} \|\iota_1-\iota_2\|_{s_1+\sigma_M(\lambda_0)}\,. \tag{6.72}
$$

Proof. The estimates (6.66) are direct consequences of (6.54) and of the ansatz (6.1). Notice that

$$\Phi^{(4)} \circ \omega \cdot \partial_\varphi \circ (\Phi^{(4)})^{-1} = \omega \cdot \partial_\varphi - \big(\omega \cdot \partial_\varphi b^{(4)}\big)\partial_x$$

and for any pseudo-differential operator $\mathrm{Op}(a(\varphi, x, \xi))$ a direct calculation shows that

$$\Phi^{(4)}\mathrm{Op}(a(\varphi, x, \xi))\big(\Phi^{(4)}\big)^{-1} = \mathrm{Op}(a(\varphi, x + b^{(4)}(\varphi), \xi)),$$

and hence, by recalling (6.53) and by the definition (6.65), one obtains (6.67) with

$$\mathrm{Op}\big(r_0^{(4)}(\varphi, x, \xi)\big) = \mathrm{Op}\big(r_0^{(3)}(\varphi, x + b^{(4)}(\varphi), \xi)\big), \quad \mathcal{R}_M^{(4)} := \Phi^{(4)}\mathcal{R}_M^{(3)}(\Phi^{(4)})^{-1}. \tag{6.73}$$

The estimates (6.68) follow by Lemma 2.1, using (6.66), (6.55) and the ansatz (6.1). The estimates (6.71) for the operator $\mathcal{R}_M^{(4)}$ follow by (6.57), (6.66) arguing as in the proof of the estimates of the remainder $\mathcal{R}_N(\tau, \varphi)$ (with β given by $b^{(4)}$) at the end of the proof of Proposition 2.28. The estimates (6.70) follow by Lemma 2.1 and (6.66). The estimates (6.69), (6.72) follow by similar arguments. □

7 KAM reduction of the linearized operator

The goal of this section is to complete the diagonalization of the Hamiltonian operator $\mathcal{L}_\omega$, started in Section 6. It remains to reduce the Hamiltonian operator $\mathcal{L}_\omega^{(4)}$ in (6.67). We are going to apply the KAM-reducibility scheme described in [9].

Recall that $\mathcal{L}_\omega^{(4)}$ is an operator acting on $H_\perp^s$. It is convenient to rename it as

$$\mathtt{L}_0 := \omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D}_0 + \mathtt{R}_0 \tag{7.1}$$

where $\omega \in \mathtt{DC}(\gamma, \tau)$ (cf. (4.4)) and in view of (6.7), (3.7), (4.3)

$$\mathtt{D}_0 := \mathrm{diag}_{j\in\mathbb{S}^\perp}(\mu_j^0), \quad \mu_j^0 := m_3(2\pi j)^3 - m_1 2\pi j - q_j(\omega), \quad q_j(\omega) := \omega_j^{kdv}\big(\nu(\omega), 0\big) - (2\pi j)^3, \tag{7.2}$$

$$\mathtt{R}_0 := -\mathrm{Op}(r_0^{(4)}) + \mathcal{R}_M^{(4)}. \tag{7.3}$$

Notice that $\mu_{-j}^0 = -\mu_j^0$ for any $j \in S_\perp$. By (3.59) we have

$$\sup_{j\in\mathbb{S}^\perp} |j||q_j|^{\mathrm{sup}}, \sup_{j\in\mathbb{S}^\perp} |j||q_j|^{\mathrm{lip}} \lesssim 1, \tag{7.4}$$

and, by (6.18), (6.66) and $\varepsilon\gamma^{-3} \le 1$,

$$|\mu_j^0 - \mu_{j'}^0|^{\mathrm{lip}} \lesssim_M |j^3 - j'^3|, \quad \forall j, j' \in \mathbb{S}^\perp. \tag{7.5}$$

The operator $\mathtt{R}_0$ satisfies the tame estimates of Lemma 7.1 below. We first *fix* the constants

$$\begin{aligned} &\mathtt{b} := [\mathtt{a}] + 2 \in \mathbb{N}, \quad \mathtt{a} := 3\tau_1 + 1, \quad \tau_1 := 2\tau + 1, \\ &\mu(\mathtt{b}) := s_0 + \mathtt{b} + \sigma_M + \sigma_M(\mathtt{b}) + 1, \quad M := 2(s_0 + \mathtt{b}) + 4, \end{aligned} \tag{7.6}$$

where the constants σ_M, $\sigma_M(\mathtt{b})$ are the ones introduced in Lemma 6.7 and where M is related to the order of smoothing of the remainder $\mathcal{R}_M^{(4)}$ in (6.67) (cf. (6.71)). Note that M only depends on the number of frequencies $|\mathbb{S}_+|$ and the diophantine constant τ.

Lemma 7.1. *Let* $\mathtt{b}$ *and* M *defined in* (7.6) *and* $S > \mathfrak{s}_M$ *with* $\mathfrak{s}_M$ *given by* (2.54).
(i) The operators $\mathtt{R}_0$, $[\mathtt{R}_0, \partial_x]$, $\partial_{\varphi_m}^{s_0}[\mathtt{R}_0, \partial_x]$, $\partial_{\varphi_m}^{s_0+\mathtt{b}}\mathtt{R}_0$, $\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathtt{R}_0, \partial_x]$, $m \in \mathbb{S}_+$, *are* Lip(γ)*-tame with tame constants*

$$\mathbb{M}_0(s) := \max_{m\in\mathbb{S}_+} \big\{\mathfrak{M}_{\mathtt{R}_0}(s), \mathfrak{M}_{[\mathtt{R}_0,\partial_x]}(s), \mathfrak{M}_{\partial_{\varphi_m}^{s_0}\mathtt{R}_0}(s), \mathfrak{M}_{\partial_{\varphi_m}^{s_0}[\mathtt{R}_0,\partial_x]}(s)\big\}, \tag{7.7}$$

$$\mathbb{M}_0(s, \mathtt{b}) := \max_{m\in\mathbb{S}_+} \big\{\mathfrak{M}_{\partial_{\varphi_m}^{s_0+\mathtt{b}}\mathtt{R}_0}(s), \mathfrak{M}_{\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathtt{R}_0,\partial_x]}(s)\big\}, \tag{7.8}$$

satisfying, for any $\mathfrak{s}_M \le s \le S$,

$$\mathfrak{M}_0(s,\mathtt{b}) := \max\{\mathbb{M}_0(s), \mathbb{M}_0(s,\mathtt{b})\} \lesssim_S \varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\mu(\mathtt{b})} \,. \tag{7.9}$$

Assuming that the ansatz (6.1) *holds with* $\mu_0 \ge \mathfrak{s}_M + \mu(\mathtt{b})$*, the latter estimate yields* $\mathfrak{M}_0(\mathfrak{s}_M,\mathtt{b}) \lesssim_S \varepsilon\gamma^{-2}$.
(ii) For any two tori $\breve{\iota}_1, \breve{\iota}_2$ *satisfying the ansatz* (6.1)*, one has for any* $m \in \mathbb{S}_+$ *and any* $\lambda \in \mathbb{N}$ *with* $\lambda \le s_0 + \mathtt{b}$

$$\|\partial^\lambda_{\varphi_m} \Delta_{12}\mathtt{R}_0\|_{\mathcal{B}(H^{\mathfrak{s}_M})},\ \|\partial^\lambda_{\varphi_m}[\Delta_{12}\mathtt{R}_0, \partial_x]\|_{\mathcal{B}(H^{\mathfrak{s}_M})} \lesssim \|\iota_1 - \iota_2\|_{\mathfrak{s}_M + \mu(\mathtt{b})} \,. \tag{7.10}$$

Proof. (i) Since the assertions for the various operators are proved in the same way, we restrict ourselves to show that there are tame constants $\mathfrak{M}_{\partial^{s_0+\mathtt{b}}_{\varphi_m}[\mathtt{R}_0,\partial_x]}(s)$, $m \in \mathbb{S}_+$, satisfying the bound in (7.9). The two operators $\mathrm{Op}(r_0^{(4)})$ and $\mathcal{R}_M^{(4)}$ in the definition (7.3) of $\mathtt{R}_0$ are treated separately. By Lemma 2.16 each operator $\partial^{s_0+\mathtt{b}}_{\varphi_m}[\mathrm{Op}(r_0^{(4)}), \partial_x] = -\mathrm{Op}\big(\partial^{s_0+\mathtt{b}}_{\varphi_m}\partial_x r_0^{(4)}\big)$, $m \in \mathbb{S}_+$, is $\mathrm{Lip}(\gamma)$-tame with a tame constant satisfying, for $s_0 \le s \le S$,

$$\begin{aligned} \mathfrak{M}_{\partial^{s_0+\mathtt{b}}_{\varphi_m}[\mathrm{Op}(r_0^{(4)}),\partial_x]}(s) &\overset{(2.31)}{\lesssim_s} \left|\mathrm{Op}\Big(\partial^{s_0+\mathtt{b}}_{\varphi_m}\partial_x r_0^{(4)}\Big)\right|^{\mathrm{Lip}(\gamma)}_{0,s,0} \lesssim_s \left|\mathrm{Op}(r_0^{(4)})\right|^{\mathrm{Lip}(\gamma)}_{0,s+s_0+\mathtt{b}+1,0} \\ &\overset{(6.68)}{\lesssim_s} \varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+s_0+\mathtt{b}+1+\sigma_M} \,. \end{aligned} \tag{7.11}$$

Next we treat $\partial^{s_0+\mathtt{b}}_{\varphi_m}[\mathcal{R}_M^{(4)}, \partial_x]$, $m \in \mathbb{S}_+$. Notice that

$$\partial^{s_0+\mathtt{b}}_{\varphi_m}[\mathcal{R}_M^{(4)}, \partial_x] = \partial^{s_0+\mathtt{b}}_{\varphi_m}\mathcal{R}_M^{(4)}\langle D\rangle\langle D\rangle^{-1}\partial_x - \langle D\rangle^{-1}\partial_x\langle D\rangle\partial^{s_0+\mathtt{b}}_{\varphi_m}\mathcal{R}_M^{(4)} \,.$$

Since there is a tame constant $\mathfrak{M}_{\langle D\rangle^{-1}\partial_x}(s)$ bounded by 1 it then follows by (6.71) that, for any $\mathfrak{s}_M \le s \le S$,

$$\mathfrak{M}_{\partial^{s_0+\mathtt{b}}_{\varphi_m}[\mathcal{R}_M^{(4)},\partial_x]}(s) \lesssim_S \varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M(\mathtt{b})} \,. \tag{7.12}$$

Combining (7.11), (7.12) and recalling the definition of $\mu(\mathtt{b})$ in (7.6) one obtains tame constants $\mathfrak{M}_{\partial^{s_0+\mathtt{b}}_{\varphi_m}[\mathtt{R}_0,\partial_x]}(s)$, $m \in \mathbb{S}_+$, satisfying the claimed bound.
(ii) The estimate (7.10) follows by similar arguments using (6.69) and (6.72) with $s_1 = \mathfrak{s}_M$. □

We perform the almost reducibility scheme for $\mathtt{L}_0$ along the scale

$$N_{-1} := 1 \,, \quad N_\nu := N_0^{\chi^\nu} \,,\ \nu \ge 0 \,, \quad \chi := 3/2 \,, \tag{7.13}$$

requiring at each induction step the second order Melnikov non-resonance conditions (7.18).

Theorem 7.2. (Almost reducibility) *There exists* $\overline{\tau} := \overline{\tau}(\tau, \mathbb{S}_+) > 0$ *so that for any* $S > \mathfrak{s}_M$*, there is* $N_0 := N_0(S,\mathtt{b}) \in \mathbb{N}$ *with the property that if*

$$N_0^{\overline{\tau}}\mathfrak{M}_0(\mathfrak{s}_M,\mathtt{b})\gamma^{-1} \le 1 \,, \tag{7.14}$$

then the following holds for any $\nu \in \mathbb{N}$*:*

(S1)$_\nu$ *There exists a Hamiltonian operator* $\mathtt{L}_\nu$*, acting on* $H^s_\perp$ *and defined for* $\omega \in \Omega^\gamma_\nu$*, of the form*

$$\mathtt{L}_\nu := \omega\cdot\partial_\varphi + \mathrm{i}\mathtt{D}_\nu + \mathtt{R}_\nu \,, \quad \mathtt{D}_\nu := \mathrm{diag}_{j\in\mathbb{S}^\perp}\mu_j^\nu \,, \quad \mu_j^\nu \in \mathbb{R} \,, \tag{7.15}$$

where for any $j \in \mathbb{S}^\perp$, μ_j^ν *is a* $\mathrm{Lip}(\gamma)$*-function of the form*

$$\mu_j^\nu(\omega) := \mu_j^0(\omega) + r_j^\nu(\omega) \,, \tag{7.16}$$

with

$$\mu^\nu_{-j} = -\mu^\nu_j \,, \qquad |r^\nu_j|^{\mathrm{Lip}(\gamma)} \le C(S)\varepsilon\gamma^{-2} \,, \tag{7.17}$$

and where $\mu_j^{(0)}$ *is defined in* (7.2)*. If* $\nu = 0$, Ω_ν^γ *is defined to be the set* $\Omega_0^\gamma := \mathtt{DC}(\gamma,\tau)$ *, and if* $\nu \geq 1$,

$$\Omega_\nu^\gamma := \Omega_\nu^\gamma(\iota) := \Big\{\omega \in \Omega_{\nu-1}^\gamma \, : |\omega\cdot\ell + \mu_j^{\nu-1} - \mu_{j'}^{\nu-1}| \geq \gamma\frac{|j^3 - j'^3|}{\langle\ell\rangle^\tau}, \forall|\ell| \leq N_{\nu-1}, j, j' \in \mathbb{S}^\perp\Big\}. \tag{7.18}$$

The operators $\mathtt{R}_\nu$ *and* $\langle\partial_\varphi\rangle^{\mathtt{b}}\mathtt{R}_\nu$ *are* $\mathrm{Lip}(\gamma)$*-modulo-tame with modulo-tame constants*

$$\mathfrak{M}_\nu^\sharp(s) := \mathfrak{M}_{\mathtt{R}_\nu}^\sharp(s)\,, \qquad \mathfrak{M}_\nu^\sharp(s,\mathtt{b}) := \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\mathtt{R}_\nu}(s)\,, \tag{7.19}$$

satisfying, for some $C_*(\mathfrak{s}_M,\mathtt{b}) > 0$*, for all* $s \in [\mathfrak{s}_M, S]$,

$$\mathfrak{M}_\nu^\sharp(s) \leq C_*(\mathfrak{s}_M,\mathtt{b})\mathfrak{M}_0(s,\mathtt{b})N_{\nu-1}^{-\mathtt{a}}\,, \qquad \mathfrak{M}_\nu^\sharp(s,\mathtt{b}) \leq C_*(\mathfrak{s}_M,\mathtt{b})\mathfrak{M}_0(s,\mathtt{b})N_{\nu-1}\,. \tag{7.20}$$

Moreover, if $\nu \geq 1$ *and* $\omega \in \Omega_\nu^\gamma$*, there exists a Hamiltonian operator* $\Psi_{\nu-1}$ *acting on* $H_\perp^s$*, so that the corresponding symplectic time one flow*

$$\Phi_{\nu-1} := \exp(\Psi_{\nu-1}) \tag{7.21}$$

conjugates $\mathtt{L}_{\nu-1}$ *to*

$$\mathtt{L}_\nu = \Phi_{\nu-1}\mathtt{L}_{\nu-1}\Phi_{\nu-1}^{-1}\,. \tag{7.22}$$

The operators $\Psi_{\nu-1}$ *and* $\langle\partial_\varphi\rangle^{\mathtt{b}}\Psi_{\nu-1}$ *are* $\mathrm{Lip}(\gamma)$*-modulo-tame with a modulo-tame constant satisfying, for all* $s \in [\mathfrak{s}_M, S]$*, (with* τ_1*,* $\mathtt{a}$ *defined in* (7.6)*)*

$$\mathfrak{M}^\sharp_{\Psi_{\nu-1}}(s) \leq \frac{C(\mathfrak{s}_M,\mathtt{b})}{\gamma}N_{\nu-1}^{\tau_1}N_{\nu-2}^{-\mathtt{a}}\mathfrak{M}_0(s,\mathtt{b})\,, \quad \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\Psi_{\nu-1}}(s) \leq \frac{C(\mathfrak{s}_M,\mathtt{b})}{\gamma}N_{\nu-1}^{\tau_1}N_{\nu-2}\mathfrak{M}_0(s,\mathtt{b})\,. \tag{7.23}$$

$(\mathbf{S2})_\nu$ *For any* $j \in \mathbb{S}^\perp$*, there exists a Lipschitz extension* $\widetilde{\mu}_j^\nu : \Omega \to \mathbb{R}$ *of* $\mu_j^\nu : \Omega_\nu^\gamma \to \mathbb{R}$*, where* $\widetilde{\mu}_j^0 = m_3(2\pi j)^3 - \widetilde{m}_1 2\pi j - q_j(\omega)$ *(cf.* (7.2)*) and* $\widetilde{m}_1 : \Omega \to \mathbb{R}$ *is an extension of* m_1 *satisfying* $|\widetilde{m}_1|^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}$*; if* $\nu \geq 1$,

$$|\widetilde{\mu}_j^\nu - \widetilde{\mu}_j^{\nu-1}|^{\mathrm{Lip}(\gamma)} \lesssim \mathfrak{M}_{\nu-1}^\sharp(\mathfrak{s}_M) \lesssim \mathfrak{M}_0(\mathfrak{s}_M,\mathtt{b})N_{\nu-1}^{-\mathtt{a}}\,.$$

$(\mathbf{S3})_\nu$ *Let* $\breve{\iota}_1$*,* $\breve{\iota}_2$ *be two tori satisfying* (6.1) *with* $\mu_0 \geq \mathfrak{s}_M + \mu(\mathtt{b})$*. Then, for all* $\omega \in \Omega_\nu^{\gamma_1}(\iota_1) \cap \Omega_\nu^{\gamma_2}(\iota_2)$ *with* $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$*, we have*

$$|||\mathtt{R}_\nu(\iota_1) - \mathtt{R}_\nu(\iota_2)|||_{\mathcal{B}(H^{\mathfrak{s}_M})} \lesssim_S N_{\nu-1}^{-\mathtt{a}}\|\iota_1 - \iota_2\|_{\mathfrak{s}_M+\mu(\mathtt{b})}, \tag{7.24}$$

$$|||\langle\partial_\varphi\rangle^{\mathtt{b}}(\mathtt{R}_\nu(\iota_1) - \mathtt{R}_\nu(\iota_2))|||_{\mathcal{B}(H^{\mathfrak{s}_M})} \lesssim_S N_{\nu-1}\|\iota_1 - \iota_2\|_{\mathfrak{s}_M+\mu(\mathtt{b})}\,. \tag{7.25}$$

Moreover, if $\nu \geq 1$*, then for any* $j \in \mathbb{S}^\perp$,

$$\big|(r_j^\nu(\iota_1) - r_j^\nu(\iota_2)) - (r_j^{\nu-1}(\iota_1) - r_j^{\nu-1}(\iota_2))\big| \lesssim |||\mathtt{R}_\nu(\iota_1) - \mathtt{R}_\nu(\iota_2)|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\,, \tag{7.26}$$

$$|r_j^\nu(\iota_1) - r_j^\nu(\iota_2)| \lesssim_S \|\iota_1 - \iota_2\|_{\mathfrak{s}_M+\mu(\mathtt{b})}\,. \tag{7.27}$$

$(\mathbf{S4})_\nu$ *Let* $\breve{\iota}_1$*,* $\breve{\iota}_2$ *be two tori as in* $(\mathbf{S3})_\nu$ *and* $0 < \rho < \gamma/2$*. Then*

$$C(S)N_{\nu-1}^\tau\|\iota_1 - \iota_2\|_{\mathfrak{s}_M+\mu(\mathtt{b})} \leq \rho \quad \Longrightarrow \quad \Omega_\nu^\gamma(\iota_1) \subseteq \Omega_\nu^{\gamma-\rho}(\iota_2)\,.$$

Theorem 7.2 implies that the symplectic invertible operator

$$U_n := \Phi_{n-1} \circ \ldots \circ \Phi_0, \quad n \geq 1\,, \tag{7.28}$$

almost diagonalizes $\mathtt{L}_0$, meaning that (7.31) below holds. The following corollary of Theorem 7.2 and Lemma 7.1 can be deduced as in [9].

Theorem 7.3. (KAM almost-reducibility) *Assume the ansatz* (6.1) *with* $\mu_0 \geq \mathfrak{s}_M + \mu(\mathtt{b})$*. Then for any* $S > \mathfrak{s}_M$ *there exist* $N_0 := N_0(S, \mathtt{b}) > 0$, $0 < \delta_0 := \delta_0(S) < 1$, *so that if*

$$N_0^{\overline{\tau}} \varepsilon \gamma^{-3} \leq \delta_0 \tag{7.29}$$

with $\overline{\tau} := \overline{\tau}(\tau, \mathbb{S}_+)$ *given by Theorem 7.2, the following holds: for any* $n \in \mathbb{N}$ *and any* ω *in*

$$\Omega_{n+1}^{\gamma} := \Omega_{n+1}^{\gamma}(\iota) = \bigcap_{\nu=0}^{n+1} \Omega_{\nu}^{\gamma} \tag{7.30}$$

with Ω_ν^γ *defined in* (7.18)*, the operator* U_n*, introduced in* (7.28)*, is well defined and* $\mathtt{L}_n := U_n \mathtt{L}_0 U_n^{-1}$ *satisfies*

$$\mathtt{L}_n = \omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D}_n + \mathtt{R}_n \tag{7.31}$$

where $\mathtt{D}_n$ *and* $\mathtt{R}_n$ *are defined in* (7.15) *(with* $\nu = n$*). The operator* $\mathtt{R}_n$ *is* $\mathrm{Lip}(\gamma)$*-modulo-tame with a modulo-tame constant*

$$\mathfrak{M}^\sharp_{\mathtt{R}_n}(s) \lesssim_S N_{n-1}^{-\mathtt{a}} (\varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\mu(\mathtt{b})}), \qquad \forall \mathfrak{s}_M \leq s \leq S\,. \tag{7.32}$$

Moreover, the operator $\mathtt{L}_n$ *is Hamiltonian,* U_n, U_n^{-1} *are symplectic, and* $U_n^{\pm 1} - \mathrm{Id}_\perp$ *are* $\mathrm{Lip}(\gamma)$*-modulo-tame with a modulo-tame constant satisfying*

$$\mathfrak{M}^\sharp_{U_n^{\pm1} - \mathrm{Id}_\perp}(s) \lesssim_S \gamma^{-1} N_0^{\tau_1} (\varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\mu(\mathtt{b})}), \quad \forall \mathfrak{s}_M \leq s \leq S\,, \tag{7.33}$$

where $\mathrm{Id}_\perp$ *denotes the identity operator on* $L^2_\perp(\mathbb{T}_1)$ *and* τ_1 *is defined in* (7.6).

7.1 Proof of Theorem 7.2

Proof of $(\mathbf{S1})_0$. Properties (7.15)-(7.17) for $\nu = 0$ follow by (7.1)-(7.2) with $r_j^0(\omega) = 0$. Moreover also (7.20) for $\nu = 0$ holds because, arguing as in Lemma 7.6 in [9], the following Lemma holds:

Lemma 7.4. $\mathfrak{M}_0^\sharp(s)$, $\mathfrak{M}_0^\sharp(s, \mathtt{b}) \lesssim_{\mathtt{b}} \mathfrak{M}_0(s, \mathtt{b})$ *where* $\mathfrak{M}_0(s, \mathtt{b})$ *is defined in* (7.9).

Proof of $(\mathbf{S2})_0$. For any $j \in \mathbb{S}^\perp$, μ_j^0 is defined in (7.2). Note that $m_3(\omega)$ and $q_j(\omega)$ are already defined on the whole parameter space Ω. By the Kirszbraun Theorem and (6.66) there is an extension $\widetilde{m}_1$ on Ω of m_1 satisfying the estimate $|\widetilde{m}_1|^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon \gamma^{-2}$. This proves $(\mathbf{S2})_0$.
Proof of $(\mathbf{S3})_0$. The estimates (7.24), (7.25) at $\nu = 0$ follows arguing as in the proof of $(\mathbf{S3})_0$ in [9].
Proof of $(\mathbf{S4})_0$. By the definition of Ω_0^γ one has $\Omega_0^\gamma(\iota_1) = \mathtt{DC}(\gamma, \tau) \subseteq \mathtt{DC}(\gamma - \rho, \tau) = \Omega_0^{\gamma-\rho}(\iota_2)$.
Iterative reductibility step. In what follows we describe how to define Ψ_ν, Φ_ν, $\mathtt{L}_{\nu+1}$ etc., at the iterative step. To simplify notation we drop the index ν and write $+$ instead of $\nu+1$. So, e.g. we write $\mathtt{L}$ for $\mathtt{L}_\nu$, $\mathtt{L}_+$ for $\mathtt{L}_{\nu+1}$, Ψ for Ψ_ν, etc. We conjugate $\mathtt{L}$ by the symplectic time one flow map

$$\Phi := \exp(\Psi) \tag{7.34}$$

generated by a Hamiltonian vector field Ψ acting in $H^s_\perp$. By a Lie expansion we get

$$\begin{aligned} \Phi \mathtt{L} \Phi^{-1} &= \Phi(\omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D})\Phi^{-1} + \Phi \mathtt{R} \Phi^{-1} \\ &= \omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D} - \omega \cdot \partial_\varphi \Psi - \mathrm{i}[\mathtt{D}, \Psi] + \Pi_N \mathtt{R} + \Pi_N^\perp \mathtt{R} - \int_0^1 \exp(\tau\Psi)[\mathtt{R}, \Psi]\exp(-\tau\Psi)\, d\tau \\ &\quad + \int_0^1 (1-\tau)\exp(\tau\Psi)\Big[\omega \cdot \partial_\varphi \Psi + \mathrm{i}[\mathtt{D}, \Psi], \Psi\Big]\exp(-\tau\Psi)\, d\tau \end{aligned} \tag{7.35}$$

where the projector Π_N is defined in (2.15) and $\Pi_N^\perp := \mathrm{Id}_\perp - \Pi_N$. We want to solve the homological equation

$$-\omega \cdot \partial_\varphi \Psi - \mathrm{i}[\mathtt{D}, \Psi] + \Pi_N \mathtt{R} = [\mathtt{R}] \qquad \text{where} \qquad [\mathtt{R}] := \mathrm{diag}_{j \in \mathbb{S}^\perp} \mathtt{R}_j^j(0)\,. \tag{7.36}$$

The solution of (7.36) is

$$\Psi_j^{j'}(\ell) := \begin{cases} \dfrac{\mathtt{R}_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell+\mu_j-\mu_{j'})} & \forall(\ell,j,j')\neq(0,j,j)\,,\ |\ell|\leq N\,,\ j,j'\in\mathbb{S}^\perp \\ 0 \quad \text{otherwise}\,. \end{cases} \tag{7.37}$$

The denominators in (7.37) are different from zero for $\omega\in\Omega_{\nu+1}^\gamma$ (cf. (7.18)).

Lemma 7.5. (Homological equations) *(i) The solution Ψ of the homological equation* (7.36)*, given by* (7.37) *for $\omega\in\Omega_{\nu+1}^\gamma$, is a* Lip$(\gamma)$*-modulo-tame operator with a modulo-tame constant satisfying*

$$\mathfrak{M}_\Psi^\sharp(s)\lesssim N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp(s)\,,\quad \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\Psi}(s)\lesssim N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp(s,\mathtt{b})\,, \tag{7.38}$$

where $\tau_1:=2\tau+1$. Moreover Ψ is Hamiltonian.
(ii) Let $\breve\imath_1$, $\breve\imath_2$ be two tori and define $\Delta_{12}\Psi:=\Psi(\imath_2)-\Psi(\imath_1)$. If $\gamma/2\leq\gamma_1,\gamma_2\leq 2\gamma$ then, for any $\omega\in\Omega_{\nu+1}^{\gamma_1}(\imath_1)\cap\Omega_{\nu+1}^{\gamma_2}(\imath_2)$,

$$|||\Delta_{12}\Psi|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\leq CN^{2\tau}\gamma^{-2}\big(|||\mathtt{R}(\imath_2)|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\|\imath_1-\imath_2\|_{\mathfrak{s}_M+\mu(\mathtt{b})}+|||\Delta_{12}\mathtt{R}|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\big)\,, \tag{7.39}$$

$$|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\Psi|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\lesssim_{\mathtt{b}}N^{2\tau}\gamma^{-2}\big(|||\langle\partial_\varphi\rangle^{\mathtt{b}}\mathtt{R}(\imath_2)|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\|\imath_1-\imath_2\|_{\mathfrak{s}_M+\mu(\mathtt{b})}+|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\mathtt{R}|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\big)\,. \tag{7.40}$$

Proof. Since $\mathtt{R}$ is Hamiltonian, one infers from Definition 2.4 and Lemma 2.5-(*iii*) that the operator Ψ defined in (7.37) is Hamiltonian as well. We now prove (7.38). Let $\omega\in\Omega_{\nu+1}^\gamma$. By (7.18), and the definition of Ψ in (7.37), it follows that for any $(\ell,j,j')\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp\times\mathbb{S}^\perp$, with $|\ell|\leq N$, $(\ell,j,j')\neq(0,j,j)$,

$$|\Psi_j^{j'}(\ell)|\lesssim\langle\ell\rangle^\tau\gamma^{-1}|\mathtt{R}_j^{j'}(\ell)| \tag{7.41}$$

and

$$\Delta_\omega\Psi_j^{j'}(\ell)=\frac{\Delta_\omega\mathtt{R}_j^{j'}(\ell)}{\delta_{\ell jj'}(\omega_1)}-\mathtt{R}_j^{j'}(\ell;\omega_2)\frac{\Delta_\omega\delta_{\ell jj'}}{\delta_{\ell jj'}(\omega_1)\delta_{\ell jj'}(\omega_2)}\,,\quad \delta_{\ell jj'}(\omega):=\mathrm{i}(\omega\cdot\ell+\mu_j-\mu_{j'})\,.$$

By (7.5), (7.16), (7.17) one gets $|\Delta_\omega\delta_{\ell jj'}|\lesssim(\langle\ell\rangle+|j^3-j'^3|)|\omega_1-\omega_2|$, and therefore, using also (7.18), we deduce that

$$|\Delta_\omega\Psi_j^{j'}(\ell)|\lesssim\langle\ell\rangle^\tau\gamma^{-1}|\Delta_\omega\mathtt{R}_j^{j'}(\ell)|+\langle\ell\rangle^{2\tau+1}\gamma^{-2}|\mathtt{R}_j^{j'}(\ell;\omega_2)||\omega_1-\omega_2|\,. \tag{7.42}$$

Recalling the definition (2.33), using (7.41), (7.42), and arguing as in the proof of the estimates (7.61) in [9, Lemma 7.7], one then deduces (7.38). The estimates (7.39)-(7.40) can be obtained by arguing similarly. $\square$

By (7.35)–(7.36) one has

$$\mathtt{L}_+=\Phi\mathtt{L}\Phi^{-1}=\omega\cdot\partial_\varphi+\mathrm{i}\mathtt{D}_++\mathtt{R}_+$$

which proves (7.22) and (7.15) at the step $\nu+1$, with

$$\begin{aligned}&\mathrm{i}\mathtt{D}_+:=\mathrm{i}\mathtt{D}+[\mathtt{R}]\,,\\ &\mathtt{R}_+=\Pi_N^\perp\mathtt{R}-\int_0^1\exp(\tau\Psi)[\mathtt{R},\Psi]\exp(-\tau\Psi)\,d\tau+\int_0^1(1-\tau)\exp(\tau\Psi)\big[\Pi_N\mathtt{R}-[\mathtt{R}],\Psi\big]\exp(-\tau\Psi)\,d\tau\,.\end{aligned} \tag{7.43}$$

The operator $\mathtt{L}_+$ has the same form as $\mathtt{L}$. More precisely, $\mathtt{D}_+$ is diagonal and $\mathtt{R}_+$ is the sum of an operator supported on high frequencies and one which is quadratic in Ψ and $\mathtt{R}$. The new normal form $\mathtt{D}_+$ has the following properties:

Lemma 7.6. (New diagonal part) *(i) The new normal form is*

$$\mathtt{D}_+=\mathtt{D}-\mathrm{i}[\mathtt{R}]\,,\qquad \mathtt{D}_+:=\mathrm{diag}_{j\in\mathbb{S}^\perp}\mu_j^+\,,\qquad \mu_j^+:=\mu_j+\mathtt{r}_j\in\mathbb{R}\,,\quad \mathtt{r}_j:=-\mathrm{i}\mathtt{R}_j^j(0)\,,\quad\forall j\in\mathbb{S}^\perp\,, \tag{7.44}$$

with

$$\mu_{-j}^+=-\mu_j^+\,,\quad |\mu_j^+-\mu_j|^{\mathrm{Lip}(\gamma)}=|\mathtt{r}_j|^{\mathrm{Lip}(\gamma)}\lesssim\mathfrak{M}^\sharp(\mathfrak{s}_M)\,.$$

(ii) For any tori $\breve\imath_1(\omega)$, $\breve\imath_2(\omega)$ and any $\omega\in\Omega_\nu^{\gamma_1}(\imath_1)\cap\Omega_\nu^{\gamma_2}(\imath_2)$, one has

$$|\mathtt{r}_j(\imath_1)-\mathtt{r}_j(\imath_2)|\lesssim|||\Delta_{12}\mathtt{R}|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\,. \tag{7.45}$$

Proof. By the definition (7.19) of $\mathfrak{M}^\sharp(\mathfrak{s}_M)$ and using (2.30) (with $\mathfrak{s}_M = s_1$) we have that $|\mu_j^+ - \mu_j|^{\mathrm{Lip}(\gamma)} \leq |\mathtt{R}_j^j(0)|^{\mathrm{Lip}(\gamma)} \lesssim \mathfrak{M}^\sharp(\mathfrak{s}_M)$. Since $\mathtt{R}(\varphi)$ is Hamiltonian, Lemma 2.5 implies that $\mathtt{r}_j = -\mathrm{i}\mathtt{R}_j^j(0)$, $j \in \mathbb{S}^\perp$, are odd in j and real. The estimate (7.45) is proved in the same way by using $|\Delta_{12}\mathtt{R}_j^j(0)| \leq C \| |\Delta_{12}\mathtt{R}| \|_{\mathcal{B}(H^{\mathfrak{s}_M})}$. $\square$

Induction. Assuming that the statements $(\mathbf{S1})_\nu$-$(\mathbf{S4})_\nu$ are true for some $\nu \geq 0$ we show in this paragraph that $(\mathbf{S1})_{\nu+1}$-$(\mathbf{S4})_{\nu+1}$ hold.

Proof of $(\mathbf{S1})_{\nu+1}$. By Lemma 7.5, for all $\omega \in \Omega_{\nu+1}^\gamma$ the solution Ψ_ν of the homological equation (7.36), defined in (7.37), is well defined and, by (7.38), (7.20), satisfies the estimates (7.23) at $\nu+1$. In particular, the estimate (7.23) for $\nu+1$, $s = \mathfrak{s}_M$ and (7.6), (7.14) imply

$$\mathfrak{M}^\sharp_{\Psi_\nu}(\mathfrak{s}_M) \lesssim_{\mathtt{b}} N_\nu^{\tau_1} N_{\nu-1}^{-\mathtt{a}} \gamma^{-1} \mathfrak{M}_0(\mathfrak{s}_M, \mathtt{b}) \leq 1 \,. \tag{7.46}$$

By Lemma 2.20 and using again Lemma 7.5 one infers that

$$\begin{aligned}
\mathfrak{M}^\sharp_{\Phi_\nu^{\pm 1}}(\mathfrak{s}_M) &\lesssim 1 \,, \\
\mathfrak{M}^\sharp_{\langle \partial_\varphi \rangle^{\mathtt{b}} \Phi_\nu^{\pm 1}}(\mathfrak{s}_M) &\lesssim 1 + \mathfrak{M}_{\langle \partial_\varphi \rangle^{\mathtt{b}} \Psi_\nu}(\mathfrak{s}_M) \lesssim 1 + N_\nu^{\tau_1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(\mathfrak{s}_M, \mathtt{b}) \,, \\
\mathfrak{M}^\sharp_{\Phi_\nu^{\pm 1}}(s) &\lesssim 1 + \mathfrak{M}^\sharp_{\Psi_\nu}(s) \lesssim_s 1 + N_\nu^{\tau_1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(s) \\
\mathfrak{M}^\sharp_{\langle \partial_\varphi \rangle^{\mathtt{b}} \Phi_\nu^{\pm 1}}(s) &\lesssim 1 + \mathfrak{M}_{\langle \partial_\varphi \rangle^{\mathtt{b}} \Psi_\nu}(s) + \mathfrak{M}^\sharp_{\Psi_\nu}(s) \mathfrak{M}_{\langle \partial_\varphi \rangle^{\mathtt{b}} \Psi_\nu}(\mathfrak{s}_M) \\
&\overset{(7.14),(7.20),(7.38)}{\lesssim} 1 + N_\nu^{\tau_1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(s, \mathtt{b}) + N_\nu^{2\tau_1} N_{\nu-1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(s) \,.
\end{aligned} \tag{7.47}$$

By Lemma 7.6, by the estimate (7.20) and Lemma 7.1, the operator $\mathtt{D}_{\nu+1}$ is diagonal and its eigenvalues $\mu_j^{\nu+1} : \Omega_{\nu+1}^\gamma \to \mathbb{R}$ satisfy (7.17) at $\nu+1$.

Now we estimate the remainder $\mathtt{R}_{\nu+1}$ defined in (7.43).

Lemma 7.7. (Nash-Moser iterative scheme) *The operator $\mathtt{R}_{\nu+1}$ is* $\mathrm{Lip}(\gamma)$*-modulo-tame with a modulo-tame constant satisfying*

$$\mathfrak{M}^\sharp_{\nu+1}(s) \lesssim N_\nu^{-\mathtt{b}} \mathfrak{M}^\sharp_\nu(s, \mathtt{b}) + N_\nu^{\tau_1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(s) \mathfrak{M}^\sharp_\nu(\mathfrak{s}_M) \,. \tag{7.48}$$

The operator $\langle \partial_\varphi \rangle^{\mathtt{b}} \mathtt{R}_{\nu+1}$ is $\mathrm{Lip}(\gamma)$*-modulo-tame with a modulo-tame constant satisfying*

$$\mathfrak{M}^\sharp_{\nu+1}(s, \mathtt{b}) \lesssim_{\mathtt{b}} \mathfrak{M}^\sharp_\nu(s, \mathtt{b}) + N_\nu^{\tau_1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(s, \mathtt{b}) \mathfrak{M}^\sharp_\nu(\mathfrak{s}_M) + N_\nu^{\tau_1} \gamma^{-1} \mathfrak{M}^\sharp_\nu(\mathfrak{s}_M, \mathtt{b}) \mathfrak{M}^\sharp_\nu(s) \,. \tag{7.49}$$

Proof. The proof follows by Lemmata 2.21, 2.19, using the estimates (7.20), (7.38), (7.47). $\square$

The estimates (7.48), (7.49), and (7.6), allow to prove that also (7.20) holds at the step $\nu+1$. It implies (see [9, Lemma 7.10])

Lemma 7.8. $\mathfrak{M}^\sharp_{\nu+1}(s) \leq C_*(\mathfrak{s}_M, \mathtt{b}) N_\nu^{-\mathtt{a}} \mathfrak{M}_0(s, \mathtt{b})$ *and* $\mathfrak{M}^\sharp_{\nu+1}(s, \mathtt{b}) \leq C_*(\mathfrak{s}_M, \mathtt{b}) N_\nu \mathfrak{M}_0(s, \mathtt{b})$.

Proof of $(\mathbf{S2})_{\nu+1}$. By Lemma 7.6, for any $j \in \mathbb{S}^\perp$, $\mu_j^{\nu+1} = \mu_j^\nu + \mathtt{r}_j^\nu$ where $|\mathtt{r}_j^\nu|^{\mathrm{Lip}(\gamma)} \lesssim \mathfrak{M}_0(\mathfrak{s}_M, \mathtt{b}) N_\nu^{-\mathtt{a}}$. Then $(\mathbf{S2})_{\nu+1}$ follows by defining $\widetilde{\mu}_j^{\nu+1} := \widetilde{\mu}_j^\nu + \widetilde{\mathtt{r}}_j^\nu$ where $\widetilde{\mathtt{r}}_j^\nu : \Omega \to \mathbb{R}$ is a Lipschitz extension of $\mathtt{r}_j^\nu$ (cf. Kirszbraun extension Theorem).

Proof of $(\mathbf{S3})_{\nu+1}$. The proof follows by induction arguing as in the proof of $(\mathbf{S2})_{\nu+1}$.

Proof of $(\mathbf{S4})_{\nu+1}$. The proof is the same as that of $(\mathbf{S3})_{\nu+1}$ in [2, Theorem 4.2]. $\square$

7.2 Almost-invertibility of $\mathcal{L}_\omega$

By (7.31), for any $\omega \in \Omega_n^\gamma$, we have that $\mathtt{L}_0 = U_n^{-1} \mathtt{L}_n U_n$ where U_n is defined in (7.28) and thus

$$\mathcal{L}_\omega = \mathcal{V}_n^{-1} \mathtt{L}_n \mathcal{V}_n \,, \qquad \mathcal{V}_n := U_n \Phi^{(4)} \cdots \Phi^{(1)} \,. \tag{7.50}$$

Lemma 7.9. *There exists $\sigma = \sigma(\tau, \mathbb{S}_+) > 0$ such that, if* (7.29) *and* (6.1) *with $\mu_0 \geq \mathfrak{s}_M + \mu(\mathtt{b}) + \sigma$ hold, then the operators $\mathcal{V}_n^{\pm 1}$ satisfy for any $\mathfrak{s}_M \leq s \leq S$ the estimate*

$$\|\mathcal{V}_n^{\pm 1} h\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \|h\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1} \gamma^{-1} \|\iota\|_{s+\mu(\mathtt{b})+\sigma}^{\mathrm{Lip}(\gamma)} \|h\|_{\mathfrak{s}_M+\sigma}^{\mathrm{Lip}(\gamma)} . \tag{7.51}$$

Proof. By the estimates (6.28), (6.44), (6.52), (6.70), using Lemmata 2.14, 2.15, 2.18 and (7.33). □

We now decompose the operator $\mathtt{L}_n$ in (7.31) as

$$\mathtt{L}_n = \mathfrak{L}_n^{<} + \mathtt{R}_n + \mathtt{R}_n^{\perp} \tag{7.52}$$

where

$$\mathfrak{L}_n^{<} := \Pi_{K_n} \big(\omega \cdot \partial_\varphi + \mathrm{i} \mathtt{D}_n \big) \Pi_{K_n} + \Pi_{K_n}^{\perp} , \quad \mathtt{R}_n^{\perp} := \Pi_{K_n}^{\perp} \big(\omega \cdot \partial_\varphi + \mathrm{i} \mathtt{D}_n \big) \Pi_{K_n}^{\perp} - \Pi_{K_n}^{\perp} , \tag{7.53}$$

the diagonal operator $\mathtt{D}_n$ is defined in (7.15) (with $\nu = n$), and $K_n := K_0^{\chi^n}$ is the scale of the nonlinear Nash-Moser iterative scheme introduced in (5.24).

Lemma 7.10. (First order Melnikov non-resonance conditions) *For all ω in*

$$\Lambda_{n+1}^{\gamma} := \Lambda_{n+1}^{\gamma}(\iota) := \big\{ \omega \in \Omega : |\omega \cdot \ell + \widetilde{\mu}_j^n| \geq 2\gamma |j|^3 \langle \ell \rangle^{-\tau} , \quad \forall |\ell| \leq K_n , \ j \in \mathbb{S}^{\perp} \big\} , \tag{7.54}$$

the operator $\mathfrak{L}_n^{<}$ in (7.53) *is invertible and*

$$\|(\mathfrak{L}_n^{<})^{-1} g\|_s^{\mathrm{Lip}(\gamma)} \lesssim \gamma^{-1} \|g\|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)} . \tag{7.55}$$

By (7.50), (7.52), Theorem 7.3, estimates (7.55), (7.56), (7.51), and using that, for all $b > 0$,

$$\|\mathtt{R}_n^{\perp} h\|_{\mathfrak{s}_M}^{\mathrm{Lip}(\gamma)} \lesssim K_n^{-b} \|h\|_{\mathfrak{s}_M+b+3}^{\mathrm{Lip}(\gamma)} , \quad \|\mathtt{R}_n^{\perp} h\|_s^{\mathrm{Lip}(\gamma)} \lesssim \|h\|_{s+3}^{\mathrm{Lip}(\gamma)} , \tag{7.56}$$

we deduce the following theorem, stating the almost-invertibility assumption of $\mathcal{L}_\omega$ of Section 5.

Theorem 7.11. (Almost-invertibility of $\mathcal{L}_\omega$) *Let $\mathtt{a}, \mathtt{b}, M$ as in* (7.6) *and $S > \mathfrak{s}_M$. There exists $\sigma = \sigma(\tau, \mathbb{S}_+) > 0$ such that, if* (7.29) *and* (6.1) *with $\mu_0 \geq \mathfrak{s}_M + \mu(\mathtt{b}) + \sigma$ hold, then, for all*

$$\omega \in \mathbf{\Omega}_{n+1}^{\gamma} := \mathbf{\Omega}_{n+1}^{\gamma}(\iota) := \Omega_{n+1}^{\gamma} \cap \Lambda_{n+1}^{\gamma} \tag{7.57}$$

(see (7.30), (7.54)*), the operator $\mathcal{L}_\omega$ defined in* (5.22) *can be decomposed as*

$$\begin{aligned} &\mathcal{L}_\omega = \mathcal{L}_\omega^{<} + \mathcal{R}_\omega + \mathcal{R}_\omega^{\perp} , \\ &\mathcal{L}_\omega^{<} := \mathcal{V}_n^{-1} \mathfrak{L}_n^{<} \mathcal{V}_n , \quad \mathcal{R}_\omega := \mathcal{V}_n^{-1} \mathtt{R}_n \mathcal{V}_n , \quad \mathcal{R}_\omega^{\perp} := \mathcal{V}_n^{-1} \mathtt{R}_n^{\perp} \mathcal{V}_n , \end{aligned} \tag{7.58}$$

where $\mathcal{L}_\omega^{<}$ is invertible and satisfies (5.28) *and the operators $\mathcal{R}_\omega$ and $\mathcal{R}_\omega^{\perp}$ satisfy* (5.26)-(5.27).

8 Proof of Theorem 4.1

Theorem 4.1 is a consequence of Theorem 8.1 below where we construct iteratively a sequence of better and better approximate solutions of the equation $\mathcal{F}_\omega(\iota, \zeta) = 0$ where $\mathcal{F}_\omega$ is defined in (4.6).

8.1 The Nash-Moser iteration

We consider the finite-dimensional subspaces of $L^2_\varphi \times L^2_\varphi \times L^2_\perp$, defined for any $n \in \mathbb{N}$ as

$$\mathtt{E}_n := \big\{ \iota(\varphi) = (\Theta, y, w)(\varphi), \ \ \Theta = \Pi_n \Theta, \ y = \Pi_n y, \ w = \Pi_n w \big\}$$

where $L^2_\varphi = L^2_\varphi(\mathbb{T}_1 \times \mathbb{R}^{\mathbb{S}_+})$ (cf. (4.8)) and where $\Pi_n := \Pi_{K_n} : L^2_\perp \to \cap_{s \geq 0} H^s_\perp$ is the projector (cf. (2.2))

$$\Pi_n : w = \sum_{\ell \in \mathbb{Z}^{\mathbb{S}_+}, j \in \mathbb{S}^{\perp}} w_{\ell,j} e^{\mathrm{i}(\ell \cdot \varphi + 2\pi j x)} \quad \mapsto \quad \Pi_n w := \sum_{|(\ell,j)| \leq K_n} w_{\ell,j} e^{\mathrm{i}(\ell \cdot \varphi + 2\pi j x)} \tag{8.1}$$

with $K_n = K_0^{\chi^n}$ (cf. (5.24)) and also denotes the corresponding one on L^2_φ, given by $L^2_\varphi \to \cap_{s\geq 0} H^s_\varphi$, $p = \sum_{\ell\in\mathbb{Z}^{\mathbb{S}_+}} p_\ell e^{\mathrm{i}\ell\cdot\varphi} \mapsto \sum_{|\ell|\leq K_n} p_\ell e^{\mathrm{i}\ell\cdot\varphi}$. Note that Π_n, $n \geq 1$, are smoothing operators for the Sobolev spaces $H^s_\perp$. In particular Π_n and $\Pi_n^\perp := \mathrm{Id} - \Pi_n$ satisfy the smoothing properties (2.3). For the Nash-Moser Theorem 8.1, stated below, we introduce the constants

$$\overline{\sigma} := \max\{\sigma_1, \sigma_2\}\,, \qquad \mathtt{b} := [\mathtt{a}] + 2\,, \qquad \mathtt{a} = 3\tau_1 + 1\,, \qquad \tau_1 = 2\tau + 1\,, \qquad \chi = 3/2\,, \tag{8.2}$$

$$\mathtt{a}_1 := \max\{12\overline{\sigma} + 13,\ p\tau + 3 + \chi(\mu(\mathtt{b}) + 2\overline{\sigma})\}, \quad \mathtt{a}_2 := \chi^{-1}\mathtt{a}_1 - \mu(\mathtt{b}) - 2\overline{\sigma}\,, \tag{8.3}$$

$$\mathtt{b}_1 := \mathtt{a}_1 + \mu(\mathtt{b}) + 3\overline{\sigma} + 4 + \frac{2}{3}\mu_1\,, \qquad \mu_1 := 3(\mu(\mathtt{b}) + 2\overline{\sigma} + 2) + 1\,, \quad S := \mathfrak{s}_M + \mathtt{b}_1\,, \tag{8.4}$$

where σ_1 is defined in Lemma 4.2, σ_2 in Theorem 5.6, and $\mathtt{a}$, $\mu(\mathtt{b})$ in (7.6). The number p is the exponent in (5.23) and is requested to satisfy

$$p\mathtt{a} > (\chi - 1)\mathtt{a}_1 + \chi(\overline{\sigma} + 4) = \frac{1}{2}\mathtt{a}_1 + \frac{3}{2}(\overline{\sigma} + 4)\,. \tag{8.5}$$

In view of the definition (8.3) of $\mathtt{a}_1$, we can define $p := p(\tau, \mathbb{S}_+)$ as

$$p := \frac{12\overline{\sigma} + 17 + \chi(\mu(\mathtt{b}) + 2\overline{\sigma})}{\mathtt{a}}\,. \tag{8.6}$$

We denote by $\|W\|_s^{\mathrm{Lip}(\gamma)} := \max\{\|\iota\|_s^{\mathrm{Lip}(\gamma)}, |\zeta|^{\mathrm{Lip}(\gamma)}\}$ the norm of a function

$$W := (\iota, \zeta) : \Omega \to \left(H^s_\varphi \times H^s_\varphi \times H^s_\perp\right) \times \mathbb{R}^{\mathbb{S}_+}\,, \quad \omega \mapsto W(\omega) = (\iota(\omega), \zeta(\omega))\,.$$

The following Nash-Moser Theorem can be proved in a by now standard way as in [9], [1].

Theorem 8.1. (Nash-Moser) *There exist $0 < \delta_0 < 1$, $C_* > 0$ so that if*

$$\varepsilon K_0^{\tau_2} < \delta_0, \quad \tau_2 := \max\{p\overline{\tau} + 3,\ 4\overline{\sigma} + 4 + \mathtt{a}_1\}\,, \quad K_0 := \gamma^{-1}, \quad \gamma := \varepsilon^{\mathfrak{a}}\,, \quad 0 < \mathfrak{a} < \frac{1}{\tau_2}\,, \tag{8.7}$$

where $\overline{\tau} := \overline{\tau}(\tau, \mathbb{S}_+)$ is defined in Theorem 7.2, then the following holds for all $n \in \mathbb{N}$:

$(\mathcal{P}1)_n$ *Let $\tilde{W}_0 := (0, 0)$. For $n \geq 1$, there exists a* $\mathrm{Lip}(\gamma)$*-function $\tilde{W}_n : \mathbb{R}^{\mathbb{S}_+} \to \mathtt{E}_{n-1} \times \mathbb{R}^{\mathbb{S}_+}$, $\omega \mapsto \tilde{W}_n(\omega) := (\tilde{\iota}_n, \tilde{\zeta}_n)$, satisfying*

$$\|\tilde{W}_n\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M + \mu(\mathtt{b}) + \overline{\sigma}} \lesssim \varepsilon\gamma^{-2}\,. \tag{8.8}$$

Let $\tilde{U}_n := U_0 + \tilde{W}_n$ where $U_0 := (\varphi, 0, 0, 0)$. For $n \geq 1$, the difference $\tilde{H}_n := \tilde{U}_n - \tilde{U}_{n-1}$, , satisfies

$$\|\tilde{H}_1\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M + \mu(\mathtt{b}) + \overline{\sigma}} \lesssim \varepsilon\gamma^{-2}\,, \qquad \|\tilde{H}_n\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M + \mu(\mathtt{b}) + \overline{\sigma}} \lesssim \varepsilon\gamma^{-2} K_{n-1}^{-\mathtt{a}_2}\,, \quad \textit{for } n \geq 2\,. \tag{8.9}$$

$(\mathcal{P}2)_n$ *Let $\mathcal{G}_0 := \Omega$ and define for $n \geq 1$,*

$$\mathcal{G}_n := \mathcal{G}_{n-1} \cap\ \mathbf{\Omega}^\gamma_n(\tilde{\iota}_{n-1})\,, \tag{8.10}$$

where $\mathbf{\Omega}^\gamma_n(\tilde{\iota}_{n-1})$ is defined in (7.57). Then for any $\omega \in \mathcal{G}_n$

$$\|\mathcal{F}_\omega(\tilde{U}_n)\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M} \leq C_*\varepsilon K_{n-1}^{-\mathtt{a}_1}\,, \qquad K_{-1} := 1. \tag{8.11}$$

$(\mathcal{P}3)_n$ (High norms) $\quad \|\tilde{W}_n\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M + \mathtt{b}_1} \leq C_*\varepsilon K_{n-1}^{\mu_1}$, $\forall \omega \in \mathcal{G}_n$.

Proof. We argue by induction. To simplify notation, we write within this proof $\|\cdot\|$ for $\|\cdot\|^{\mathrm{Lip}(\gamma)}$.

Step 1: *Proof of* $(\mathcal{P}1, \mathcal{P}2, \mathcal{P}3)_0$. Note that $(\mathcal{P}1)_0$ and $(\mathcal{P}3)_0$ are trivially satisfied and hence it remains to verify (8.11) at $n = 0$. By (4.6), (4.10), (4.3), and Lemma 4.2, there exists $C_* > 0$ large enough so that $\|\mathcal{F}_\omega(U_0)\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M} \leq \varepsilon C_*$.

STEP 2: *Proof of the induction step.* Assuming that $(\mathcal{P}1,\mathcal{P}2,\mathcal{P}3)_n$ hold for some $n \geq 0$, we have to prove that $(\mathcal{P}1,\mathcal{P}2,\mathcal{P}3)_{n+1}$ hold. We are going to define the approximation $\tilde{U}_{n+1}$ by a modified Nash-Moser scheme. To this aim, we prove the almost-approximate invertibility of the linearized operator

$$L_n := L_n(\omega) := d_{\iota,\zeta}\mathcal{F}_\omega(\tilde{\iota}_n(\omega)) \tag{8.12}$$

by applying Theorem 5.6 to $L_n(\omega)$. To prove that the inversion assumptions (5.25)-(5.28) hold, we apply Theorem 7.11 with $\iota = \tilde{\iota}_n$.

By choosing ε small enough it follows by (8.7) that $N_0 = K_0^p = \gamma^{-p} = \varepsilon^{-p\mathtt{a}}$ satisfies the requirement of Theorem 7.11 and that the smallness condition (7.29) holds. Therefore Theorem 7.11 applies, and we deduce that (5.25)-(5.28) hold for all $\omega \in \mathbf{\Omega}^\gamma_{n+1}(\tilde{\iota}_n)$, see (7.57).

Now we apply Theorem 5.6 to the linearized operator $L_n(\omega)$ with $\Omega_o = \mathbf{\Omega}^\gamma_{n+1}(\tilde{\iota}_n)$ and $S = \mathfrak{s}_M + \mathtt{b}_1$, see (8.4). It implies the existence of an almost-approximate inverse $\mathbf{T}_n := \mathbf{T}_n(\omega, \tilde{\iota}_n(\omega))$ satisfying

$$\|\mathbf{T}_n g\|_s \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} \gamma^{-2}\big(\|g\|_{s+\overline{\sigma}} + K_0^{\tau_1 p}\gamma^{-1}\|\tilde{\iota}_n\|_{s+\mu(\mathtt{b})+\overline{\sigma}}\|g\|_{\mathfrak{s}_M+\overline{\sigma}}\big)\,, \quad \forall \mathfrak{s}_M \leq s \leq \mathfrak{s}_M + \mathtt{b}_1\,, \tag{8.13}$$

where we used that $\overline{\sigma} \geq \sigma_2$ (cf. (8.2)), σ_2 is the loss of regularity constant appearing in the estimate (5.43), and $N_0 = K_0^p$. Furthermore, by (8.7), (8.8) one obtains that

$$K_0^{\tau_1 p}\gamma^{-1}\|\tilde{W}_n\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}} \leq 1\,, \tag{8.14}$$

therefore (8.13) specialized for $s = \mathfrak{s}_M$ becomes

$$\|\mathbf{T}_n g\|_{\mathfrak{s}_M} \lesssim_{\mathtt{b}_1} \gamma^{-2}\|g\|_{\mathfrak{s}_M+\overline{\sigma}}\,. \tag{8.15}$$

For all $\omega \in \mathcal{G}_{n+1} = \mathcal{G}_n \cap \mathbf{\Lambda}^\gamma_{n+1}(\tilde{\iota}_n)$ (see (8.10)), we define

$$U_{n+1} := \tilde{U}_n + H_{n+1}\,, \qquad H_{n+1} := (\widehat{\iota}_{n+1}, \widehat{\zeta}_{n+1}) := -\mathbf{\Pi}_n\mathbf{T}_n\Pi_n\mathcal{F}_\omega(\tilde{U}_n) \in \mathtt{E}_n \times \mathbb{R}^{\mathbb{S}_+} \tag{8.16}$$

where $\mathbf{\Pi}_n$ is defined by (see (8.1))

$$\mathbf{\Pi}_n(\iota,\zeta) := (\Pi_n\iota, \zeta)\,, \quad \mathbf{\Pi}_n^\perp(\iota,\zeta) := (\Pi_n^\perp\iota, 0)\,, \quad \forall(\iota,\zeta)\,. \tag{8.17}$$

We show that the iterative scheme in (8.16) is rapidly converging. We write

$$\mathcal{F}_\omega(U_{n+1}) = \mathcal{F}_\omega(\tilde{U}_n) + L_n H_{n+1} + Q_n \tag{8.18}$$

where $L_n := d_{\iota,\zeta}\mathcal{F}_\omega(\tilde{U}_n)$ and Q_n is defined by (8.18). Then, by the definition of H_{n+1} in (8.16), we have (recall also (8.17))

$$\begin{aligned}\mathcal{F}_\omega(U_{n+1}) &= \mathcal{F}_\omega(\tilde{U}_n) - L_n\mathbf{\Pi}_n\mathbf{T}_n\Pi_n\mathcal{F}_\omega(\tilde{U}_n) + Q_n \\ &= \mathcal{F}_\omega(\tilde{U}_n) - L_n\mathbf{T}_n\Pi_n\mathcal{F}_\omega(\tilde{U}_n) + L_n\mathbf{\Pi}_n^\perp\mathbf{T}_n\Pi_n\mathcal{F}_\omega(\tilde{U}_n) + Q_n \\ &= \Pi_n^\perp\mathcal{F}_\omega(\tilde{U}_n) + R_n + Q_n + P_n\end{aligned} \tag{8.19}$$

where

$$R_n := L_n\mathbf{\Pi}_n^\perp\mathbf{T}_n\Pi_n\mathcal{F}_\omega(\tilde{U}_n)\,, \qquad P_n := -(L_n\mathbf{T}_n - \mathrm{Id})\Pi_n\mathcal{F}_\omega(\tilde{U}_n)\,. \tag{8.20}$$

We first note that for any $\omega \in \Omega$, $s \geq \mathfrak{s}_M$ one has by the triangular inequality, (4.6), Lemma 4.2, and (8.2), (8.8)

$$\|\mathcal{F}_\omega(\tilde{U}_n)\|_s \lesssim_s \|\mathcal{F}_\omega(U_0)\|_s + \|\mathcal{F}_\omega(\tilde{U}_n) - \mathcal{F}_\omega(U_0)\|_s \lesssim_s \varepsilon + \|\tilde{W}_n\|_{s+\overline{\sigma}} \tag{8.21}$$

and, by (8.8), (8.7), (8.11)

$$K_0^{\tau_1 p}\gamma^{-1}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M} \leq 1\,. \tag{8.22}$$

We now prove the following inductive estimates of Nash-Moser type.

Lemma 8.2. *For all $\omega \in \mathcal{G}_{n+1}$ we have, setting $\mu_2 := \mu(\mathtt{b}) + 3\overline{\sigma} + 3$,*

$$\|\mathcal{F}_\omega(U_{n+1})\|_{\mathfrak{s}_M} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{\mu_2-\mathtt{b}_1}(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}) + K_n^{4\overline{\sigma}+4}\|\mathcal{F}_\omega(\tilde{U}_n)\|^2_{\mathfrak{s}_M} + \varepsilon K_{n-1}^{-p\mathtt{a}}K_n^{\overline{\sigma}+4}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M} \tag{8.23}$$

$$\|W_1\|_{\mathfrak{s}_M+\mathtt{b}_1} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_0^2\varepsilon\,, \qquad \|W_{n+1}\|_{\mathfrak{s}_M+\mathtt{b}_1} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+2\overline{\sigma}+2}(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1})\,, \ n \geq 1\,. \tag{8.24}$$

Proof. We first estimate H_{n+1} defined in (8.16).

Estimates of H_{n+1}. By (8.16) and (2.3), (8.13), (8.8), we get

$$\|H_{n+1}\|_{\mathfrak{s}_M+\mathtt{b}_1} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} \gamma^{-2}\big(K_n^{\overline{\sigma}}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M+\mathtt{b}_1} + K_n^{\mu(\mathtt{b})+2\overline{\sigma}}K_0^{\tau_1 p}\gamma^{-1}\|\tilde{\imath}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}\big)$$

$$\overset{(8.21),(8.22)}{\lesssim_{\mathfrak{s}_M+\mathtt{b}_1}} K_n^{\mu(\mathtt{b})+2\overline{\sigma}}\gamma^{-2}\big(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}\big) \tag{8.25}$$

$$\overset{\gamma^{-1}=K_0\le K_n}{\lesssim_{\mathfrak{s}_M+\mathtt{b}_1}} K_n^{\mu(\mathtt{b})+2\overline{\sigma}+2}\big(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}\big)\,, \tag{8.26}$$

$$\|H_{n+1}\|_{\mathfrak{s}_M} \overset{(8.15)}{\lesssim_{\mathfrak{s}_M+\mathtt{b}_1}} \gamma^{-2}K_n^{\overline{\sigma}}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}\,. \tag{8.27}$$

Next we estimate the terms Q_n in (8.18) and P_n, R_n in (8.20) in $\|\ \|_{\mathfrak{s}_M}$ norm.

Estimate of Q_n. By (8.8), (8.16), (2.3), (8.27), (8.11), and since $\chi 2\overline{\sigma} - \mathtt{a}_1 \le 0$ (see (8.3)), we deduce that $\|\tilde{W}_n + tH_{n+1}\|_{\mathfrak{s}_M+\overline{\sigma}} \lesssim \varepsilon\gamma^{-2}K_0^{2\overline{\sigma}}$ for all $t \in [0,1]$. Since $\gamma^{-1} = K_0$, by (8.7) we can apply Lemma 4.2 and by Taylor's formula, using (8.18), (4.6), (8.27), (2.3), and $\gamma^{-1} = K_0 \le K_n$, we get

$$\|Q_n\|_{\mathfrak{s}_M} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} \|H_{n+1}\|^2_{\mathfrak{s}_M+\overline{\sigma}} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{4\overline{\sigma}+4}\|\mathcal{F}_\omega(\tilde{U}_n)\|^2_{\mathfrak{s}_M}\,. \tag{8.28}$$

Estimate of P_n. By (5.44), $L_n\mathbf{T}_n - \mathrm{Id} = \mathcal{P}(\tilde{\imath}_n) + \mathcal{P}_\omega(\tilde{\imath}_n) + \mathcal{P}_\omega^\perp(\tilde{\imath}_n)$. Accordingly, we decompose P_n in (8.20) as $P_n = -P_n^{(1)} - P_{n,\omega} - P_{n,\omega}^\perp$, where

$$P_n^{(1)} := \Pi_n\mathcal{P}(\tilde{\imath}_n)\Pi_n\mathcal{F}_\omega(\tilde{U}_n), \qquad P_{n,\omega} := \Pi_n\mathcal{P}_\omega(\tilde{\imath}_n)\Pi_n\mathcal{F}_\omega(\tilde{U}_n), \qquad P_{n,\omega}^\perp := \Pi_n\mathcal{P}_\omega^\perp(\tilde{\imath}_n)\Pi_n\mathcal{F}_\omega(\tilde{U}_n).$$

By (2.3),

$$\begin{aligned}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M+\overline{\sigma}} &\le \|\Pi_n\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M+\overline{\sigma}} + \|\Pi_n^\perp\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M+\overline{\sigma}} \\ &\le K_n^{\overline{\sigma}}(\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M} + K_n^{-\mathtt{b}_1}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M+\mathtt{b}_1}).\end{aligned} \tag{8.29}$$

By (5.45), (8.14), (8.29), and using that (8.21), (8.22), $\gamma^{-1} = K_0 \le K_n$ we obtain

$$\begin{aligned}\|P_n^{(1)}\|_{\mathfrak{s}_M} &\lesssim_{\mathfrak{s}_M+\mathtt{b}_1} \gamma^{-3}K_n^{2\overline{\sigma}}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}(\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M} + K_n^{-\mathtt{b}_1}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M+\mathtt{b}_1}) \\ &\lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{2\overline{\sigma}+3}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}(\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M} + K_n^{\overline{\sigma}-\mathtt{b}_1}(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1})) \\ &\lesssim_{\mathfrak{s}_0+\mathtt{b}_1} K_n^{2\overline{\sigma}+3}\|\mathcal{F}_\omega(\tilde{U}_n)\|^2_{\mathfrak{s}_M} + K_n^{3\overline{\sigma}+3-\mathtt{b}_1}(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1})\,.\end{aligned} \tag{8.30}$$

By (5.46), (8.14), (8.8), (2.3), we have

$$\|P_{n,\omega}\|_{\mathfrak{s}_M} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} \varepsilon\gamma^{-4}N_{n-1}^{-\mathtt{a}}K_n^{\overline{\sigma}}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M} \overset{\gamma^{-1}=K_0\le K_n}{\lesssim_{\mathfrak{s}_0+\mathtt{b}_1}} \varepsilon N_{n-1}^{-\mathtt{a}}K_n^{\overline{\sigma}+4}\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}\,, \tag{8.31}$$

where $\mathtt{a}$ is in (8.2). By (5.47), (2.3), (8.4), (8.11), (8.22) and then using (8.21), $\gamma^{-1} = K_0 \le K_n$, we get

$$\begin{aligned}\|P_{n,\omega}^\perp\|_{\mathfrak{s}_M} &\lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+2\overline{\sigma}-\mathtt{b}_1}\gamma^{-2}(\|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M+\mathtt{b}_1} + \varepsilon\|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}) \\ &\lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+3\overline{\sigma}+2-\mathtt{b}_1}(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}).\end{aligned} \tag{8.32}$$

Estimate of R_n. By the definition (8.12) of L_n one has that for any $\widehat{U} = (\widehat{\imath}, \widehat{\zeta})$, $L_n\widehat{U}$ is given by

$$\begin{aligned}L_n\widehat{U} &= \omega\cdot\partial_\varphi\widehat{\imath} - d_\iota X_{\mathcal{H}_\varepsilon}\big((\varphi,0,0)+\tilde{\imath}_n\big)[\widehat{\imath}] - (0,\widehat{\zeta},0) \\ &\overset{(4.10)}{=} \omega\cdot\partial_\varphi\widehat{\imath} - d_\iota X_{\mathcal{N}}\big((\varphi,0,0)+\tilde{\imath}_n\big)[\widehat{\imath}] - d_\iota X_{\mathcal{P}_\varepsilon}\big((\varphi,0,0)+\tilde{\imath}_n\big)[\widehat{\imath}] - (0,\widehat{\zeta},0)\end{aligned} \tag{8.33}$$

where we recall that $d_\iota X_{\mathcal{N}}\big((\varphi,0,0)+\tilde{\imath}_n\big)[\widehat{\imath}] = \big(\Omega^{kdv}_{\mathbb{S}_+}(\mu)[\widehat{y}],\,0\,,\Omega^{kdv}(\mu,D)[\widehat{w}]\big)$. By the estimate of $d_\iota X_{\mathcal{P}_\varepsilon}$ of Lemma 4.2, one then obtains $\|L_n\widehat{U}\|_{\mathfrak{s}_M} \lesssim \|\widehat{U}\|_{\mathfrak{s}_M+\overline{\sigma}}$. Using (8.20), (8.13), (8.8), (2.3) and then (8.14), (8.21), (8.22), $\gamma^{-1} = K_0 \le K_n$, we get

$$\|R_n\|_{\mathfrak{s}_M} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+3\overline{\sigma}+2-\mathtt{b}_1}(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}). \tag{8.34}$$

Estimate of $\mathcal{F}_\omega(U_{n+1})$. By (8.19), (2.3), (8.21), (8.28), (8.30)-(8.32), (8.34), (8.8), we get (8.23). By (8.16) and (8.13) we now deduce the bound (8.24) for $W_1 := H_1$. Indeed

$$\|W_1\|_{\mathfrak{s}_M+\mathtt{b}_1} = \|H_1\|_{\mathfrak{s}_M+\mathtt{b}_1} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} \gamma^{-2}\|\mathcal{F}_\omega(U_0)\|_{\mathfrak{s}_M+\mathtt{b}_1+\overline{\sigma}} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} \varepsilon\gamma^{-2} \overset{\gamma^{-1}=K_0}{\lesssim} K_0^2\varepsilon\,.$$

Estimate (8.24) for $W_{n+1} := \tilde{W}_n + H_{n+1}$, $n \geq 1$, follows by (8.26). □

By Lemma 8.2 we get the following lemma, where for clarity we write $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$ instead of $\|\cdot\|_s$ as above.

Lemma 8.3. *For any $\omega \in \mathcal{G}_{n+1}$*

$$\|\mathcal{F}_\omega(U_{n+1})\|_{\mathfrak{s}_M}^{\mathrm{Lip}(\gamma)} \leq C_*\varepsilon K_n^{-\mathtt{a}_1}\,, \qquad \|W_{n+1}\|_{\mathfrak{s}_M+\mathtt{b}_1}^{\mathrm{Lip}(\gamma)} \leq C_* K_n^{\mu_1}\varepsilon\,, \tag{8.35}$$

$$\|H_1\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}\,, \quad \|H_{n+1}\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}K_n^{\mu(\mathtt{b})+2\overline{\sigma}}K_{n-1}^{-\mathtt{a}_1}\,, \quad n \geq 1\,. \tag{8.36}$$

Proof. First note that, by (8.10), if $\omega \in \mathcal{G}_{n+1}$, then $\omega \in \mathcal{G}_n$ and so (8.11) and the inequality in $(\mathcal{P}3)_n$ holds. Then the first inequality in (8.35) follows by (8.23), $(\mathcal{P}2)_n$, $(\mathcal{P}3)_n$, $\gamma^{-1} = K_0 \leq K_n$, and by (8.3), (8.4), (8.5)-(8.6). For $n = 0$ we use also (8.7).

The second inequality in (8.35) for $n = 0$ follows directly from the bound for W_1 in (8.24), since $\mu_1 \geq 2$, see (8.4) and $C_* > 0$ large enough (i.e., ε small enough); the second inequality in (8.35) for $n \geq 1$ is proved inductively by taking (8.24), $(\mathcal{P}3)_n$, and the choice of μ_1 in (8.4) into account and by choosing K_0 large enough.

Since $H_1 = W_1$, the first inequality in (8.36) follows since $\|H_1\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}} \lesssim \gamma^{-2}\|\mathcal{F}_\omega(U_0)\|_{\mathfrak{s}_M+\mu(\mathtt{b})+2\overline{\sigma}} \lesssim \varepsilon\gamma^{-2}$. If $n \geq 1$, estimate (8.36) follows by (2.3), (8.27) and (8.11). □

Denote by $\tilde{H}_{n+1}$ a Lip(γ)-extension of $(H_{n+1})_{|\mathcal{G}_{n+1}}$ to the whole set Ω of parameters, provided by the Kirzbraun theorem. Then $\tilde{H}_{n+1}$ satisfies the same bound as H_{n+1} in (8.36) and therefore, by the definition of $\mathtt{a}_2$ in (8.3), the estimate (8.9) holds at $n+1$.

Finally we define the functions

$$\tilde{W}_{n+1} := \tilde{W}_n + \tilde{H}_{n+1}\,, \quad \tilde{U}_{n+1} := \tilde{U}_n + \tilde{H}_{n+1} = U_0 + \tilde{W}_n + \tilde{H}_{n+1} = U_0 + \tilde{W}_{n+1}\,,$$

which are defined for all $\omega \in \Omega$. Note that for any $\omega \in \mathcal{G}_{n+1}$, $\tilde{W}_{n+1} = W_{n+1}$, $\tilde{U}_{n+1} = U_{n+1}$. Therefore $(\mathcal{P}2)_{n+1}$, $(\mathcal{P}3)_{n+1}$ are proved by Lemma 8.3. Moreover by (8.9), which at this point has been proved up to the step $n+1$, we have

$$\|\tilde{W}_{n+1}\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \leq \sum_{k=1}^{n+1}\|\tilde{H}_k\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \leq C_*\varepsilon\gamma^{-2}$$

and thus (8.8) holds also at the step $n+1$. This completes the proof of Theorem 8.1. □

We now deduce Theorem 4.1. Let $\gamma = \varepsilon^{\mathfrak{a}}$ with $\mathfrak{a} \in (0, \mathfrak{a}_0)$ and $\mathfrak{a}_0 := 1/\tau_2$ where τ_2 is defined in (8.7). Then the smallness condition (8.7) holds for $0 < \varepsilon < \varepsilon_0$ small enough and Theorem 8.1 applies. Passing to the limit for $n \to \infty$ we deduce the existence of a function $U_\infty(\omega) = (\breve{\imath}_\infty(\omega), \zeta_\infty(\omega))$, $\omega \in \Omega$, such that $\mathcal{F}_\omega(U_\infty(\omega)) = 0$ for any ω in the set

$$\bigcap_{n\geq 0}\mathcal{G}_n = \mathcal{G}_0 \cap \bigcap_{n\geq 1}\mathbf{\Omega}_{n+1}^{\gamma}(\tilde{\imath}_{n-1}) \overset{(7.57)}{=} \mathcal{G}_0 \cap \Big[\bigcap_{n\geq 1}\Lambda_n^{\gamma}(\tilde{\imath}_{n-1})\Big] \cap \Big[\bigcap_{n\geq 1}\Omega_n^{\gamma}(\tilde{\imath}_{n-1})\Big]\,. \tag{8.37}$$

Moreover

$$\|U_\infty - U_0\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}\,, \quad \|U_\infty - \tilde{U}_n\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}K_n^{-\mathtt{a}_2}\,, \quad n \geq 1\,. \tag{8.38}$$

Formula (5.5) implies that $\zeta_\infty(\omega) = 0$ for ω belonging to the set (8.37), and therefore $\breve{\imath}_\omega := \breve{\imath}_\infty(\omega)$ is an invariant torus for the Hamiltonian vector field $X_{\mathcal{H}_\varepsilon}$ filled by quasi-periodic solutions with frequency ω. It remains only to prove the measure estimate (4.9).

8.2 Measure estimates

Arguing as in [9] one proves the following two lemmata.

Lemma 8.4. *The set*

$$\mathcal{G}_\infty := \mathcal{G}_0 \cap \Big[\bigcap_{n\geq 1} \Lambda_n^{2\gamma}(\iota_\infty)\Big] \cap \Big[\bigcap_{n\geq 1} \Omega_n^{2\gamma}(\iota_\infty)\Big] \tag{8.39}$$

is contained in $\mathcal{G}_n$ *for any* $n \geq 0$*, and hence* $\mathcal{G}_\infty \subseteq \bigcap_{n\geq 0} \mathcal{G}_n$.

For any $j \in \mathbb{S}^\perp$, the sequence $\widetilde{\mu}_j^n : \Omega \to \mathbb{R}$, $n \geq 0$, in Theorem 7.2-$(\mathbf{S2})_n$ is a Cauchy sequence with respect to the norm $|\cdot|^{\mathrm{Lip}(\gamma)}$. We denote the limit by μ_j^∞,

$$\mu_j^\infty := \lim_{n\to\infty} \widetilde{\mu}_j^n(\iota_\infty), \quad j \in \mathbb{S}^\perp. \tag{8.40}$$

By Theorem 7.2 one has for any $j \in \mathbb{S}^\perp$,

$$\mu_{-j}^\infty = -\mu_j^\infty, \qquad |\mu_j^\infty - \widetilde{\mu}_j^n(\iota_\infty)|^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2} N_{n-1}^{-\mathtt{a}}, \ n \geq 0. \tag{8.41}$$

Lemma 8.5. *The set*

$$\begin{aligned}\Omega_\infty^\gamma := \Big\{\omega \in \mathtt{DC}(4\gamma,\tau) : |\omega\cdot\ell + \mu_j^\infty - \mu_{j'}^\infty| \geq \frac{4\gamma|j^3 - j'^3|}{\langle \ell\rangle^\tau}, \ \forall(\ell,j,j') \in \mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp\times\mathbb{S}^\perp, \\ |\omega\cdot\ell + \mu_j^\infty| \geq \frac{4\gamma|j|^3}{\langle\ell\rangle^\tau}, \ \forall(\ell,j) \in \mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp\Big\}\end{aligned} \tag{8.42}$$

is contained in $\mathcal{G}_\infty$, $\Omega_\infty^\gamma \subseteq \mathcal{G}_\infty$*, where* $\mathcal{G}_\infty$ *is defined in* (8.39).

In view of Lemma 8.4 and 8.5, it suffices to estimate the Lebesgue measure $|\Omega \setminus \Omega_\infty^\gamma|$ of $\Omega\setminus\Omega_\infty^\gamma$.

Proposition 8.6. (Measure estimates) *Let* $\tau > |\mathbb{S}_+| + 2$*. Then there is* $\mathfrak{a} \in (0,1)$ *so that for* $\varepsilon\gamma^{-3}$ *sufficiently small, one has* $|\Omega\setminus\Omega_\infty^\gamma| \lesssim \gamma^{\mathfrak{a}}$.

The remaining part of this section is devoted to prove Proposition 8.6. By (8.42), we have

$$\Omega\setminus\Omega_\infty^\gamma = \Omega\setminus\mathtt{DC}(4\gamma,\tau) \ \cup \bigcup_{(\ell,j,j')\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp\times\mathbb{S}^\perp (\ell,j,j')\neq(0,j,j)} \mathcal{R}_{\ell,j,j'} \ \cup \bigcup_{(\ell,j)\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp} \mathcal{Q}_{\ell,j} \tag{8.43}$$

where $\mathcal{R}_{\ell,j,j'}$, $\mathcal{Q}_{\ell,j}$ denote the 'resonant' sets

$$\mathcal{R}_{\ell,j,j'} := \Big\{\omega\in\mathtt{DC}(4\gamma,\tau) \ : \ |\omega\cdot\ell + \mu_j^\infty - \mu_{j'}^\infty| < \frac{4\gamma|j^3 - j'^3|}{\langle\ell\rangle^\tau}\Big\}, \tag{8.44}$$

$$\mathcal{Q}_{\ell,j} := \Big\{\omega\in\mathtt{DC}(4\gamma,\tau) : |\omega\cdot\ell + \mu_j^\infty| < \frac{4\gamma|j|^3}{\langle\ell\rangle^\tau}\Big\}. \tag{8.45}$$

Notice that $\mathcal{R}_{\ell,j,j} = \emptyset$. Furthermore, it is well known that $|\Omega\setminus\mathtt{DC}(4\gamma,\tau)| \lesssim \gamma$. In order to prove Proposition 8.6 we shall use the following asymptotic properties of $\mu_j^\infty(\omega)$. For any ω in $\mathtt{DC}(4\gamma,\tau)$, we have $\widetilde{\mu}_j^0(\iota_\infty) = \mu_j^0(\iota_\infty)$ and we write $\mu_j^\infty(\omega) = \mu_j^0(\iota_\infty) + r_j^\infty(\omega)$, where by (7.2), $m_3^\infty := m_3(\iota_\infty)$, $m_1^\infty := m_1(\iota_\infty)$,

$$\mu_j^0(\iota_\infty) = m_3^\infty(\omega)(2\pi j)^3 - m_1^\infty(\omega)2\pi j - q_j(\omega).$$

On $\mathtt{DC}(4\gamma,\tau)$, the following estimates hold

$$\begin{aligned}&|m_3^\infty + 1|^{\mathrm{Lip}(\gamma)} \overset{(6.18)}{\lesssim} \varepsilon, \qquad |m_1^\infty|^{\mathrm{Lip}(\gamma)} \overset{(6.66)}{\lesssim} \varepsilon\gamma^{-2}, \\ &\sup_{j\in\mathbb{S}^\perp}|j||q_j|^{\sup}, \sup_{j\in\mathbb{S}^\perp}|j||q_j|^{\mathrm{lip}} \overset{(7.4)}{\lesssim} 1, \qquad |r_j^\infty|^{\mathrm{Lip}(\gamma)} \overset{(8.41)}{\lesssim} \varepsilon\gamma^{-2}.\end{aligned} \tag{8.46}$$

From the latter estimates one infers the following standard lemma see [2, Lemma 5.3]).

Lemma 8.7. *(i) If $\mathcal{R}_{\ell,j,j'} \neq \emptyset$, then $|j^3 - j'^3| \leq C\langle \ell \rangle$ for some $C > 0$. In particular one has $j^2 + j'^2 \leq C\langle \ell \rangle$.*
(ii) If $\mathcal{Q}_{\ell,j} \neq \emptyset$, then $|j|^3 \leq C\langle \ell \rangle$ for some $C > 0$.

Lemma 8.7 can be used to estimate $|\mathcal{R}_{\ell,j,j'}|$ and $|\mathcal{Q}_{\ell,j}|$ for $|\ell|$ sufficiently large.

Lemma 8.8. *(i) If $\mathcal{R}_{\ell,j,j'} \neq \emptyset$, then there exists $C_1 > 0$ with the following property: if $|\ell| \geq C_1$, then $|\mathcal{R}_{\ell,j,j'}| \lesssim \gamma |j^3 - j'^3| \langle \ell \rangle^{-(\tau+1)}$.*
(ii) If $\mathcal{Q}_{\ell,j} \neq \emptyset$, then there exists $C_1 > 0$ with the following property: if $|\ell| \geq C_1$, then $|\mathcal{Q}_{\ell,j}| \lesssim \gamma |j|^3 \langle \ell \rangle^{-(\tau+1)}$.

Proof. We only prove item (i) since item (ii) can be proved in a similar way. Assume that $\mathcal{R}_{\ell,j,j'} \neq \emptyset$. Let $\bar{\omega}$ such that $\bar{\omega} \cdot \ell = 0$ and introduce the real valued function $s \mapsto \phi_{\ell,j,k}(s)$,

$$\phi_{\ell,j,j'}(s) := f_{\ell,j,j'}\big(\bar{\omega} + s\frac{\ell}{|\ell|}\big), \qquad f_{\ell,j,j'}(\omega) := \omega \cdot \ell + \mu_j^\infty(\omega) - \mu_{j'}^\infty(\omega).$$

Using that by Lemma 8.7, $|j^3 - j'^3| \leq C\langle \ell \rangle$, one infers from (8.46) that, for $\varepsilon\gamma^{-2}$ small enough and $|\ell| \geq C_1$ with C_1 large enough, $|\phi_{\ell,j,j'}(s_2) - \phi_{\ell,j,j'}(s_1)| \geq \frac{|\ell|}{2}|s_2 - s_1|$. Since $\mathtt{DC}(4\gamma, \tau)$ is bounded one sees by standard arguments that

$$\big|\big\{s \in \mathbb{R} \,:\, \bar{\omega} + s\frac{\ell}{|\ell|} \in \mathcal{R}_{\ell,j,j'}\big\}\big| \lesssim \gamma |j^3 - j'^3| \langle \ell \rangle^{-(\tau+1)}.$$

The claimed estimate then follows by applying Fubini's theorem. $\square$

It remains to estimate the Lebesgue measure of the resonant sets $\mathcal{R}_{\ell,j,j'}$ and $\mathcal{Q}_{\ell,j}$ for $|\ell| \leq C_1$.

Lemma 8.9. *Assume that $|\ell| \leq C_1$ and that $\varepsilon\gamma^{-3}$ is small enough. Then the following holds:*
(i) If $\mathcal{R}_{\ell,j,j'} \neq \emptyset$, then there are constants $\mathfrak{a} \in (0,1)$ and $C_2 > 0$ so that $|j|, |j'| \leq C_2$ and $|\mathcal{R}_{\ell,j,j'}| \lesssim \gamma^{\mathfrak{a}}$.
(ii) If $\mathcal{Q}_{\ell,j} \neq \emptyset$ then there are constants $\mathfrak{a} \in (0,1)$ and $C_2 > 0$ so that $|j| \leq C_2$ and $|\mathcal{Q}_{\ell,j}| \lesssim \gamma^{\mathfrak{a}}$.

Proof. We only prove item (i) since item (ii) can be proved in a similar way. If $|\ell| \leq C_1$ and $\mathcal{R}_{\ell,j,j'} \neq \emptyset$, Lemma 8.7-$(i)$ implies that there is a constant C_2 such that $|j|, |j'| \leq j^2 + j'^2 \leq C_2$. For $\varepsilon\gamma^{-3}$ small enough one sees, using (8.46), the definition (7.2) of μ_j^0, and the bounds $|\ell| \leq C_1, |j|, |j'| \leq C_2$, that $|\mu_j^\infty - \omega_j^{kdv}| \lesssim \varepsilon\gamma^{-2} \lesssim \gamma$, implying that for some constant $C_3 > 0$,

$$\mathcal{R}_{\ell,j,j'} \subset \big\{\omega \in \Omega \;:\; |\omega \cdot \ell + \omega_j^{kdv}(\nu(\omega), 0) - \omega_{j'}^{kdv}(\nu(\omega), 0)| \leq C_3\gamma\big\}. \tag{8.47}$$

By Lemma 3.8, the function $\omega \mapsto \omega \cdot \ell + \omega_j^{kdv}(\nu(\omega), 0) - \omega_{j'}^{kdv}(\nu(\omega), 0)$ is real analytic and not identically zero. Hence by the Weierstrass preparation theorem (cf. the proof of [8, Lemma 9.7]), we deduce that the measure of the set on the right hand side of (8.47) is smaller than $\gamma^{\mathfrak{a}}$ for some $\mathfrak{a} \in (0,1)$ and γ small enough. $\square$

By (8.43) and Lemmata 8.8–8.9 we deduce that

$$|\Omega \setminus \Omega_\infty^\gamma| \lesssim \gamma^{\mathfrak{a}} + \gamma \sum_{|\ell| \geq C_1, |j|, |j'| \leq C\langle \ell \rangle} \langle \ell \rangle^{-\tau} \lesssim \gamma^{\mathfrak{a}},$$

where we used the assumption that $\tau - 2 > |\mathbb{S}_+|$. This concludes the proof of Proposition 8.6.

References

[1] P. Baldi, M. Berti, E. Haus, R. Montalto, *Time quasi-periodic gravity water waves in finite depth*, Inventiones Math. 214 (2), 739–911, 2018.

[2] P. Baldi, M. Berti, R. Montalto, *KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation*, Math. Annalen 359, 471–536, 2014.

[3] P. Baldi, M. Berti, R. Montalto, *KAM for autonomous quasi-linear perturbations of KdV*, Ann. Inst. H. Poincaré Analyse Non. Lin. 33, no. 6, 1589–1638, 2016.

[4] P. Baldi, M. Berti, R. Montalto, *KAM for autonomous quasi-linear perturbations of mKdV*, Bollettino Unione Matematica Italiana, 9, 143–188, 2016.

[5] M. Berti, L. Biasco, M. Procesi, *KAM theory for the Hamiltonian DNLW*, Ann. Sci. Éc. Norm. Supér. (4), Vol. 46, fascicule 2, 301–373, 2013.

[6] M. Berti, L. Biasco, M. Procesi, *KAM theory for the reversible derivative wave equation*, Arch. Rational Mech. Anal., 212, 905–955, 2014.

[7] M. Berti, P. Bolle, *A Nash-Moser approach to KAM theory*, Fields Institute Communications, special volume "Hamiltonian PDEs and Applications", 255–284, 2015.

[8] M. Berti, T. Kappeler, R. Montalto, *Large KAM tori for perturbations of the defocusing NLS equation*, Astérisque 403, viii+148 pp., 2018.

[9] M. Berti, R. Montalto, *Quasi-periodic standing wave solutions of gravity capillary standing water waves*, to appear in Memoirs of the Amer. Math. Society, MEMO 891, arXiv:1602.02411, 2016.

[10] R.. Bikbaev, S. Kuksin, *On the parametrisation of finite gap solutions by frequency vector and wave number vector and a theorem by I. Krichever*, Lett. Math. Phys. 28, 115-122, 1993.

[11] R. Feola, M. Procesi, *Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations*, J. Diff. Eq., 259, no. 7, 3389–3447, 2015.

[12] F. Giuliani, *Quasi-periodic solutions for quasi-linear generalized KdV equations*, J. Differential Equations 262, 10, 5052–5132, 2017.

[13] J. Liu, X. Yuan, *A KAM Theorem for Hamiltonian Partial Differential Equations with Unbounded Perturbations*, Comm. Math. Phys, 307 (3), 629–673, 2011.

[14] T. Kappeler, R. Montalto, *Canonical coordinates with tame estimates for the defocusing NLS equation on the circle*, Int. Math. Res. Not. IMRN, 5, 1473–1531, 2018.

[15] T. Kappeler, R. Montalto, *Normal form coordinates for the KdV equation having expansions in terms of pseudo-differential operators*, Comm. Math Phys. DOI: 10.1007/s00220-019-03498-1, 2019.

[16] T. Kappeler, J. Pöschel, *KdV & KAM*, Springer-Verlag, 2003.

[17] T. Kappeler, B. Schaad, P. Topalov, *Qualitative features of periodic solutions of KdV*, Comm. Part. Diff. Eqs. 38, (9), 1626–1673, 2013.

[18] T. Kappeler, B. Schaad, P. Topalov, *Semi-linearity of the nonlinear Fourier transform of the defocusing NLS equation*, Int. Math. Res. Notices, 7212 –7229, 2016.

[19] S. Kuksin, *A KAM theorem for equations of the Korteweg-de Vries type*, Rev. Math. Phys., 10, 3, 1–64, 1998.

[20] S. Kuksin, *Analysis of Hamiltonian PDEs*, Oxford University Press, 2000.

[21] J. Zhang, M. Gao, X. Yuan, *KAM tori for reversible partial differential equations*, Nonlinearity 24, 1189–1228, 2011.