Arch. Rational Mech. Anal. 239 (2021) 1395–1500
Digital Object Identifier (DOI) https://doi.org/10.1007/s00205-020-01596-2

Large KAM Tori for Quasi-linear Perturbations of KdV

Massimiliano Berti, Thomas Kappeler & Riccardo Montalto

Communicated by N. Masmoudi

Abstract

In this paper we prove the persistence of space periodic multi-solitons of arbitrary size under any quasi-linear Hamiltonian perturbation, which is smooth and sufficiently small. This answers positively a longstanding question of whether KAM techniques can be further developed to prove the existence of quasi-periodic solutions of arbitrary size of strongly nonlinear perturbations of integrable PDEs.

Contents

1. Introduction

The Korteweg–de Vries (KdV) equation

$$\partial_t u = -\partial_x^3 u + 6u\partial_x u \tag{1.1}$$

is one of the most important model equations for dispersive phenomena with numerous applications in physics. The seminal discovery in the late sixties that (1.1) admits infinitely many conservation laws [28,30], and the development of the inverse scattering transform method [18], led to the modern theory of infinite dimensional integrable systems (for example [13,16] and references therein).

One of the most distinguished features of (1.1) is the existence of sharply localized traveling waves of arbitrarily large amplitudes and particle like properties. Kruskal and Zabusky, who discovered them in numerical experiments in the early sixties, both on the real line and in the periodic setup (cf. [24]), coined the name solitons for them. More generally, they found solutions, which are localized near finitely many points in space. In the periodic setup, these solutions are referred to as *periodic multi-solitons* or *finite gap* solutions. Due to their importance in applications, various stability aspects, in particular long time asymptotics, have been extensively studied. A major question concerns the persistence of the multi-solitons under perturbations. In the last thirty years, KAM methods pioneered by Kolmogorov, Arnold, and Moser to treat perturbations of integrable systems of finite dimension, were developed for PDEs. Most of the work focused on small amplitude solutions or semilinear perturbations. It has been a longstanding question from experts in PDEs and in infinite dimensional dynamical systems whether KAM results hold also for solutions of arbitrary size under quasi-linear perturbations, called strongly nonlinear in [26], of integrable PDEs.

The aim of this paper is to prove the first persistence result of periodic multi-solitons of KdV of *arbitrary* size under *strongly nonlinear* perturbations—see Theorem 1.1 below. Note that in this case, it was not even known if there exist solutions of the perturbed equation which are global in time.

To describe the class of perturbations of the KdV equation considered, we recall that (1.1), with space periodic variable x in $\mathbb{T}_1 := \mathbb{R}/\mathbb{Z}$, can be written in Hamiltonian form

$$\partial_t u = \partial_x \nabla H^{kdv}(u), \quad H^{kdv}(u) := \int_{\mathbb{T}_1} \frac{1}{2}(\partial_x u)^2(x) + u^3(x)\, \mathrm{d}x, \tag{1.2}$$

where ∇H^{kdv} denotes the L^2-gradient of H^{kdv} and ∂_x is the Poisson structure, corresponding to the Poisson bracket, defined for functionals F, G by $\{F, G\} := \int_{\mathbb{T}_1} \nabla F\, \partial_x \nabla G\, \mathrm{d}x$.

We consider *quasi-linear* perturbations of (1.1) of the form

$$\partial_t u = -\partial_x^3 u + 6u\partial_x u + \varepsilon a(x, u(x), \partial_x u(x))\partial_x^3 u + \cdots, \tag{1.3}$$

where $0 < \varepsilon < 1$ is a small parameter and $\cdots$ comprises terms which are ε-small and contain x-derivatives of u up to second order. We assume that the perturbation is *Hamiltonian*, namely $\varepsilon a\partial_x^3 u + \cdots = \varepsilon\partial_x \nabla P$, where ∇P is the L^2-gradient of a functional of the form

$$P(u) := \int_{\mathbb{T}_1} f(x, u(x), \partial_x u(x))\, \mathrm{d}x, \quad f : \mathbb{T}_1 \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \ \ \mathcal{C}^\infty\text{-smooth}. \tag{1.4}$$

Note that the nonlinear vector field

$$\partial_x \nabla P(u) = -(\partial_{u_x}^2 f)(x, u(x), \partial_x u(x))\partial_x^3 u + \cdots, \quad u_x := \partial_x u \tag{1.5}$$

has the same order as the linear one $\partial_x^3 u$ in (1.1). When written as a Hamiltonian PDE, (1.3) takes the form

$$\partial_t u = \partial_x \nabla H_\varepsilon(u) \tag{1.6}$$

with Hamiltonian

$$H_\varepsilon(u) := H^{kdv}(u) + \varepsilon P(u). \tag{1.7}$$

To state our main result, we first need to introduce some more notation. Note that the mean $u \mapsto \int_{\mathbb{T}_1} u(x)\, \mathrm{d}x$ is a prime integral for (1.6). We restrict our attention to functions with zero average (cf. Remark (R1) below) and choose as phase spaces for (1.6) the scale of Sobolev spaces $H_0^s(\mathbb{T}_1)$, $s \geq 0$,

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

where

$$H^s(\mathbb{T}_1) := \Big\{u(x) = \sum_{n\in\mathbb{Z}} u_n e^{\mathrm{i}2\pi nx} \ : \ \|u\|_{H_x^s} := \Big(\sum_{n\in\mathbb{Z}} \langle n\rangle^{2s} |u_n|^2\Big)^{\frac{1}{2}} < \infty, \ u_{-n} = \overline{u_n} \ \ \forall n \in \mathbb{Z}\Big\} \tag{1.8}$$

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 $L_0^2(\mathbb{T}_1)$ is given by

$$\mathcal{W}_{L_0^2}(u, v) := \int_{\mathbb{T}_1} (\partial_x^{-1} u) v\, \mathrm{d}x, \quad \partial_x^{-1} u = \sum_{n\neq 0} \frac{1}{\mathrm{i}n} u_n e^{\mathrm{i}2\pi nx}, \quad \forall u, v \in L_0^2(\mathbb{T}_1) \tag{1.9}$$

Note that the Hamiltonian vector field $X_H(u) = \partial_x \nabla H(u)$, associated with a Hamiltonian H, is determined by $\mathrm{d}H(u)[\cdot] = \mathcal{W}_{L_0^2}(X_H, \cdot)$.

$\mathbb{S}_+$-*gap potentials* According to [21], the KdV equation (1.1) on $\mathbb{T}_1$ is an integrable PDE in the strongest possible sense, meaning that it admits globally defined canonical coordinates on $H_0^0(\mathbb{T}_1)$, so that (1.1) can be solved by quadrature, see Theorem 3.1 in Section 3 for a precise statement. These coordinates, referred to as Birkhoff coordinates, are particularly suited to describe the finite gap solutions of

KdV. Each of these solutions is contained in a finite dimensional integrable subsystem $\mathcal{M}_{\mathbb{S}_+}$, of dimension $2|\mathbb{S}_+|$, with $\mathbb{S}_+$ being a finite subset of $\mathbb{N}_+ := \{1, 2, \dots\}$. The integrable subsystem $\mathcal{M}_{\mathbb{S}_+}$ can be described in terms of action angle coordinates $\theta := (\theta_n)_{n\in\mathbb{S}_+}$, $I := (I_n)_{n\in\mathbb{S}_+}$ as follows: there exists a real analytic, canonical diffeomorphism

$$\Psi_{\mathbb{S}_+} : \mathbb{T}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+}_{>0} \to \mathcal{M}_{\mathbb{S}_+}, (\theta, I) \mapsto q(\theta, \cdot; I), \quad \mathbb{T}^{\mathbb{S}_+} := (\mathbb{R}/2\pi\mathbb{Z})^{\mathbb{S}_+}, \tag{1.10}$$

(cf. (3.5)) so that the pull-back of the KdV Hamiltonian, $H^{kdv} \circ \Psi_{\mathbb{S}_+}$, is a real analytic function of the actions I alone. Elements in $\mathcal{M}_{\mathbb{S}_+}$ are referred to as $\mathbb{S}_+$-gap potentials. The function $q(\theta, x) \equiv q(\theta, x; I)$ is real analytic. In action angle coordinates, any solution of (1.1) on $\mathcal{M}_{\mathbb{S}_+}$ is given by

$$\theta(t) = \theta^{(0)} - \omega^{kdv}(\nu)t, \quad I(t) = \nu,$$

where $\theta^{(0)} \in \mathbb{T}^{\mathbb{S}_+}$ denotes the initial angles, $\nu \in \mathbb{R}^{\mathbb{S}_+}_{>0}$ the initial actions, and $\omega^{kdv}(\nu)$ the frequency vector

$$\omega^{kdv}(\nu) := \partial_I (H^{kdv} \circ \Psi_{\mathbb{S}_+})(0, \nu) \in \mathbb{R}^{\mathbb{S}_+}. \tag{1.11}$$

(Cf. Section 3.1 for more details.) The corresponding finite-gap solution of (1.1) on $\mathcal{M}_{\mathbb{S}_+}$ is then given by

$$t \mapsto q\big(\theta^{(0)} - \omega^{kdv}(\nu)t,\ x;\ \nu\big) \tag{1.12}$$

and hence is quasi-periodic in time. The map

$$\mathbb{R}^{\mathbb{S}_+}_{>0} \to \mathbb{R}^{\mathbb{S}_+},\ \nu \mapsto \omega^{kdv}(\nu), \tag{1.13}$$

is a local diffeomorphism (see Remark 3.10). In the entire paper, $\Xi \subset \mathbb{R}^{\mathbb{S}_+}_{>0}$ will always denote an open, nonempty set with the property that the map $\Xi \to \mathbb{R}^{\mathbb{S}_+}$, $\nu \mapsto \omega^{kdv}(\nu)$, defined by (1.11), is a diffeomorphism onto its image and that its closure is a compact subset of $\mathbb{R}^{\mathbb{S}_+}_{>0}$. Then, for some $\delta > 0$ small enough,

$$\Xi + B_{\mathbb{S}_+}(\delta) \subseteq \mathbb{R}^{\mathbb{S}_+}_{>0}, \tag{1.14}$$

where $B_{\mathbb{S}_+}(\delta)$ denotes the ball in $\mathbb{R}^{\mathbb{S}_+}$ of radius δ, centered at the origin. Furthermore we introduce the Sobolev spaces of periodic, real valued functions $H^s \equiv H^s(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1, \mathbb{R})$,

$$\begin{aligned} H^s := \Big\{ g = \sum_{(\ell,j)\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{Z}} g_{\ell,j}\, e^{\mathrm{i}(\ell\cdot\varphi + 2\pi j x)} \ :\ \overline{g_{\ell,j}} = g_{-(\ell,j)},\ \forall\,(\ell, j) \in \mathbb{Z}^{\mathbb{S}_+} \times \mathbb{Z}, \\ \|g\|_s := \Big(\sum_{(\ell,j)\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{Z}} |g_{\ell,j}|^2 \langle \ell, j\rangle^{2s} \Big)^{1/2} < +\infty, \quad \langle \ell, j\rangle := \max\{1, |\ell|, |j|\} \Big\}. \end{aligned} \tag{1.15}$$

Note that by the Sobolev embedding theorem, $H^s \subset \mathcal{C}^0(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1, \mathbb{R})$ for any $s > (|\mathbb{S}_+| + 1)/2$ where $\mathcal{C}^0(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1, \mathbb{R})$ denotes the Banach space of continuous functions endowed with the supremum norm.

The main result of this paper is Theorem 1.1 below. It says that for ε small enough and for ν in a subset Ξ_ε of Ξ of asymptotically full Lebesgue measure, there is a quasi-periodic solution of equation (1.6) close to the finite gap solution $q(-\omega^{kdv}(\nu)t, x; \nu)$ of (1.1). More precisely, the following holds:

Theorem 1.1. *Let f be a function in $\mathcal{C}^\infty(\mathbb{T}_1 \times \mathbb{R} \times \mathbb{R}, \mathbb{R})$, $\mathbb{S}_+$ a finite subset of $\mathbb{N}_+$, and $\mathfrak{b}$ a real number in $(0, 1)$. Then there exist $\bar{s} > (|\mathbb{S}_+| + 1)/2$, and $0 < \varepsilon_0 < 1$ so that the following holds: there exists a decreasing family of measurable subsets $\Xi_\varepsilon \subseteq \Xi$, $0 < \varepsilon \le \varepsilon_0$, with asymptotically full measure, that is*

$$\lim_{\varepsilon \to 0} |\Xi \setminus \Xi_\varepsilon| = 0, \tag{1.16}$$

with the property that for any $\nu \in \Xi_\varepsilon$, the perturbed KdV equation (1.6) admits a quasi-periodic solution $t \mapsto u_\varepsilon(\omega_\varepsilon(\nu)t, x; \nu)$ with frequency vector $\omega_\varepsilon(\nu) = -\omega^{kdv}(\nu) \in \mathbb{R}^{\mathbb{S}_+}$, where $u_\varepsilon(\cdot, \cdot; \nu) \in H^{\bar{s}}(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1, \mathbb{R})$ and

$$\|u_\varepsilon(\cdot, \cdot; \nu) - q(\cdot, \cdot; \nu)\|_{\bar{s}} \lesssim \varepsilon^{1-\mathfrak{b}}. \tag{1.17}$$

Here, $q(\theta, x; \nu)$, $\theta \in \mathbb{T}^{\mathbb{S}_+}$, is the $\mathbb{S}_+$-gap potential in $\mathcal{M}_{\mathbb{S}_+}$, defined in (1.10). The quasi-periodic solution $t \mapsto u_\varepsilon(\omega_\varepsilon(\nu)t, x; \nu)$ is linearly stable.

We make the following remarks:

(R0) Since the Hamiltonian vector field in (1.6) is autonomous, any translate $u_\varepsilon(\omega_\varepsilon(\nu)t + \theta^{(0)}, x; \nu)$, $\theta^{(0)} \in \mathbb{T}^{\mathbb{S}_+}$, of $u_\varepsilon(\omega_\varepsilon(\nu)t, x; \nu)$ is also a solution of the perturbed KdV equation (1.6).

(R1) The result of Theorem 1.1 holds for any density f in $\mathcal{C}^{s_*}(\mathbb{T}_1 \times \mathbb{R} \times \mathbb{R}, \mathbb{R})$ with s_* large enough and for any family of $\mathbb{S}_+$-gap solutions of KdV with average c (cf. [21, page 112]). We assume in this paper that f is $\mathcal{C}^\infty$ and $c = 0$ merely to simplify the exposition.

(R2) The methods developed to prove Theorem 1.1 are quite general. We expect that analogous results can also be proved for equations in the KdV hierarchy as well as for the defocusing NLS and equations in the NLS hierarchy such as the defocusing mKdV equation.

Theorem 1.1 is proved at the end of the paper in Section 8.3. It is deduced from Theorem 4.1 (Section 4), which is proved by applying a Nash–Moser iteration scheme (Section 8.1) and by establishing the measure estimates of Section 8.2. Before describing the main ideas of the proof in detail, we first comment on the novelty of our result.

1. The first KAM results for (1.1) were proved by Kuksin [25] (cf. also [26]) and Kappeler–Pöschel [21] for finite gap solutions of *arbitrary* size for *semilinear* perturbations of the KdV equation. This means that the density f of (1.4) does *not* depend on u_x, and hence $\partial_x \nabla P(u) = \partial_u^2 f(x, u(x)) u_x + \cdots$ depends only on u and u_x. (Note that in addition, the dependence on u_x is linear.) Subsequently, Liu–Yuan [29] proved KAM results for semilinear perturbations of small amplitude solutions of the derivative NLS and the Benjamin–Ono equations whereas Zhang–Gao–Yuan [32] proved analogous results for the reversible derivative

NLS equation. More recently, Berti–Biasco–Procesi [6]–[7] proved existence of small quasi-periodic solutions of derivative Klein-Gordon equations. For the NLS and the beam equations in higher space dimension, KAM results were obtained by Eliasson–Kuksin [15] and, respectively, Eliasson-Grébert–Kuksin [14]. In all of these works, the perturbations are required to be semilinear.
On the other hand, the results in Baldi–Berti–Montalto [3,4], for *quasi-linear* perturbations of the KdV and mKdV equations concern only *small amplitude* solutions. The proofs of these results make use of pseudo-differential calculus and rely in a decisive manner on the differential nature of these equations. The latter property cannot be read off in the action-angle coordinates outside a neighborhood of the origin. Furthermore, also the results of Giuliani [19] for the generalized KdV equation, the ones of Feola–Procesi [17] for the NLS equation, and the ones of Berti–Montalto [10] and Baldi–Berti–Haus–Montalto [1] for water waves concern small amplitude solutions.
Thus the challenging problem of the persistence of the finite gap solutions of (1.1) of *arbitrary* size under *strongly nonlinear* perturbations (1.5) remained completely open.

2. In [9], we used the "one-smoothing property" of the Birkhoff coordinates of the defocusing NLS equation on $\mathbb{T}_1$, established in [23], to prove a KAM result for *semilinear* perturbations. This property is used to deal with the difficulties related to the double "asymptotic multiplicity" of the frequencies. For the KdV equation, a "one-smoothing property" has been proved near the equilibrium in [27] and then in general in [22], however it is not sufficient for dealing with the quasi-linear perturbations (1.5).
3. The proof of Theorem 1.1 uses the canonical coordinates constructed in [20] near any given compact family of $\mathbb{S}_+$-gap potentials in $\mathcal{M}_{\mathbb{S}_+}$, reviewed in Section 3.2. These coordinates admit an expansion in terms of pseudo-differential operators up to a remainder of arbitrary negative order. Due to its length, this part of the proof of Theorem 1.1 has been published in a separate paper [20]. In Section 3.3 we show that the linearization of the Hamiltonian vector field X_{H_ε}, when expressed in these coordinates, admits an expansion in terms of pseudo-differential operators. This property is one of the key ingredients for implementing the Nash–Moser iteration scheme as explained in the subsequent paragraph.

Ideas of the Proof Theorem 1.1 is proved by means of a Nash–Moser iterative scheme to construct, for any ν belonging to a suitable subset Ξ_ε of Ξ, a quasi-periodic solution of (1.6) with frequency vector $\omega = -\omega^{kdv}(\nu)$ near the $\mathbb{S}_+$-gap solution $t \mapsto q\big(-\omega^{kdv}(\nu)t,\ x;\ \nu\big)$ of the KdV equation (1.1) (cf. (1.12)), which evolves on the torus $\mathcal{T}_\nu := \Psi_{\mathbb{S}_+}(\mathbb{T}^{\mathbb{S}_+} \times \{\nu\})$ (cf. (1.10), (3.5)). The subset Ξ_ε is obtained by imposing along the iterative scheme suitable non-resonance conditions. In particular we will always assume ω to be diophantine, meaning that there exist positive constants $0 < \gamma < 1$ and $\tau > |\mathbb{S}_+| - 1$ so that $|\omega \cdot \ell| \geqq \gamma |\ell|^{-\tau}$ for any $\ell \in \mathbb{Z}^{\mathbb{S}_+} \setminus \{0\}$.

The starting point of our proof is to express the perturbed KdV equation (1.6) in the canonical coordinates $(\theta, y, w) \in \mathbb{T}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times L^2_\perp(\mathbb{T}_1)$, constructed in [20], in a neighborhood of a torus $\mathcal{T}_\nu$. Here

$$L^2_\perp(\mathbb{T}_1) := \Big\{ w = \sum_{n \in \mathbb{S}^\perp} w_n e^{\mathrm{i}2\pi n x} \in L^2_0(\mathbb{T}_1) \Big\}, \quad \mathbb{S}^\perp := \mathbb{Z} \setminus \big(\mathbb{S}_+ \cup (-\mathbb{S}_+) \cup \{0\}\big). \qquad (1.18)$$

To be more precise, denote by $B_\perp(\delta)$ the open ball in $L^2_\perp(\mathbb{T}_1)$, centered at 0, of radius δ, and by $B_{\mathbb{S}_+}(\delta)$ the one in $\mathbb{R}^{\mathbb{S}_+}$ (cf. (1.14)). According to [20], for $\nu \in \Xi$ and $0 < \delta < 1$ sufficiently small, there exists a canonical coordinate chart

$$\Psi_\nu : \mathcal{V}(\delta) \to L^2_0(\mathbb{T}_1), \;\; (\theta, y, w) \mapsto \Psi_\nu(\theta, y, w), \quad \mathcal{V}(\delta) := \mathbb{T}^{\mathbb{S}_+} \times B_{\mathbb{S}_+}(\delta) \times B_\perp(\delta),$$

so that the following key properties hold (cf. Theorem 3.2):

(P1) $\Psi_\nu(\theta, y, 0) = \Psi_{\mathbb{S}_+}(\theta, \nu + y)$ for any $(\theta, y) \in \mathbb{T}^{\mathbb{S}_+} \times B_{\mathbb{S}_+}(\delta)$ (with $\Psi_{\mathbb{S}_+}$ as in (1.10), (3.5));

(P2) $\Psi_\nu(\theta, y, w) \in H^s_0(\mathbb{T}_1)$ for any $(\theta, y, w) \in \mathbb{T}^{\mathbb{S}_+} \times B_{\mathbb{S}_+}(\delta) \times \big(B_\perp(\delta) \cap H^s(\mathbb{T}_1)\big)$ and $s \in \mathbb{N}$;

(P3) equation (1.6) takes the form $(\dot\theta, \dot y, \dot w) = X_{\mathcal{H}_\varepsilon}$ where the Hamiltonian vector field $X_{\mathcal{H}_\varepsilon}$ is given by

$$X_{\mathcal{H}_\varepsilon} = (-\nabla_y \mathcal{H}_\varepsilon, \nabla_\theta \mathcal{H}_\varepsilon, \partial_x \nabla_w \mathcal{H}_\varepsilon), \quad \mathcal{H}_\varepsilon := H_\varepsilon \circ \Psi_\nu;$$

for $\varepsilon = 0$, the manifold $\{w = 0\}$ is invariant for the constant vector field $X_{\mathcal{H}_0} = (-\omega^{kdv}(\nu), 0, 0)$;

(P4) Ψ_ν admits an expansion in terms of pseudo-differential operators, up to regularizing operators satisfying tame estimates, as stated in Theorem 3.2-(AE1) (note that in the estimates Theorem 3.2-(Est1), the dependence with respect to the highest Sobolev norm is linear);

(P5) the linearization of $(\dot\theta, \dot y, \dot w) = X_{\mathcal{H}_0}$ along the manifold $\{y = 0, w = 0\}$ is in diagonal form with coefficients depending only on ν; more specifically $\partial_t \widehat\theta = -\Omega^{kdv}_{\mathbb{S}_+}(\nu)\widehat y$, $\partial_t \widehat y = 0$, $\partial_t \widehat w = \partial_x \Omega^{kdv}(D; \nu)\widehat w$, (cf. the normal form Hamiltonian (3.12)).

As a consequence of $(P1)$–$(P3)$, for $\varepsilon = 0$, the curve $t \mapsto (-\omega^{kdv}(\nu)t, 0, 0)$ is a solution of $(\dot\theta, \dot y, \dot w) = X_{\mathcal{H}_0}$, evolving on the torus $\mathbb{T}^{\mathbb{S}_+} \times \{0\} \times \{0\}$, which is invariant under the flow of $X_{\mathcal{H}_0}$. We look for a quasi-periodic solution of $(\dot\theta, \dot y, \dot w) = X_{\mathcal{H}_\varepsilon}$ near the torus $\mathbb{T}^{\mathbb{S}_+} \times \{0\} \times \{0\}$, with frequency vector $\omega = -\omega^{kdv}(\nu)$, of the form $\breve\iota(\omega t)$ where

$$\breve\iota : \mathbb{R}^{\mathbb{S}_+} \to \mathbb{R}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times H^s_\perp(\mathbb{T}_1), \quad H^s_\perp(\mathbb{T}_1) := H^s(\mathbb{T}_1) \cap L^2_\perp(\mathbb{T}_1),$$

with s sufficiently large, is the lift

$$\breve\iota(\varphi) = (\varphi, 0, 0) + \iota(\varphi), \quad \iota(\varphi) = (\theta(\varphi) - \varphi, y(\varphi), w(\varphi))$$

of a torus embedding and ι is $(2\pi\mathbb{Z})^{\mathbb{S}_+}$-periodic. Thus the unknown function ι satisfies (cf. (4.6))

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

and the map $t \mapsto \Psi_\nu(\breve\iota(\omega t)) \in H^s_0(\mathbb{T}_1)$ is a quasi-periodic solution of (1.6). The equation $\mathcal{F}_\omega(\iota) = 0$ is solved by a Nash–Moser iteration scheme. The core of this scheme is the construction of an approximate right inverse of the linearized operator $d\mathcal{F}_\omega$ at an embedding $\breve\iota(\varphi) = (\theta(\varphi), y(\varphi), w(\varphi))$, near $\breve\iota_0(\varphi) = (\varphi, 0, 0)$, and the proof that it satisfies suitable tame estimates, cf. Theorem 5.7. One of the main issues is to construct an approximate inverse of the linear operator, acting on $L^2_\perp(\mathbb{T}_1)$,

$$\mathcal{L}^{(0)}_\omega = \omega \cdot \partial_\varphi - \partial_x d_\perp \nabla_w \mathcal{H}_\varepsilon(\breve\iota(\varphi)) \,,$$

where $d_\perp$ denotes the differential with respect to w. We achieve this goal by reducing $\mathcal{L}^{(0)}_\omega$ to a linear diagonal operator with constant coefficients. Using properties $(P4)$ and $(P5)$ we prove that

(P6) the linearized Hamiltonian operator $\partial_x d_\perp \nabla_w \mathcal{H}_\varepsilon$ in a neighborhood of a $\mathbb{S}_+$-gap potential is close to $\partial_x \Omega^{kdv}(D; \nu)$ (acting in $H^s_\perp(\mathbb{T}_1)$) and it admits an expansion in terms of classical pseudo-differential operators, up to smoothing remainders which satisfy tame estimates in $H^s_\perp(\mathbb{T}_1)$ – see Lemmata 3.5 and 3.7 in Section 3.3.

Property $(P6)$ allows us to use pseudo-differential techniques, developed in [1,3, 10], to reduce $\mathcal{L}^{(0)}_\omega$ to a diagonal one with constant coefficients up to smoothing remainders. Actually, using $(P6)$ we prove that the operator $\mathcal{L}^{(0)}_\omega$ has the form (cf. Lemma 6.2)

$$\mathcal{L}^{(0)}_\omega = \omega \cdot \partial_\varphi - \Pi_\perp\Big(a^{(0)}_3 \partial_x^3 + 2(a^{(0)}_3)_x \partial_x^2 + a^{(0)}_1 \partial_x + \sum_{k=0}^{M} a^{(0)}_{-k} \partial_x^{-k} + Q^{kdv}_{-1}(D; \omega)\Big) + \mathcal{R}^{(0)}_M, \tag{1.19}$$

where $\Pi_\perp$ is the L^2-orthogonal projector onto the subspace $L^2_\perp(\mathbb{T}_1)$, the coefficients $a^{(0)}_{-k}(\varphi, x)$, $k = -3, \ldots, M$, are real valued functions, $a^{(0)}_3 \sim -1$, and $\mathcal{R}^{(0)}_M$ is a φ-dependent regularizing operator which satisfies tame estimates in the Sobolev spaces $H^s(\mathbb{T}^{\mathbb{S}_+}_\varphi \times \mathbb{T}_1)$. The order M of regularization has to be sufficiently large to ensure the convergence of the KAM iterative reducibility scheme, carried out in Section 7; M is fixed in (7.6) and depends on the cardinality $|\mathbb{S}_+|$ of $\mathbb{S}_+$ and on the diophantine exponent τ of the frequency vector ω. We point out that the term $Q^{kdv}_{-1}(D; \omega)$ in (1.19) is not ε-small, since the finite gap solutions considered might have large amplitudes. More precisely, $Q^{kdv}_{-1}(D; \omega)$ is the Fourier multiplier, acting on $L^2_\perp(\mathbb{T}_1)$, with symbol $\omega^{kdv}_n - (2\pi n)^3$ (cf. (3.62)), which takes into account the difference between the KdV-frequencies and the frequencies $(2\pi n)^3$, $n \in \mathbb{Z}$, of the Airy equation $\partial_t v = -\partial_x^3 v$. We also mention that the pseudo-differential operator $\sum_{k=0}^{M} a^{(0)}_{-k} \partial_x^{-k}$ is not present in [3], since in the latter paper only small amplitude finite gap solutions are considered.

In order to show that the regularizing operator $\mathcal{R}^{(0)}_M$ in (1.19) is tame (which is a key property for the convergence of a Nash–Moser iterative scheme), we prove in Section 3.2 novel results of independent interest concerning the extensions of the differential of the canonical coordinates of [20] to Sobolev spaces $H^s(\mathbb{T}_1)$, of negative order $s < 0$ (cf. Corollaries 3.3 and 3.4).

The special form (1.19) allows to find preliminary transformations which diagonalize $\mathcal{L}_\omega^{(0)}$ up to a pseudo-differential operator of order zero plus a regularizing remainder (see Section 6). More precisely we conjugate $\mathcal{L}_\omega^{(0)}$ to the Hamiltonian operator (cf. (6.69))

$$\mathcal{L}_\omega^{(4)} = \omega \cdot \partial_\varphi - \big(m_3 \partial_x^3 + m_1 \partial_x + \mathrm{Op}(r_0^{(4)}) + Q_{-1}^{kdv}(D;\omega)\big) + \mathcal{R}_M^{(4)}, \quad (1.20)$$

where $m_3 + 1$ and m_1 are real constants, which are ε-small, $\mathrm{Op}(r_0^{(4)})$ is a pseudo-differential operator of order 0 and $\mathcal{R}_M^{(4)}$ is a regularizing operator satisfying tame estimates. The map which conjugates $\mathcal{L}_\omega^{(0)}$ to $\mathcal{L}_\omega^{(4)}$ is obtained by the composition of the transformations introduced in Sections 6.2–6.5. These transformations, inspired by [3], are Fourier integral operators given by symplectic flows of linear Hamiltonian transport PDEs or pseudo-differential maps. In particular, we point out that in order to conjugate the pseudo-differential terms $a_{-k}^{(0)} \partial_x^{-k}$ under the transport flow used in Section 6.3, we need a quantitative version of the Egorov theorem, which is stated and proved in Section 2.5. We remark that in contrast to [3], we implement in Section 6.2 the time-quasi-periodic reparametrization *before* the conjugation with the transport flow to avoid a technical difficulty in the conjugation of the remainders obtained in the Egorov theorem. Furthermore, we mention that related transformations have been developed in [5] for proving upper bounds for the growth of Sobolev norms for certain classes of PDEs.

At this point, using properties of the KdV frequencies that are recorded in Section 3.4, we are able to perform in Section 7 a KAM reducibility scheme to complete the diagonalization of the operator $\mathcal{L}_\omega^{(4)}$ in (1.20) for most values of ν. Since the variable coefficients term $-\mathrm{Op}(r_0^{(4)}) + \mathcal{R}_M^{(4)}$ in (1.20) (which is renamed $\mathtt{R}_0$ in (7.3)) is proven to satisfy the tame estimates of Lemma 7.1, such a KAM reducibility scheme can be implemented along the lines developed in Berti-Montalto [10]. See Theorem 7.3 for details.

Finally in Section 8 we implement a standard Nash–Moser iterative scheme to construct a solution of $\mathcal{F}_\omega(\iota) = 0$ for all frequency vectors ω satisfying "Melnikov nonresonance conditions", cf. (8.37). By the results of Section 8.2, using properties of the KdV frequencies, we prove that the set of such non-resonant frequencies has asymptotically full measure as $\varepsilon \to 0$.

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

$$\|f\|_{H_\varphi^s X} := \|f\|_{L_\varphi^2 X} + \max_{|\beta|=s} \|\partial_\varphi^\beta f\|_{L_\varphi^2 X}.$$

In case $s = 0$, we often write $L_\varphi^2 X$ instead of $H_\varphi^0 X$. Occasionally, we denote the Sobolev space $H^s(\mathbb{T}_1) \equiv H^s(\mathbb{T}_1, \mathbb{R})$ in (1.8) by H_x^s and write L_x^2 for H_x^0. The space L^2_x is endowed with the standard L^2-inner product $\big(f, g\big)_{L^2_x}$ given by

$$\big(f, g\big)_{L^2_x} := \int_{\mathbb{T}_1} f(x)g(x)\,\mathrm{d}x. \tag{1.21}$$

Note that the Sobolev space $H^s \equiv H^s(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1, \mathbb{R})$ defined in (1.15) is an algebra for the product of functions if $s \geq s_0$ where throughout the paper, s_0 is defined as

$$s_0 := \left[\frac{|\mathbb{S}_+| + 1}{2}\right] + 1, \tag{1.22}$$

where [ ] denotes the integer part. For any $s \geqq 0$, denote by h^s_0 the sequence space

$$\begin{aligned} h^s_0 &:= \big\{z = (z_n)_{n\in\mathbb{Z}} \in h^s : z_0 = 0\big\}, \\ h^s &:= \big\{z = (z_n)_{n\in\mathbb{Z}},\ z_n \in \mathbb{C} : \|z\|_s^2 < \infty,\ \overline{z_n} = z_{-n},\ \forall n \in \mathbb{Z}\big\}, \end{aligned} \tag{1.23}$$

where $\|z\|_s^2 := \sum_{n\in\mathbb{Z}} \langle n\rangle^{2s}|z_n|^2 < \infty$. 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 $u_n := \int_{\mathbb{T}_1} u(x)e^{-\mathrm{i}2\pi nx}\,\mathrm{d}x$ for any $n \in \mathbb{Z}$ and by $\mathcal{F}^{-1} : h^0 \to L^2(\mathbb{T}_1)$ its inverse. Furthermore, we denote by $\Pi_0^\perp$ the L^2-orthogonal projector onto the subspace of functions with zero average $L^2_0(\mathbb{T}_1)$. We set

$$\begin{aligned} &H^s_\perp(\mathbb{T}_1) := H^s(\mathbb{T}_1) \cap L^2_\perp(\mathbb{T}_1), \\ &H^s_\perp \equiv H^s_\perp(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1) := \big\{u \in H^s(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1) : u(\varphi, \cdot) \in L^2_\perp(\mathbb{T}_1)\big\} \end{aligned} \tag{1.24}$$

where $L^2_\perp(\mathbb{T}_1)$ is defined in (1.18). Often we write $L^2_\perp$ for $H^0_\perp$. The space $H^0_\perp(\mathbb{T}_1)$ is also denoted by $L^2_\perp(\mathbb{T}_1)$. By $\Pi_\perp$ we denote the L^2-orthogonal projector onto $L^2_\perp(\mathbb{T}_1)$, $\Pi_\perp : L^2(\mathbb{T}_1) \to L^2_\perp(\mathbb{T}_1)$. Let

$$\begin{aligned} \mathcal{E}_s &:= \mathbb{T}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times H^s_\perp(\mathbb{T}_1), \quad \mathcal{E} \equiv \mathcal{E}_0, \\ E_s &:= \mathbb{R}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times H^s_\perp(\mathbb{T}_1), \quad E \equiv E_0. \end{aligned} \tag{1.25}$$

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 (1.25). Note 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

$$\big\langle(\widehat{\theta}_1, \widehat{y}_1, \widehat{w}_1), (\widehat{\theta}_2, \widehat{y}_2, \widehat{w}_2)\big\rangle := \widehat{\theta}_1 \cdot \widehat{\theta}_2 + \widehat{y}_1 \cdot \widehat{y}_2 + \big(\widehat{w}_1, \widehat{w}_2\big)_{L^2_x}. \tag{1.26}$$

By a slight abuse of notation, $\Pi_\perp$ also denotes the projector of E_s onto its third component, $\Pi_\perp : E_s \to H^s_\perp(\mathbb{T}_1)$, $(\widehat{\theta}, \widehat{y}, \widehat{w}) \mapsto \widehat{w}$. Furthermore, we denote by $d_\perp$ the differential with respect to w of any map, defined on an open set of $\mathcal{E}_s$, taking values in some Banach space.

For any $0 < \delta < 1$, we denote $B_{\mathbb{S}_+}(\delta)$ 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 $B_\perp^0(\delta)$. These balls are used to define the following open neighborhoods in $\mathcal{E}_s$, $s \in \mathbb{N}$,

$$\mathcal{V}^s(\delta) := \mathbb{T}_1^{\mathbb{S}_+} \times B_{\mathbb{S}_+}(\delta) \times B_\perp^s(\delta), \quad \mathcal{V}(\delta) \equiv \mathcal{V}^0(\delta), \quad 0 < \delta < 1 . \tag{1.27}$$

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, $[A, B] := AB - BA$ and for a linear operator A, acting on an Hilbert space H, by $A^\top$ the transpose of A with respect to the scalar product of H. In case A is invertible, the transpose of the inverse A^{-1} of A is denoted by $A^{-\top}$.
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 $\Delta_\omega f : \Omega \times \Omega \to X$, $(\omega_1, \omega_2) \mapsto f(\omega_1) - f(\omega_2)$.

2. Preliminaries

The goal of this section is to record analytical tools used throughout the paper. In Section 2.1 we introduce function spaces of functions of the variables φ, x, depending on a parameter ω in a Lipschitz continuous way, and state their main properties. In addition, we introduce the classes of φ-dependent linear operators used in the paper, and the subclasses of Hamiltonian and of symplectic ones. In Section 2.2 we review the notion of periodic pseudo-differential operators and basic elements of their calculus. They are a key tool in Section 6, for the reduction of the linearized operators obtained along the Nash–Moser iteration, to operators with constant coefficients, up to smoothing remainders. In Section 2.3 we discuss the notion of tame and modulo-tame operators, introduced in [10] as a technical tool in order to facilitate the derivations of tame estimates, mainly needed to setup the KAM reducibility scheme in Section 7. The results of Sections 2.4 and 2.5 are new. In Section 2.4 we prove tame estimates for compositions of functions and operators with a torus embedding $\breve{\imath} : \mathbb{T}^{\mathbb{S}_+} \to \mathbb{T}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times H_\perp^s(\mathbb{T}_1)$, acting in spaces of functions of the variables $(\varphi, x) \in \mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1$. These tame estimates are at the heart of the convergence of the KAM reducibility scheme, proved in Section 7 and of the Nash–Moser iteration, proved in Section 8.1. In Section 2.5, we prove a version of the Egorov theorem with quantitative tame estimates needed in the reduction step of Section 6.3.

2.1. Function Spaces and Linear Operators

In the paper we consider real or complex functions $u(\varphi, x; \omega)$, $(\varphi, x) \in \mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1$, which are Lipschitz continuous with respect to the parameter $\omega \in \Omega$, where Ω is a subset of $\mathbb{R}^{\mathbb{S}_+}$. In the sequel, we will often suppress ω in $u(\varphi, x; \omega)$ to make notation lighter. Given $0 < \gamma < 1$ and $s \geq 0$, we define the norm

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

where $\| \|_s$ is the norm of the Sobolev space H^s, defined in (1.15), and $u(\omega_1) = u(\cdot, \cdot; \omega_1)$. 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 by Π_N the *smoothing* operators on H^s,

$$(\Pi_N u)(\varphi, x) := \sum_{|(\ell, j)| \leqq N} u_{\ell, j} e^{\mathrm{i}(\ell \cdot \varphi + 2\pi j x)},$$
$$u_{\ell, j} = \frac{1}{(2\pi)^{|\mathbb{S}_+|}} \int_{\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1} u(\varphi, x) e^{-\mathrm{i}(\ell \cdot \varphi + 2\pi j x)} d\varphi \mathrm{d}x, \tag{2.2}$$

and let $\Pi_N^\perp := \mathrm{Id} - \Pi_N$. For any $\alpha \geqq 0$ and $s \in \mathbb{R}$, the operators Π_N and $\Pi_N^\perp$ satisfy the standard estimates

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

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

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

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

Lemma 2.1. (Product and composition)

(i) *For any $s \geqq s_0 = [(|\mathbb{S}_+| + 1)/2] + 1$ (cf. (1.22)),*

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

(ii) *Let $\beta(\cdot, \cdot; \omega) : \mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1 \to \mathbb{R}$ with $\|\beta\|_{2s_0+2}^{\mathrm{Lip}(\gamma)} \leqq \delta(s_0)$ 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 $s \geqq s_0 + 1$,*

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

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

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

(iii) *Let $\alpha(\cdot; \omega) : \mathbb{T}^{\mathbb{S}_+} \to \mathbb{R}$ with $\|\alpha\|_{2s_0+2}^{\mathrm{Lip}(\gamma)} \leqq \delta(s_0)$ 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 $s \geqq s_0 + 1$,*

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

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

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

Remark 2.2. Note that for $s > s_0$, the bound $C(s)\|u\|_s^{\mathrm{Lip}(\gamma)}\|v\|_{s_0}^{\mathrm{Lip}(\gamma)} + C(s_0)\|u\|_{s_0}^{\mathrm{Lip}(\gamma)}\|v\|_s^{\mathrm{Lip}(\gamma)}$ of $\|uv\|_s^{\mathrm{Lip}(\gamma)}$ in (2.5) is *linear* in $\|u\|_s^{\mathrm{Lip}(\gamma)}$ and $\|v\|_s^{\mathrm{Lip}(\gamma)}$. It is this property of the estimate which is referred to as "tame".

Proof. Item (i) follows from (2.72) in [10] and (ii)–(iii) follow from [10, Lemma 2.30]. □

If the vector $\omega \in \mathbb{R}^{\mathbb{S}_+}$ is diophantine, that is $|\omega\cdot\ell| \geqq \gamma/|\ell|^\tau$ for any $\ell \in \mathbb{Z}^{\mathbb{S}_+}\setminus\{0\}$, the equation $\omega\cdot\partial_\varphi v = u$ with $u(\varphi, x)$ satisfying $u_{0,j} = 0$ for any $j \in \mathbb{Z}$, has the periodic solution

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

and it satisfies the standard estimate (cf. [9, Lemma 2.2])

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

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 $H^s(\mathbb{T}^d)$.

Lemma 2.3. (Composition operator, [10, Lemma 2.31]) *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 $\|v\|_{s_0(d)}^{\mathrm{Lip}(\gamma)} \leqq 1$ where $s_0(d) > d/2$, then, for any $s \geqq s_0(d)$,*

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

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

Next we discuss classes of linear operators used in this paper. 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 let A act on functions $u(\varphi, x)$ of space-time. In this way we get an element in $\mathcal{L}(L^2(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1, \mathbb{C}))$, again denoted by A, which is defined by

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

We say that the linear operator A is *real* if it maps real valued functions to 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 2.4. 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

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

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 $\mathrm{i}(\ell_m-\ell'_m)A_j^{j'}(\ell-\ell')$.

Hamiltonian and symplectic operators will play an important role in the reduction procedure of linearized operators, implemented in Sections 6 and 7. They are defined as follows.

Definition 2.5. (*Hamiltonian and symplectic operators*)

(i) A φ-dependent family of linear operators $X(\varphi)$, $\varphi\in\mathbb{T}^{\mathbb{S}_+}$, densely defined in $L_0^2(\mathbb{T}_1)$, 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. By a slight abuse of terminology, $\omega\cdot\partial_\varphi - \partial_x G(\varphi)$ is also said to be a Hamiltonian operator.

(ii) A φ-dependent family of linear operators $A(\varphi): L_0^2(\mathbb{T}_1)\to L_0^2(\mathbb{T}_1)$, $\forall\varphi\in\mathbb{T}^{\mathbb{S}_+}$, is symplectic if

$$\mathcal{W}_{L_0^2}(A(\varphi)u, A(\varphi)v) = \mathcal{W}_{L_0^2}(u,v), \quad \forall u,v\in L_0^2(\mathbb{T}_1),$$

where the symplectic 2-form $\mathcal{W}_{L_0^2}$ is the one defined in (1.9).

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 one which is again Hamiltonian. Self-adjoint operators and real ones are characterized in terms of their matrix elements as follows:

Lemma 2.6. *A family of linear operators $R(\varphi)$, $\varphi\in\mathbb{T}^{\mathbb{S}_+}$ with Fourier series $R(\varphi)=\sum_{\ell\in\mathbb{Z}^{\mathbb{S}_+}} R(\ell)e^{\mathrm{i}\ell\cdot\varphi}$, is*

(i) self-adjoint *if and only if* $\overline{R_j^{j'}(\ell)} = R_{j'}^{j}(-\ell)$, $\forall j,j'\in\mathbb{Z}$, $\ell,\ell'\in\mathbb{Z}^{\mathbb{S}_+}$;

(ii) real *if and only if* $\overline{R_{j'}^{j}(\ell)} = R_{-j'}^{-j}(-\ell)$, $\forall j,j'\in\mathbb{Z}$, $\ell,\ell'\in\mathbb{Z}^{\mathbb{S}_+}$;

(iii) Real and self-adjoint *if and only if* $R_j^{j'}(\ell) = R_{-j'}^{-j}(\ell)$, $\forall j,j'\in\mathbb{Z}$, $\ell,\ell'\in\mathbb{Z}^{\mathbb{S}_+}$.

The next lemma describes the structure of specific linear Hamiltonian operators (cf. Definition 2.5), and will be used in Lemmata 6.2 and 6.5.

Lemma 2.7. *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*

$$X = a_3(x)\partial_x^3 + a_2(x)\partial_x^2 + a_1(x)\partial_x + \text{bounded operator} \tag{2.16}$$

where $a_3, a_2, a_1 \in \mathcal{C}^\infty(\mathbb{T}_1, \mathbb{R})$. *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 $L_0^2(\mathbb{T}_1)$ satisfying $\mathcal{A} = \mathcal{A}^\top$. Since by (2.16), $\mathcal{A} = \partial_x^{-1} X = a_3(x)\partial_{xx} + \big(-(a_3)_x + a_2\big)\partial_x + \ldots$ and $\mathcal{A}^\top = -X^\top \partial_x^{-1} = a_3(x)\partial_{xx} + \big(3(a_3)_x - a_2\big)\partial_x + \ldots$, the identity $\mathcal{A} = \mathcal{A}^\top$ implies that $a_2 = 2(a_3)_x$. □

2.2. Pseudo-differential Operators

In this section we introduce the class of pseudo-differential operators, acting on functions on $\mathbb{T}_1$, which are used in this paper, and discuss their basic calculus, following [10]. (Note however that in [10], the space variable x is in $\mathbb{R}/(2\pi\mathbb{Z})$ whereas in this paper it is in $\mathbb{T}_1$.)

Definition 2.8. (*Pseudo-differential operators, symbols*) 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 $\alpha, \beta \in \mathbb{N}$,

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

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 $u(x) = \sum_{j\in\mathbb{Z}} u_j e^{\mathrm{i}2\pi jx}$ to

$$A[u](x) \equiv (Au)(x) := \sum\nolimits_{j\in\mathbb{Z}} a(x, j) u_j e^{\mathrm{i}2\pi jx}.$$

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 $a(x)$, that is, $A : u(x) \mapsto a(x)u(x)$ and we also write a for A. If a is independent of x, the operator $A = \mathrm{Op}(a)$ is referred to as Fourier multiplier. In particular, $\langle D \rangle$ denotes the Fourier multiplier with symbol $\langle \xi \rangle := \max\{1, |\xi|\}$

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 $\omega \in \Omega$, 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 $A(\varphi)$, $\mathrm{Op}(a(\varphi, \cdot))$, or $a(\varphi, x, D; \omega)$.

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

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

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.

Following [10, Definition 2.11], we introduce for any $s \geq 0$ the norm of a symbol $a(\varphi, x, \xi; \omega)$ in S^m, which controls the regularity in (φ, x) and the decay in ξ of a and its derivatives $\partial_\xi^\beta a \in S^{m-\beta}$, $0 \leqq \beta \leqq \alpha$, 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 as in [10], we consider the difference quotient instead of the derivative with respect to ω, and write $| \; |_{m,s,\alpha}^{\mathrm{Lip}(\gamma)}$ instead of $| \; |_{m,s,\alpha}^{1,\gamma}$.

Definition 2.9. (*Norm of pseudo-differential operators*) Let $A(\omega) := a(\varphi, x, D; \omega) \in OPS^m$ be a family of pseudo-differential operators with symbols $a(\varphi, x, \xi; \omega) \in S^m$ of order $m \in \mathbb{R}$. For $\gamma \in (0,1)$, $\alpha \in \mathbb{N}$, $s \geqq 0$, 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 $|A(\omega)|_{m,s,\alpha} := \max_{0 \leqq \beta \leqq \alpha} \sup_{\xi \in \mathbb{R}} \|\partial_\xi^\beta a(\cdot, \cdot, \xi; \omega)\|_s \langle \xi \rangle^{-m+\beta}$.

The pseudo-differential norm $|\cdot|_{m,s,\alpha}^{\mathrm{Lip}(\gamma)}$ satisfies the following elementary properties: for any $s \leqq s'$, $\alpha \leqq \alpha'$, and $m \leqq m'$,

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

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

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

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 $N \geqq 0$, 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 [10, Lemma 2.13]:

Lemma 2.10. (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* $A \circ B$ *is the pseudo-differential operator of order* $m + m'$, *associated to the symbol* $\sigma_{AB}(\varphi, x, \xi; \omega)$ *which satisfies, for any* $\alpha \in \mathbb{N}$, $s \geqq s_0$,

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

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

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

By (2.23) the commutator $[A, B]$ of $A = a(x, D) \in OPS^m$ and $B = b(x, D) \in OPS^{m'}$ is a pseudo-differential operator of order $m + m' - 1$, and Lemma 2.10 yields (cf. [10, Lemma 2.15]).

Lemma 2.11. (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* $s \geqq s_0$ *satisfies*

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

In the case of operators of the special form $a\partial_x^m$, Lemmas 2.10 and 2.11 simplify as follows:

Lemma 2.12. (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* $K_{n,m} \in \mathbb{R}$, $0 \le n \le N$, *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* $s \geqq s_0$, $\alpha \in \mathbb{N}$,

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

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

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

Proof. The results follow from the asymptotic expansion formula (2.23) and Lemma 2.10. □

Finally we give the following result on the exponential of a pseudo-differential operator of order 0.

Lemma 2.13. (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. *Furthermore, for any* $s \geqq s_0$, $\alpha \in \mathbb{N}$, *there is a constant* $C(s, \alpha) > 0$ *so that*

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

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

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

Therefore

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

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. □

2.3. Lip(γ)-*Tame and Modulo-Tame Operators*

In this section we review the notions and the main properties of Lip(γ)-σ-tame and Lip(γ)-modulo-tame operators, introduced in [10, Section 2.2]. (Again, unlike [10], we consider difference quotients instead of first order derivatives with respect to ω.)

Definition 2.14. (Lip(γ)-σ-tame) Let $\sigma \geq 0$ and $0 < \gamma < 1$. A linear operator $A = A(\omega)$ as in (2.12) is said to be Lip(γ)-σ-tame if there exist numbers s_1, S with $s_0 \leq s_1 < S$ and a non-decreasing function $[s_1, S] \to [0, +\infty)$, $s \mapsto \mathfrak{M}_A(s)$, so that, for any $s_1 \leqq s \leqq S$ and $u \in H^{s+\sigma}$,

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

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 σ, referred to as loss of derivatives. We will not explicitly record this dependence.

Remark 2.15. In the sequel, often we will not explicitly record the domain of definition $[s_1, S]$ of the Lip(γ)-σ-tame constant $\mathfrak{M}_A(s)$ in order to make the statements lighter. Similarly, we will always assume that $0 < \gamma < 1$, without stating it explicitly.

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 any $j' \in \mathbb{Z}$, $\ell' \in \mathbb{Z}^{\mathbb{S}_+}$ and any $\omega_1, \omega_2 \in \Omega$, $\omega_1 \neq \omega_2$,

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

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

Lemma 2.16. (Composition, [10, Lemma 2.20]) *Let A, B be a* Lip(γ)-σ_A*-tame and respectively, a* Lip(γ)-σ_B*-tame operator 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*

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

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 2.17. (Action on H^s, [10, Lemma 2.22]) *Let $A := A(\omega)$ be a* Lip(γ)-σ-*tame operator with tame constant $\mathfrak{M}_A(s)$. Then for any $s \in [s_1, S]$, for any family of Sobolev functions $u := u(\omega) \in H^{s+\sigma}$, Lipschitz continuous with respect to ω, the following tame estimates hold:*

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

Pseudo-differential operators are tame operators. We will use, in particular, the following lemma:

Lemma 2.18. *Let* $a(\varphi, x, \xi; \omega) \in S^0$ *be a family of symbols that are Lipschitz continuous with respect to* ω. *If* $A = a(\varphi, x, D; \omega)$ *satisfies* $|A|_{0,s,0}^{\mathrm{Lip}(\gamma)} < +\infty$ *for any* $s \geqq s_0$, *then* A *is* Lip(γ)*-tame with a tame constant satisfying*

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

As a consequence, for any $s \geq s_0$,

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

Proof. See [10, Lemma 2.21] for the proof of (2.31). The estimate (2.32) then follows from Lemma 2.17. □

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.

Definition 2.19. (Lip(γ)*-modulo-tame*) A linear operator $A := A(\omega)$ as in (2.12) is Lip(γ)-modulo-tame if there exist numbers s_1, S with Let $s_0 \leq s_1 < S$ and a non-decreasing function $[s_1, S] \to [0, +\infty)$, $s \mapsto \mathfrak{M}_A^\sharp(s)$, so that for any $s_1 \leqq s \leqq S$ and $u \in H^s$,

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

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

Similarly as mentioned in Remark 2.15, the domain of definition of $\mathfrak{M}_A^\sharp(s)$ will often not be explicitly recorded. By Definition 2.19, if B is a Lip(γ)-modulo-tame operator and A is a linear operator satisfying $|A_j^{j'}(\ell)| \leqq |B_j^{j'}(\ell)|$, then A is Lip(γ)-modulo-tame with a modulo-tame constant $\mathfrak{M}_A^\sharp(s)$ satisfying $\mathfrak{M}_A^\sharp(s) \leqq \mathfrak{M}_B^\sharp(s)$. Moreover, by comparing Definitions 2.19 and 2.14 (for $\sigma = 0$) one deduces the following lemma (cf. [10, Lemma 2.24] for details):

Lemma 2.20. *An operator* A *which 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.

The class of Lip(γ)-modulo-tame operators (Definition 2.19) is closed under the operations coming up in the KAM reduction procedure, namely: sum and composition (Lemma 2.21); projections (Lemma 2.23); solution of the homological equation (Lemma 7.5). Let us give the precise statement of the first property.

Lemma 2.21. (Sum and composition [10, Lemma 2.25]) *Let* A, B *be* Lip(γ)*-modulo-tame operators with modulo-tame constants* $\mathfrak{M}_A^\sharp(s)$ *and, respectively ,* $\mathfrak{M}_B^\sharp(s)$. *Then* $A+B$ *is* Lip(γ)*-modulo-tame with a modulo-tame constant* $\mathfrak{M}_{A+B}^\sharp(s)$ *satisfying*

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

The composed operator $A \circ B$ *is* Lip(γ)*-modulo-tame with a modulo-tame constant satisfying, for some* $C \geq 1$,

$$\mathfrak{M}^{\sharp}_{AB}(s) \leqq C \left(\mathfrak{M}^{\sharp}_{A}(s)\mathfrak{M}^{\sharp}_{B}(s_1) + \mathfrak{M}^{\sharp}_{A}(s_1)\mathfrak{M}^{\sharp}_{B}(s)\right). \tag{2.35}$$

Assume in addition that $\langle \partial_\varphi \rangle^{\mathtt{b}} A$, $\langle \partial_\varphi \rangle^{\mathtt{b}} B$ *(see Definition 2.4) are* Lip(γ)*-modulo-tame with modulo-tame constants* $\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} A}(s)$ *and, respectively,* $\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} B}(s)$*. Then* $\langle \partial_\varphi \rangle^{\mathtt{b}}(AB)$ *is* Lip(γ)*-modulo-tame with a modulo-tame constant* $\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}}(AB)}(s)$*, bounded for some* $C(\mathtt{b}) \geqq 1$ *by*

$$C(\mathtt{b})\Big(\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} A}(s)\mathfrak{M}^{\sharp}_{B}(s_1) + \mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} A}(s_1)\mathfrak{M}^{\sharp}_{B}(s) + \mathfrak{M}^{\sharp}_{A}(s)\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} B}(s_1) + \mathfrak{M}^{\sharp}_{A}(s_1)\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} B}(s)\Big). \tag{2.36}$$

Iterating the tame estimates (2.35) and (2.36) for the composition of operators we get that, for any $n \geqq 2$,

$$\mathfrak{M}^{\sharp}_{A^n}(s) \leqq \left(2C\mathfrak{M}^{\sharp}_{A}(s_1)\right)^{n-1}\mathfrak{M}^{\sharp}_{A}(s), \tag{2.37}$$

and

$$\begin{aligned}\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} A^n}(s) \leqq (4C(\mathtt{b})C)^{n-1}\Big(&\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} A}(s)\big[\mathfrak{M}^{\sharp}_{A}(s_1)\big]^{n-1} \\ &+\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} A}(s_1)\mathfrak{M}^{\sharp}_{A}(s)\big[\mathfrak{M}^{\sharp}_{A}(s_1)\big]^{n-2}\Big).\end{aligned} \tag{2.38}$$

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

Lemma 2.22. (Exponential map) *Let* A *and* $\langle \partial_\varphi \rangle^{\mathtt{b}} A$ *be* Lip(γ)*-modulo-tame operators and assume that* $\mathfrak{M}^{\sharp}_{A} : [s_1, S] \to [0, +\infty)$ *is a modulo-tame constant satisfying*

$$\mathfrak{M}^{\sharp}_{A}(s_1) \leqq 1. \tag{2.39}$$

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

$$\begin{aligned}\mathfrak{M}^{\sharp}_{\Phi^{\pm 1} - \mathrm{Id}}(s) &\lesssim \mathfrak{M}^{\sharp}_{A}(s), \\ \mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}}(\Phi^{\pm 1} - \mathrm{Id})}(s) &\lesssim_{\mathtt{b}} \mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} A}(s) + \mathfrak{M}^{\sharp}_{A}(s)\mathfrak{M}^{\sharp}_{\langle \partial_\varphi \rangle^{\mathtt{b}} A}(s_1).\end{aligned} \tag{2.40}$$

Proof. In view of the identity $\Phi^{\pm 1} - \mathrm{Id} = \sum_{n \geqq 1} \frac{(\pm A)^n}{n!}$ and the assumption (2.39) the claimed estimates follow by (2.37)–(2.38). □

Along the KAM reducibility scheme of Section 7.1 we need the following estimates for the operator $\Pi_N^{\perp} A := A - \Pi_N A$ where $\Pi_N A$ is the smoothed operator, defined in (2.15):

Lemma 2.23. (Smoothing, [10, Lemma 2.27]) *Suppose that* $\langle \partial_\varphi \rangle^{\mathtt{b}} A$, $\mathtt{b} \geqq 0$, *is* Lip(γ)*-modulo-tame. Then the operator* $\Pi_N^\perp A$ *(cf. Definition 2.4) is* Lip(γ)*-modulo-tame with a modulo-tame constant satisfying*

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

We will also encounter linear operators of the form $h \mapsto (a_2, h)_{L^2_x} a_1$ where a_1, a_2 are smooth functions. According to the next lemma, such operators are modulo-tame regularizing. Recall that $\langle D \rangle$ denotes the Fourier multiplier with symbol $\langle \xi \rangle$.

Lemma 2.24. *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* $\mathcal{R} : L^2_x \to L^2_x$, $h \mapsto (a_2, h)_{L^2_x} a_1$. *Then for any* $\lambda \in \mathbb{N}^{\mathbb{S}_+}$ *and* $n_1, n_2 \geqq 0$, *the operator* $\langle D \rangle^{n_1} \partial_\varphi^\lambda \mathcal{R} \langle D \rangle^{n_2}$ *is* Lip(γ)*-tame with a tame constant satisfying, for some* $\sigma \equiv \sigma(n_1, n_2, \lambda) > 0$ *and, for any* $s \geq s_0$,

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

Proof. For any $n_1, n_2 \geqq 0$, $\lambda \in \mathbb{N}^{\mathbb{S}_+}$, $h \in L^2_x$, one has that, for some combinatorial constants c_{λ_1, λ_2},

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

where we used that the operator $\langle D \rangle$ is self-adjoint. The lemma then follows by (2.5). □

2.4. Tame Estimates

In this section we record tame estimates for compositions of functions and operators with the lift $\breve{\iota}$ of a torus embedding $\mathbb{T}^{\mathbb{S}_+} \to \mathcal{E}_s$ (cf. (1.25)) of the form

$$\breve{\iota}(\varphi) = (\varphi, 0, 0) + \iota(\varphi) \quad \text{where} \quad \iota(\varphi) = (\Theta(\varphi), y(\varphi), w(\varphi)) \ \text{ is } \ (2\pi\mathbb{Z})^{\mathbb{S}_+}\text{-periodic},$$

endowed with the norm $\|\iota\|_s^{\mathrm{Lip}(\gamma)} := \|\Theta\|_{H^s_\varphi}^{\mathrm{Lip}(\gamma)} + \|y\|_{H^s_\varphi}^{\mathrm{Lip}(\gamma)} + \|w\|_s^{\mathrm{Lip}(\gamma)}$. The main results are Lemmas 2.25 and 2.26, whose relevance is described in Remark 2.27.

Note that the norm $\|\cdot\|_s$ of the Sobolev space $H^s = H^s(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1)$, introduced 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 that by interpolation estimates (which are proved using Young's inequality), one has

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

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

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

If X is a Banach space and H a Hilbert space, the following Sobolev embedding results hold: for any $s_1 \in \mathbb{N}$,

$$\begin{aligned} &H^{s+s_1}(\mathbb{T}^{\mathbb{S}_+}, X) \hookrightarrow \mathcal{C}^{s_1}(\mathbb{T}^{\mathbb{S}_+}, X), \quad \forall s > |\mathbb{S}_+|, \\ &H^{s+s_1}(\mathbb{T}^{\mathbb{S}_+}, H) \hookrightarrow \mathcal{C}^{s_1}(\mathbb{T}^{\mathbb{S}_+}, H), \quad \forall s > |\mathbb{S}_+|/2. \end{aligned} \tag{2.45}$$

For the convenience of the reader, let us prove the two cases for $s_1 = 0$. To see that $H^s(\mathbb{T}^{\mathbb{S}_+}, X) \hookrightarrow \mathcal{C}^0(\mathbb{T}^{\mathbb{S}_+}, X)$ for $s > |\mathbb{S}_+|$, consider for any given $u \in H^s(\mathbb{T}^{\mathbb{S}_+}, X)$ its Fourier expansion $u(\varphi) = \sum_{\ell \in \mathbb{Z}^{\mathbb{S}_+}} u_\ell e^{\mathrm{i}\ell \cdot \varphi}$, $u_\ell \in X$. Integrating by parts, one has

$$\| u_\ell \|_X = \Big\| \frac{1}{(2\pi)^{|\mathbb{S}_+|}} \int_{\mathbb{T}^{\mathbb{S}_+}} u(\varphi) e^{-\mathrm{i}\ell \cdot \varphi} \, \mathrm{d}\varphi \Big\|_X \lesssim_s \langle \ell \rangle^{-s} \| u \|_{H^s_\varphi X}, \quad \forall \, \ell \in \mathbb{Z}^{\mathbb{S}_+}.$$

The claimed embedding then follows, since $s > |\mathbb{S}_+|$:

$$\| u \|_{C^0_\varphi X} \leqq \sum_{\ell \in \mathbb{Z}^{\mathbb{S}_+}} \| u_\ell \|_X \lesssim_s \| u \|_{H^s_\varphi X} \sum_{\ell \in \mathbb{Z}^{\mathbb{S}_+}} \langle \ell \rangle^{-s} \lesssim_s \| u \|_{H^s_\varphi X}.$$

To see that $H^s(\mathbb{T}^{\mathbb{S}_+}, H) \hookrightarrow \mathcal{C}^0(\mathbb{T}^{\mathbb{S}_+}, H)$ for $s > |\mathbb{S}_+|/2$, use Plancherel's identity $\| u \|^2_{H^s_\varphi X} \sim_s \sum_{\ell \in \mathbb{Z}^{\mathbb{S}_+}} \| u_\ell \|_X^2 \langle \ell \rangle^{2s}$ to conclude that $\sum_{\ell \in \mathbb{Z}^{\mathbb{S}_+}} \| u_\ell \|_X \leqq (\sum_{\ell \in \mathbb{Z}^{\mathbb{S}_+}} \| u_\ell \|_X^2 \langle \ell \rangle^{2s})^{1/2} (\sum_{\ell \in \mathbb{Z}^{\mathbb{S}_+}} \langle \ell \rangle^{-2s})^{1/2} \lesssim_s \| u \|_{H^s_\varphi X}$.

On the Banach spaces $\mathcal{C}^s(\mathbb{T}^{\mathbb{S}_+}, X)$ the following interpolation inequalities hold: for any integer $0 \leqq k \leqq s$,

$$\| f \|_{\mathcal{C}^k_\varphi X} \lesssim_s \| f \|_{\mathcal{C}^0_\varphi X}^{1-\frac{k}{s}} \| f \|_{\mathcal{C}^s_\varphi X}^{\frac{k}{s}}. \tag{2.46}$$

In the next lemma, we assume the tame estimates (2.47) for the function $\mathfrak{x} = (\theta, y, w) \mapsto a(\mathfrak{x})$ in the x-variable only, and we deduce tame estimates for the composed function $a(\breve{\iota}(\varphi))$ in the variables (φ, x). Recall that $\mathcal{E}_s$, E_s are defined in (1.25) and $\mathcal{V}^s(\delta)$ in (1.27). Let Ω be an open bounded subset of $\mathbb{R}^{\mathbb{S}_+}$. In more detail, the following holds:

Lemma 2.25. (Tame estimates for functions) *Let $\sigma > 0$ and assume that, for any $s \geqq 0$, 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 $|\alpha| \leqq 1$, and $l \geqq 1$, the tame estimates*

$$\begin{aligned} &\| \partial_\omega^\alpha a(\mathfrak{x}; \omega) \|_{H^s_x} \lesssim_s 1 + \| w \|_{H^{s+\sigma}_x}, \\ &\| d^l \partial_\omega^\alpha a(\mathfrak{x}; \omega)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l] \|_{H^s_x} \lesssim_{s,l,\alpha} \sum_{j=1}^{l} \Big(\| \widehat{\mathfrak{x}}_j \|_{E_{s+\sigma}} \prod_{n \neq j} \| \widehat{\mathfrak{x}}_n \|_{E_\sigma} \Big) \end{aligned} \tag{2.47}$$

$$+ \|w\|_{H_x^{s+\sigma}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_\sigma}.$$

Then for any torus embedding $\breve{\iota}$ with $\|\iota\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} \leqq \delta$ and any maps $\widehat{\iota}, \widehat{\iota}_1, \widehat{\iota}_2 : \mathbb{T}^{\mathbb{S}_+} \to E_{s+s_0+\sigma}$, the following tame estimates hold for any $s \geqq 0$:

(i)
$$\begin{aligned}
&\|a(\breve{\iota})\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s 1 + \|\iota\|_{s+s_0+\sigma}^{\mathrm{Lip}(\gamma)},\\
&\|da(\breve{\iota})[\widehat{\iota}]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\widehat{\iota}\|_{s+s_0+\sigma}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+s_0+\sigma}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)},\\
&\|d^2a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\widehat{\iota}_1\|_{s+s_0+\sigma}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}_2\|_{s_0+\sigma}\\
&\qquad + \|\widehat{\iota}_1\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}_2\|_{s+s_0+\sigma}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+s_0+\sigma}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}_1\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}_2\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}.
\end{aligned} \tag{2.48}$$

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

(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*

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

Proof. (i) It suffices to prove the estimates for $\|d^2a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|_s$ and $\|d^2a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|_s^{\mathrm{lip}}$ in (2.48) since the ones for $a(\breve{\iota})$ and $da(\breve{\iota})$ then follow by Taylor expansions. By the hypothesis (2.47) with $l = 2$, $\alpha = 0$, we have, for any $\varphi \in \mathbb{T}^{\mathbb{S}_+}$, $s \geqq 0$,

$$\begin{aligned}
\|d^2a(\breve{\iota}(\varphi))[\widehat{\iota}_1(\varphi), \widehat{\iota}_2(\varphi)]\|_{H_x^s} \lesssim_s \|\widehat{\iota}_1(\varphi)\|_{E_{s+\sigma}} \|\widehat{\iota}_2(\varphi)\|_{E_\sigma} + \|\widehat{\iota}_1(\varphi)\|_{E_\sigma} \|\widehat{\iota}_2(\varphi)\|_{E_{s+\sigma}}\\
+ \|\iota(\varphi)\|_{E_{s+\sigma}} \|\widehat{\iota}_1(\varphi)\|_{E_\sigma} \|\widehat{\iota}_2(\varphi)\|_{E_\sigma}.
\end{aligned} \tag{2.49}$$

Squaring the expressions on the left and right hand side of (2.49) and then integrating them with respect to φ, one concludes, using (2.42), (2.43), and the Sobolev embedding (2.45), that

$$\begin{aligned}
&\|d^2a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|_{L^2_\varphi H^s_x} \lesssim_s \|\widehat{\iota}_1\|_{s+\sigma} \|\widehat{\iota}_2\|_{s_0+\sigma}\\
&+ \|\widehat{\iota}_1\|_{s_0+\sigma} \|\widehat{\iota}_2\|_{s+\sigma} + \|\iota\|_{s+\sigma} \|\widehat{\iota}_1\|_{s_0+\sigma} \|\widehat{\iota}_2\|_{s_0+\sigma}.
\end{aligned} \tag{2.50}$$

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

$$\begin{aligned}
\|d^2a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|_{\mathcal{C}^s_\varphi L^2_x} \lesssim_s \|\widehat{\iota}_1\|_{s_0+\sigma} \|\widehat{\iota}_2\|_{s+s_0+\sigma}\\
+ \|\widehat{\iota}_1\|_{s+s_0+\sigma} \|\widehat{\iota}_2\|_{s_0+\sigma} + \|\iota\|_{s+s_0+\sigma} \|\widehat{\iota}_1\|_{s_0+\sigma} \|\widehat{\iota}_2\|_{s_0+\sigma},
\end{aligned} \tag{2.51}$$

so that the estimate for $\|d^2a(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]\|_s$ stated in (2.48) follows by (2.50), (2.51), 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 (2.51). By the Leibnitz rule, for any $\beta \in \mathbb{N}^{\mathbb{S}_+}$, $0 \leqq |\beta| \leqq s$,

$$\partial_\varphi^\beta \Big(d^2a(\breve{\iota}(\varphi))[\widehat{\iota}_1(\varphi), \widehat{\iota}_2(\varphi)] \Big)$$

$$= \sum_{\beta_1+\beta_2+\beta_3=\beta} c_{\beta_1,\beta_2,\beta_3} \partial_\varphi^{\beta_1}(d^2 a(\breve{\iota}(\varphi)))\big[\partial_\varphi^{\beta_2}\widehat{\iota}_1(\varphi), \partial_\varphi^{\beta_3}\widehat{\iota}_2(\varphi)\big], \tag{2.52}$$

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

$$\partial_\varphi^{\beta_1}(d^2 a(\breve{\iota}(\varphi)))\big[\partial_\varphi^{\beta_2}\widehat{\iota}_1(\varphi), \partial_\varphi^{\beta_3}\widehat{\iota}_2(\varphi)\big]$$
$$= \sum_{\substack{1\leqq m\leqq|\beta_1| \\ \alpha_1+\cdots+\alpha_m=\beta_1}} c_{\alpha_1,\cdots,\alpha_m} d^{m+2}a(\breve{\iota}(\varphi))\big[\partial_\varphi^{\alpha_1}\breve{\iota}(\varphi), \cdots, \partial_\varphi^{\alpha_m}\breve{\iota}(\varphi), \partial_\varphi^{\beta_2}\widehat{\iota}_1(\varphi), \partial_\varphi^{\beta_3}\widehat{\iota}_2(\varphi)\big]$$

for suitable combinatorial constants $c_{\alpha_1,\cdots,\alpha_m}$. Then, by (2.47) with $l = m+2$, $\alpha = 0$, we have the bound

$$\begin{aligned} &\|\partial_\varphi^{\beta_1}(d^2 a(\breve{\iota}))[\partial_\varphi^{\beta_2}\widehat{\iota}_1, \partial_\varphi^{\beta_3}\widehat{\iota}_2]\|_{\mathcal{C}^0_\varphi L^2_x} \\ &\lesssim_\beta \sum_{\substack{1\leqq m\leqq|\beta_1| \\ \alpha_1+\cdots+\alpha_m=\beta_1}} (1+\|\iota\|_{\mathcal{C}^{|\alpha_1|}_\varphi E_\sigma})\cdots(1+\|\iota\|_{\mathcal{C}^{|\alpha_m|}_\varphi E_\sigma})\|\widehat{\iota}_1\|_{\mathcal{C}^{|\beta_2|}_\varphi E_\sigma}\|\widehat{\iota}_2\|_{\mathcal{C}^{|\beta_3|}_\varphi E_\sigma}. \end{aligned} \tag{2.53}$$

Arguing as in the proof of the formula (75) in [9], 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 (2.46), we get

$$\begin{aligned} &(1+\|\iota\|_{\mathcal{C}^{|\alpha_1|}_\varphi E_\sigma})\cdots(1+\|\iota\|_{\mathcal{C}^{|\alpha_m|}_\varphi E_\sigma})\|\widehat{\iota}_1\|_{\mathcal{C}^{|\beta_2|}_\varphi E_\sigma}\|\widehat{\iota}_2\|_{\mathcal{C}^{|\beta_3|}_\varphi E_\sigma} \\ &\lesssim_s \|\widehat{\iota}_1\|_{\mathcal{C}^0_\varphi E_\sigma}^{1-\frac{|\beta_2|}{|\beta|}}\|\widehat{\iota}_1\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma}^{\frac{|\beta_2|}{|\beta|}}\|\widehat{\iota}_2\|_{\mathcal{C}^0_\varphi E_\sigma}^{1-\frac{|\beta_3|}{|\beta|}}\|\widehat{\iota}_2\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma}^{\frac{|\beta_3|}{|\beta|}}\prod_{j=1}^m(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})^{1-\frac{|\alpha_j|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma})^{\frac{|\alpha_j|}{|\beta|}} \\ &\lesssim_s \|\widehat{\iota}_1\|_{\mathcal{C}^0_\varphi E_\sigma}^{1-\frac{|\beta_2|}{|\beta|}}\|\widehat{\iota}_1\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma}^{\frac{|\beta_2|}{|\beta|}}\|\widehat{\iota}_2\|_{\mathcal{C}^0_\varphi E_\sigma}^{1-\frac{|\beta_3|}{|\beta|}}\|\widehat{\iota}_2\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma}^{\frac{|\beta_3|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})^{m-\sum_{j=1}^m\frac{|\alpha_j|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^{|\beta|}_\varphi E_\sigma})^{\sum_{j=1}^m\frac{|\alpha_j|}{|\beta|}}. \end{aligned} \tag{2.54}$$

By (2.45), (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

$$\begin{aligned} (2.54) &\lesssim_s \|\widehat{\iota}_1\|_{\mathcal{C}^0_\varphi E_\sigma}^{\frac{|\beta_1|+|\beta_3|}{|\beta|}}\|\widehat{\iota}_1\|_{\mathcal{C}^s_\varphi E_\sigma}^{\frac{|\beta_2|}{|\beta|}}\|\widehat{\iota}_2\|_{\mathcal{C}^0_\varphi E_\sigma}^{\frac{|\beta_1|+|\beta_2|}{|\beta|}}\|\widehat{\iota}_2\|_{\mathcal{C}^s_\varphi E_\sigma}^{\frac{|\beta_3|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})^{\frac{|\beta_2|+|\beta_3|}{|\beta|}}(1+\|\iota\|_{\mathcal{C}^s_\varphi E_\sigma})^{\frac{|\beta_1|}{|\beta|}} \\ &\lesssim_s \Big(\|\widehat{\iota}_1\|_{\mathcal{C}^0_\varphi E_\sigma}\|\widehat{\iota}_2\|_{\mathcal{C}^0_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^s_\varphi E_\sigma})\Big)^{\frac{|\beta_1|}{|\beta|}}\Big(\|\widehat{\iota}_1\|_{\mathcal{C}^s_\varphi E_\sigma}\|\widehat{\iota}_2\|_{\mathcal{C}^0_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})\Big)^{\frac{|\beta_2|}{|\beta|}} \\ &\quad\times\Big(\|\widehat{\iota}_1\|_{\mathcal{C}^0_\varphi E_\sigma}\|\widehat{\iota}_2\|_{\mathcal{C}^s_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma})\Big)^{\frac{|\beta_3|}{|\beta|}} \end{aligned}$$

and, by the iterated Young inequality with exponents $|\beta|/|\beta_1|$, $|\beta|/|\beta_2|$, $|\beta|/|\beta_3|$, we conclude that the expression on the line (2.54) is bounded by

$$\begin{aligned} &\|\widehat{\iota}_1\|_{\mathcal{C}^0_\varphi E_\sigma}\|\widehat{\iota}_2\|_{\mathcal{C}^0_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^s_\varphi E_\sigma}) + \|\widehat{\iota}_1\|_{\mathcal{C}^s_\varphi E_\sigma}\|\widehat{\iota}_2\|_{\mathcal{C}^0_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma}) \\ &\quad+\|\widehat{\iota}_1\|_{\mathcal{C}^0_\varphi E_\sigma}\|\widehat{\iota}_2\|_{\mathcal{C}^s_\varphi E_\sigma}(1+\|\iota\|_{\mathcal{C}^0_\varphi E_\sigma}) \\ &\overset{(2.45),(2.43)}{\lesssim_s} \|\iota\|_{s+s_0+\sigma}\|\widehat{\iota}_1\|_{s_0+\sigma}\|\widehat{\iota}_2\|_{s_0+\sigma} \end{aligned}$$

$$+ \|\widehat{\iota}_1\|_{s+s_0+\sigma} \|\widehat{\iota}_2\|_{s_0+\sigma} + \|\widehat{\iota}_1\|_{s_0+\sigma} \|\widehat{\iota}_2\|_{s+s_0+\sigma}.$$

Clearly, the term (2.53) satisfies the same type of bound as (2.54). The left hand side in (2.52) with $\beta_1 = 0$ is estimated in the same way and thus (2.51) is proved.
PROOF (ii)–(iii). Let $\varphi \mapsto \breve{\iota}(\varphi) = (\theta(\varphi), y(\varphi), w(\varphi))$ be a torus embedding. If $a(\theta, 0, 0) = 0$, one has by the mean value theorem $a(\breve{\iota}) = \int_0^1 da(\breve{\iota}_t)[\widehat{\iota}]\, dt$ with $\breve{\iota}_t := (1 - t)(\theta(\varphi), 0, 0) + t\breve{\iota}(\varphi)$ and $\widehat{\iota} := (0, y(\varphi), w(\varphi))$. If $a(\theta, 0, 0)$, $\partial_y a(\theta, 0, 0)$, $\partial_w a(\theta, 0, 0)$ vanish, we write $a(\breve{\iota}) = \int_0^1 (1-t) d^2 a(\breve{\iota}_t)[\widehat{\iota}, \widehat{\iota}]\, dt$.

Items (ii)–(iii) follow by (i), noting that $\|\widehat{\iota}\|_s^{\mathrm{Lip}(\gamma)} = \|(0, y(\cdot), w(\cdot))\|_s^{\mathrm{Lip}(\gamma)} \lesssim \|\iota\|_s^{\mathrm{Lip}(\gamma)}$ for any $s \geqq 0$. □

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

$$\mathfrak{s}_M := \max\{s_0, M + 1\}. \tag{2.55}$$

Lemma 2.26. (Tame estimates for smoothing operators) *Assume that, for any $M \geqq 0$, there is $\sigma_M \geqq 0$ so that the following holds:*

- **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 $|\alpha| \leqq 1$,

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

and, for any $l \geqq 1$,

$$\|d^l \partial_\omega^\alpha \mathcal{R}(\mathfrak{x}; \omega)[\widehat{w}][\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{s+M+1}} \lesssim_{s,M,l} \|\widehat{w}\|_{H_x^s} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}$$
$$+ \|\widehat{w}\|_{L_x^2} \Big(\|w\|_{H_x^{s+\sigma_M}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} + \sum_{j=1}^{l} \big(\|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{n \neq j} \|\widehat{\mathfrak{x}}_n\|_{E_{\sigma_M}} \big) \Big).$$

- **Assumption B.** *For any $-M - 1 \leqq s \leqq 0$, 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 $\mathcal{C}^\infty$ with respect 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 $|\alpha| \leqq 1$, and $l \geqq 1$,

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

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

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 $\breve{\iota}(\varphi) = (\varphi, 0, 0) + \iota(\varphi)$ with $\|\iota\|_{s_0+\sigma_M(\lambda)}^{\mathrm{Lip}(\gamma)} \leqq \delta$ and any $n_1, n_2 \in \mathbb{N}$ satisfying $n_1 + n_2 \leqq M + 1$, 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* $\mathfrak{s}_M \leqq s \leqq S$,

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

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

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

(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* $\mathfrak{s}_M \leqq s \leqq S$,

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

Remark 2.27. Let us comment on Lemma 2.26 and its applications. Under the above assumptions A and B for the operator valued map $\mathfrak{x} = (\theta, y, w) \mapsto \mathcal{R}(\mathfrak{x})$, with $\mathcal{R}(\mathfrak{x})$ acting on spaces of functions of the x-variable only, we obtained tame estimates for the composed operator $\mathcal{R}(\breve{\imath}(\varphi))$, acting on spaces of functions in the variables (φ, x). Assumption B regarding the action of $\mathcal{R}(\mathfrak{x})$ on negative Sobolev spaces is used to prove that also $\langle D\rangle^{n_1}\mathcal{R}(\breve{\imath}(\varphi))\langle D\rangle^{n_2}$ with $n_1 + n_2 \leqq M+1$ is a modulo-tame operator. Lemma 2.26 will be used in the proof of Lemma 6.4 to show that the remainder $\mathcal{R}_M^{(1)}$ in the expansion (6.23) is a tame operator satisfying (6.31). The verification that $\mathcal{R}_M^{(1)}$ satisfies the assumptions A and B of Lemma 2.26 is proved by applying Lemmata 3.5 and 3.7.

Proof. Since items (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 $n_1 + n_2 \leqq M+1$, set $\mathcal{Q} := \langle D\rangle^{n_1}\mathcal{R}\langle D\rangle^{n_2}$. Assumption A implies that for any $s \geqq M+1$ and any $\mathfrak{x} \in \mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}$, the operator $\mathcal{Q}(\mathfrak{x})$ is in $\mathcal{B}(H^s_x)$ and for any $\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l \in E_{s+\sigma_M}$ with $l \geqq 1$, and $\widehat{w} \in H^s_x$,

$$\begin{aligned}
&\|\mathcal{Q}(\mathfrak{x})[\widehat{w}]\|_{H^s_x} \lesssim_{s,M} \|\widehat{w}\|_{H^s_x} + \|w\|_{H^{s+\sigma_M}_x}\|\widehat{w}\|_{H^{M+1}_x},\\
&\|d^l\big(\mathcal{Q}(\mathfrak{x})[\widehat{w}]\big)[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l]\|_{H^s_x} \lesssim_{s,M,l} \|\widehat{w}\|_{H^s_x}\prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\\
&\quad + \|\widehat{w}\|_{H^{M+1}_x}\Big(\|\mathfrak{x}\|_{E_{s+\sigma_M}}\prod_{j=1}^l\|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} + \sum_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}}\prod_{n\neq j}\|\widehat{\mathfrak{x}}_n\|_{E_{\sigma_M}}\Big).
\end{aligned} \tag{2.56}$$

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

$$\|\mathcal{Q}(\mathfrak{x})\|_{\mathcal{B}(L^2_x)} \lesssim_M 1, \quad \|d^l\mathcal{Q}(\mathfrak{x})[\widehat{\mathfrak{x}}_1,\dots,\widehat{\mathfrak{x}}_l]\|_{\mathcal{B}(L^2_x)} \lesssim_{M,l} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}. \tag{2.57}$$

One computes by Leibniz's rule that

$$\partial_\varphi^\lambda\big(\mathrm{d}\mathcal{Q}(\breve\imath(\varphi))[\widehat\imath(\varphi)]\big) = \sum_{\substack{0\leqq k\leqq|\lambda| \\ \lambda_1+\ldots+\lambda_{k+1}=\lambda}} c_{\lambda_1,\ldots,\lambda_{k+1}} d^{k+1}\mathcal{Q}(\breve\imath(\varphi))[\partial_\varphi^{\lambda_1}\breve\imath(\varphi),\ldots,\partial_\varphi^{\lambda_k}\breve\imath(\varphi),\partial_\varphi^{\lambda_{k+1}}\widehat\imath(\varphi)], \tag{2.58}$$

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 (2.56), we have, for $s\geqq M+1$,

$$\begin{aligned}
&\|d^{k+1}\mathcal{Q}(\breve\imath(\varphi))[\partial_\varphi^{\lambda_1}\breve\imath(\varphi),\ldots,\partial_\varphi^{\lambda_k}\breve\imath(\varphi),\partial_\varphi^{\lambda_{k+1}}\widehat\imath(\varphi)][\widehat w(\varphi)]\|_{H^s_x} \\
&\lesssim_{s,M,k} \|\widehat w(\varphi)\|_{H^s_x}\|\partial_\varphi^{\lambda_{k+1}}\widehat\imath(\varphi)\|_{E_{\sigma_M}}\prod_{n=1}^k\|\partial_\varphi^{\lambda_n}\breve\imath(\varphi)\|_{E_{\sigma_M}} \\
&\quad+\|\widehat w(\varphi)\|_{H^{M+1}_x}\Big(\|\iota(\varphi)\|_{E_{s+\sigma_M}}\|\partial_\varphi^{\lambda_{k+1}}\widehat\imath(\varphi)\|_{E_{\sigma_M}}\prod_{n=1}^k\|\partial_\varphi^{\lambda_n}\breve\imath(\varphi)\|_{E_{\sigma_M}} \\
&\quad+\sum_{j=1}^k\|\partial_\varphi^{\lambda_j}\breve\imath(\varphi)\|_{E_{s+\sigma_M}}\Big(\prod_{n\neq j}\|\partial_\varphi^{\lambda_n}\breve\imath(\varphi)\|_{E_{\sigma_M}}\Big)\|\partial_\varphi^{\lambda_{k+1}}\widehat\imath(\varphi)\|_{E_{\sigma_M}} \\
&\quad+\|\partial_\varphi^{\lambda_{k+1}}\widehat\imath(\varphi)\|_{E_{s+\sigma_M}}\prod_{n=1}^k\|\partial_\varphi^{\lambda_n}\breve\imath(\varphi)\|_{E_{\sigma_M}}\Big).
\end{aligned}\tag{2.59}$$

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

$$\|\partial_\varphi^\mu\breve\imath(\varphi)\|_{E_s}\lesssim 1+\|\partial_\varphi^\mu\iota\|_{C^0_\varphi E_s}\lesssim 1+\|\iota\|_{s+s_0+|\mu|}. \tag{2.60}$$

Hence (2.58)–(2.59) and $\|\cdot\|_{L^2_\varphi H^s_x}\lesssim\|\cdot\|_s$ imply that for any $\breve\imath$ with $\|\iota\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_M(\lambda)}\leqq\delta$ and any $s\geqq M+1$,

$$\begin{aligned}
&\|\partial_\varphi^\lambda\big(\mathrm{d}\mathcal{Q}(\breve\imath(\varphi))[\widehat\imath(\varphi)]\big)[\widehat w(\varphi)]\|_{L^2_\varphi H^s_x} \\
&\lesssim_{s,M,\lambda}\|\widehat w\|_s\|\widehat\imath\|_{s_0+\sigma_M(\lambda)}+\|\widehat w\|_{M+1}\big(\|\iota\|_{s+\sigma_M(\lambda)}\|\widehat\imath\|_{s_0+\sigma_M(\lambda)}+\|\widehat\imath\|_{s+\sigma_M(\lambda)}\big)
\end{aligned}\tag{2.61}$$

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 $s\in\mathbb{N}$, $\beta\in\mathbb{N}^{\mathbb{S}_+}$, $|\beta|\leqq s$, 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 (2.58) we have

$$\begin{aligned}
&\partial_\varphi^{\beta+\lambda}\big(d\mathcal{Q}(\breve\imath(\varphi))[\widehat\imath(\varphi)]\big) \\
&=\sum_{\substack{0\leqq k\leqq|\beta|+|\lambda| \\ \alpha_1+\ldots+\alpha_{k+1}=\beta+\lambda}} c_{\alpha_1,\ldots,\alpha_{k+1}}d^{k+1}\mathcal{Q}(\breve\imath(\varphi))[\partial_\varphi^{\alpha_1}\breve\imath(\varphi),\ldots,\partial_\varphi^{\alpha_k}\breve\imath(\varphi),\partial_\varphi^{\alpha_{k+1}}\widehat\imath(\varphi)]
\end{aligned}\tag{2.62}$$

where $c_{\alpha_1,\ldots,\alpha_{k+1}}$ are combinatorial constants. By (2.57) and (2.60) one obtains that

$$\|d^{k+1}\mathcal{Q}(\breve\imath(\varphi))[\partial_\varphi^{\alpha_1}\breve\imath(\varphi),\ldots,\partial_\varphi^{\alpha_k}\breve\imath(\varphi),\partial_\varphi^{\alpha_{k+1}}\widehat\imath(\varphi)]\|_{L^2_\varphi\mathcal{B}(L^2_x)}$$

$$\lesssim_{\beta,\lambda} \prod_{j=1}^{k}(1+\|\iota\|_{|\alpha_j|+\eta_M})\|\widehat{\iota}\|_{|\alpha_{k+1}|+\eta_M} \tag{2.63}$$

for some $\eta_M > 0$. Using the interpolation inequality (2.4), and arguing as in the proof of the formula (75) in [9], we have, for any $\breve{\iota}$ with $\|\iota\|_{\eta_M} \leqq 1$ and any $j = 1, \ldots, k$,

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

Then by (2.63) and since $\sum_{j=1}^{k}|\alpha_j| + |\alpha_{k+1}| = |\beta+\lambda|$, it follows that

$$\begin{aligned}
&\|d^{k+1}\mathcal{Q}(\breve{\iota}(\varphi))[\partial_\varphi^{\alpha_1}\breve{\iota}(\varphi), \ldots, \partial_\varphi^{\alpha_k}\breve{\iota}(\varphi), \partial_\varphi^{\alpha_{k+1}}\widehat{\iota}(\varphi)]\|_{L^2_\varphi\mathcal{B}(L^2_x)} \\
&\lesssim_{s,\lambda} (1+\|\iota\|_{|\beta+\lambda|+\eta_M})^{\frac{\sum_{j=1}^{k}|\alpha_j|}{|\beta+\lambda|}}\|\widehat{\iota}\|_{\eta_M}^{1-\frac{|\alpha_{k+1}|}{|\beta+\lambda|}}\|\widehat{\iota}\|_{|\beta+\lambda|+\eta_M}^{\frac{|\alpha_{k+1}|}{|\beta+\lambda|}} \\
&\lesssim_{s,\lambda} \Big((1+\|\iota\|_{|\beta+\lambda|+\eta_M})\|\widehat{\iota}\|_{\eta_M}\Big)^{\frac{\sum_{j=1}^{k}|\alpha_j|}{|\beta+\lambda|}}\|\widehat{\iota}\|_{|\beta+\lambda|+\eta_M}^{\frac{|\alpha_{k+1}|}{|\beta+\lambda|}} \\
&\lesssim_{s,\lambda} \|\widehat{\iota}\|_{|\beta+\lambda|+\eta_M} + \|\iota\|_{|\beta+\lambda|+\eta_M}\|\widehat{\iota}\|_{\eta_M}
\end{aligned} \tag{2.64}$$

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 (2.62) and (2.64) we obtain

$$\|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])\|_{H^s_\varphi\mathcal{B}(L^2_x)} \lesssim_{s,M,\lambda} \|\widehat{\iota}\|_{s+|\lambda|+\eta_M} + \|\iota\|_{s+|\lambda|+\eta_M}\|\widehat{\iota}\|_{\eta_M}. \tag{2.65}$$

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 [9, Lemma 2.12] that for any $\breve{\iota}$ with $\|\iota\|_{2s_0+|\lambda|+\eta_M} \leqq 1$ and any $s \geqq s_0$,

$$\begin{aligned}
\|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])[\widehat{w}]\|_{H^s_\varphi L^2_x} &\lesssim_s \|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])\|_{H^{2s_0}_\varphi\mathcal{B}(L^2_x)}\|\widehat{w}\|_{H^s_\varphi L^2_x} \\
&\quad + \|\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])\|_{H^{s+s_0}_\varphi\mathcal{B}(L^2_x)}\|\widehat{w}\|_{H^{s_0}_\varphi L^2_x} \\
&\overset{(2.65)}{\lesssim_{s,M}} \|\widehat{w}\|_s\|\widehat{\iota}\|_{2s_0+|\lambda|+\eta_M} \\
&\quad + \|\widehat{w}\|_{s_0}\big(\|\widehat{\iota}\|_{s+s_0+|\lambda|+\eta_M} + \|\iota\|_{s+s_0+|\lambda|+\eta_M}\|\widehat{\iota}\|_{2s_0+|\lambda|+\eta_M}\big).
\end{aligned} \tag{2.66}$$

Increasing the constant $\sigma_M(\lambda)$ in (2.61) if needed, one infers from the estimates (2.61), (2.66) that for any $s \geqq \mathfrak{s}_M = \max\{s_0, M+1\}$, $\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\iota})[\widehat{\iota}])$ satisfies

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

$$+\|\widehat{w}\|_{\mathfrak{s}_M}\big(\|\widehat{\imath}\|_{s+\sigma_M(\lambda)}+\|\imath\|_{s+\sigma_M(\lambda)}\|\widehat{\imath}\|_{s_0+\sigma_M(\lambda)}\big). \quad (2.67)$$

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

$$\gamma \frac{\|\partial_\varphi^\lambda \Delta_\omega(d\mathcal{Q}(\breve{\imath})[\widehat{\imath}])[\widehat{w}]\|_s}{|\omega_1-\omega_2|} \lesssim_{s,M,\lambda} \|\widehat{w}\|_s \|\widehat{\imath}\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_M(\lambda)} + \|\widehat{w}\|_{\mathfrak{s}_M}\big(\|\widehat{\imath}\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M(\lambda)}+\|\imath\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M(\lambda)}\|\widehat{\imath}\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_M(\lambda)}\big). \quad (2.68)$$

It then follows from (2.67) and (2.68) that there exists a tame constant $\mathfrak{M}_{\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\imath})[\widehat{\imath}])}(s)$ for $\partial_\varphi^\lambda(d\mathcal{Q}(\breve{\imath})[\widehat{\imath}])$ satisfying the estimate stated in item (ii).
(iii) Since $\mathcal{R}(\theta,0,0)=0$, one has by the mean value theorem $\mathcal{R}(\breve{\imath}) = \int_0^1 \mathrm{d}\mathcal{R}(\breve{\imath}_t)[\widehat{\imath}]\,\mathrm{d}t$ with $\breve{\imath}_t=(1-t)(\theta(\varphi),0,0)+t\breve{\imath}(\varphi)$ and $\widehat{\imath}(\varphi):=(0,y(\varphi),w(\varphi))$. Since $\|\widehat{\imath}\|_s \lesssim \|\imath\|_s$ for any $s \geqq 0$, item (iii) is thus a direct consequence of (ii). □

2.5. Egorov Type Theorems

In this section we 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. The main result is an Egorov type theorem, stated in Proposition 2.31, saying that such a conjugated operator is again a pseudo-differential operator, up to a smoothing remainder; it is used in Section 6.3.

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

$$\partial_\tau \Phi(\tau_0,\tau,\varphi) = B(\tau,\varphi)\Phi(\tau_0,\tau,\varphi), \quad \Phi(\tau_0,\tau_0,\varphi)=\mathrm{Id}, \quad (2.69)$$

where $B(\tau,\varphi)$ is the transport operator, given by

$$B(\tau,\varphi) := \Pi_\perp\big(b(\tau,\varphi,x)\partial_x + b_x(\tau,\varphi,x)\big), \qquad b \equiv b(\tau,\varphi,x) := \frac{\beta(\varphi,x)}{1+\tau\beta_x(\varphi,x)}, \quad (2.70)$$

$\Pi_\perp$ is the L^2_x-orthogonal projector $L^2_x \to L^2_\perp(\mathbb{T}_1)$, and $\beta(\varphi,x) \equiv \beta(\varphi,x;\omega)$ is a real valued function, which is $\mathcal{C}^\infty$ with respect to the variables (φ,x) and Lipschitz continuous 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

$$\Phi(\tau_0,\tau,\varphi)=\Phi(\tau,\varphi)\circ\Phi(\tau_0,\varphi)^{-1}. \quad (2.71)$$

By standard hyperbolic estimates, equation (2.69) is well-posed. The flow $\Phi(\tau_0,\tau,\varphi)$ has the following properties:

Lemma 2.28. (Transport flow) *Let $\lambda_0 \in \mathbb{N}$, $S > s_0$. For any $\lambda \in \mathbb{N}$ with $\lambda \leqq \lambda_0$, any $n_1,n_2 \in \mathbb{R}$ with $n_1+n_2=-\lambda-1$, and any $s \geqq s_0$, there exist constants $\sigma(\lambda_0,n_1,n_2)>0$ and $\delta \equiv \delta(S,\lambda_0,n_1,n_2) \in (0,1)$ so that the following holds: if $\beta(\varphi,x))$ satisfies*

$$\|\beta\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma(\lambda_0,n_1,n_2)} \leqq \delta, \quad (2.72)$$

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

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

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

$$\mathfrak{M}_{\langle D\rangle^{n_1}\partial^{\lambda}_{\varphi_m}(\Phi(\tau_0,\tau,\varphi)-\mathrm{Id})\langle D\rangle^{n_2}}(s) \lesssim_{S,\lambda_0,n_1,n_2} \|\beta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma(\lambda_0,n_1,n_2)},$$
$$\forall s_0 \leqq s \leqq S, \quad \forall \tau_0,\tau\in[0,1]. \tag{2.74}$$

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

$$\|\langle D\rangle^{n_1}\partial^{\lambda}_{\varphi_m}\Delta_{12}\Phi(\tau_0,\tau,\varphi)\langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})}$$
$$\lesssim_{s_1,\lambda_0,n_1,n_2} \|\Delta_{12}\beta\|_{s_1+\sigma(n_1,n_2)}, \quad \forall\,\tau_0,\tau\in[0,1], \tag{2.75}$$

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 (2.73) is similar to the one of Propositions A.7, A.10 and A.11 in [10] and hence we omit it. (Essentially the only difference is that the vector field (2.70) is of order 1, whereas the vector field considered in [10] is of order $\frac{1}{2}$.) Using (2.73) we now prove (2.74). By (2.69), one has that $\Phi(\tau_0,\tau,\varphi)-\mathrm{Id}=\int_{\tau_0}^{\tau}B(t,\varphi)\Phi(\tau_0,t,\varphi)\,\mathrm{d}t$. Then, for any $\lambda\in\mathbb{N}$ with $\lambda\leqq\lambda_0$ and any $n_1,n_2\in\mathbb{R}$ with $n_1+n_2=-\lambda-2$, one has, by Leibniz' rule that,

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

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 (2.70) of B, using Lemmata 2.10, 2.18, 2.30-(i), and (2.73), one has that for any $s\geqq s_0$,

$$\mathfrak{M}_{\langle D\rangle^{n_1}\partial^{\lambda_1}_{\varphi_m}B\langle D\rangle^{-1-n_1-\lambda_1}}(s)\lesssim_s |\langle D\rangle^{n_1}B\langle D\rangle^{-1-n_1-\lambda_1}|^{\mathrm{Lip}(\gamma)}_{0,s+\lambda_1,0}\lesssim_{s,\lambda_1,n_1}\|\beta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma(\lambda_1,n_1)},$$
$$\mathfrak{M}_{\langle D\rangle^{-1-n_2-\lambda_2}\partial^{\lambda_2}_{\varphi_m}\Phi(\tau_0,t,\varphi)\langle D\rangle^{n_2}}(s)\lesssim_{s,\lambda_2,n_1,n_2} 1+\|\beta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma(\lambda_2,n_1,n_2)}. \tag{2.76}$$

Then (2.74) follows by (2.76), Lemma 2.16 and (2.72). The estimate (2.75) follows by similar arguments. □

For what follows we need to study the solutions of the characteristic ODE $\partial_\tau x=-b(\tau,\varphi,x)$ associated with the transport operator defined in (2.70).

Lemma 2.29. (Characteristic flow) *The characteristic flow* $\gamma^{\tau_0,\tau}(\varphi, x)$ *defined by*

$$\partial_\tau \gamma^{\tau_0,\tau}(\varphi, x) = -b\big(\tau, \varphi, \gamma^{\tau_0,\tau}(\varphi, x)\big), \quad \gamma^{\tau_0,\tau_0}(\varphi, x) = x, \tag{2.77}$$

is given by

$$\gamma^{\tau_0,\tau}(\varphi, x) = x + \tau_0\beta(\varphi, x) + \breve{\beta}(\tau, \varphi, x + \tau_0\beta(\varphi, x)), \tag{2.78}$$

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

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 (2.78). □

Lemma 2.30. *There are constants* σ *and* $\delta > 0$ *so that the following holds: if* $\|\beta\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma} \leqq \delta$, *then*

(i) $\|b\|^{\mathrm{Lip}(\gamma)}_s \lesssim_s \|\beta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma}$ *for any* $s \geqq s_0$;
(ii) $\|\gamma^{\tau_0,\tau}(\varphi, x) - x\|^{\mathrm{Lip}(\gamma)}_s \lesssim_s \|\beta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma}$ *for any* $\tau_0, \tau \in [0, 1]$ *and* $s \geqq s_0$;
(iii) *for any* $\|\beta_j\|_{s_1+\sigma} \leq \delta$, $j = 1, 2$ *with* $s_1 > s_0$, $\Delta_{12}b := b(\cdot; \beta_2) - b(\cdot; \beta_1)$ *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*

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

Proof. Item (i) follows from the definition of b in (2.70) and Lemma 2.3. Item (ii) can be deduced from Lemma 2.1 and (2.78) and item (iii) follows by similar arguments. □

The main result of this section is the Egorov type theorem below, 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 (2.69), remains a pseudo-differential operator with a homogenous asymptotic expansion up to a regularizing remainder satisfying the quantitative tame estimate (2.83).

Proposition 2.31. (Egorov) *Let* $N, \lambda_0 \in \mathbb{N}$ *and* $S > s_0$ *be given and assume that* $\beta(\cdot, \cdot; \omega)$ *and* $a(\cdot, \cdot; \omega)$ *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* $C_0 > 0$ *so that the following holds: if*

$$\|\beta\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_N(\lambda_0)} \leqq \delta, \quad \|a\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_N(\lambda_0)} \leqq C_0, \tag{2.79}$$

then for any $m \in \mathbb{Z}$, *the conjugated operator*

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

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

$$\mathcal{P}(\varphi) = \sum_{i=0}^{N} p_{m-i}(\varphi, x; \omega)\partial_x^{m-i} + \mathcal{R}_N(\varphi) \tag{2.80}$$

with the following properties:

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

$$p_m(\varphi, x; \omega) = \Big((1 + \breve{\beta}_y(\varphi, y; \omega))^m a(\varphi, y; \omega)\Big)|_{y=x+\beta(\varphi,x;\omega)} \quad (2.81)$$

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

2. *For any* $s \geqq s_0$ *and* $i = 1, \dots, N$*,*

$$\|p_m - a\|_s^{\mathrm{Lip}(\gamma)},\ \|p_{m-i}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,N} \|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)} + \|a\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)} \|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)}. \quad (2.82)$$

3. *For any* $k \in \mathbb{S}_+$*, any* $\lambda \in \mathbb{N}$ *with* $\lambda \leqq \lambda_0$*, and any* $n_1, n_2 \in \mathbb{N}$ *with* $n_1 + n_2 + \lambda_0 \leqq N - 1 - m$*, 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*

$$\mathfrak{M}_{\langle D\rangle^{n_1} \partial_{\varphi_k}^{\lambda} \mathcal{R}_N(\varphi)\langle D\rangle^{n_2}}(s) \lesssim_{S,N,\lambda_0} \|\beta\|_{s+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)} + \|a\|_{s+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)} \|\beta\|_{s_0+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)}, \quad \forall s_0 \leqq s \leqq S. \quad (2.83)$$

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

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

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

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

where $\Delta_{12} a = a_2 - a_1$*,* $\Delta_{12}\beta = \beta_2 - \beta_1$*, and* $\Delta_{12}\mathcal{R}_N = \mathcal{R}_N^{(2)} - \mathcal{R}_N^{(1)}$ *with* $\mathcal{R}_N^{(j)}$ *denoting the remainder in* (2.80)*, corresponding to* a_j*,* β_j *for* $j = 1, 2$*.*

Proof. The L_x^2-orthogonal projector $\Pi_\perp : L^2(\mathbb{T}_1) \to L_\perp^2(\mathbb{T}_1)$ 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 $\mathbb{S} \cup \{0\}$. We then decompose the operator $B(\tau, \varphi) = \Pi_\perp(b(\tau, \varphi, x)\partial_x + b_x(\tau, \varphi, x))$ as $B(\tau, \varphi) = B_1(\tau, \varphi) + B_\infty(\tau, \varphi)$ with

$$\begin{aligned} B_1(\tau, \varphi) &:= b(\tau, \varphi, x)\partial_x + b_x(\tau, \varphi, x), \\ B_\infty(\tau, \varphi) &:= \mathrm{Op}(b_\infty(\tau, \varphi, x, \xi)) \in OPS^{-\infty} \end{aligned} \quad (2.84)$$

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

$$|B_\infty|_{-m,s,\alpha}^{\mathrm{Lip}(\gamma)} \lesssim_{m,s,\alpha} \|\beta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}. \quad (2.85)$$

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

$$\partial_\tau \mathcal{P}(\tau, \varphi) = [B(\tau, \varphi), \mathcal{P}(\tau, \varphi)], \quad \mathcal{P}(0, \varphi) = \mathcal{P}_0(\varphi) = a(\varphi, x; \omega)\partial_x^m. \quad (2.86)$$

Indeed, one has

$$
\begin{aligned}
\partial_\tau \mathcal{P}(\tau,\varphi) &= \partial_\tau \Phi(\tau,\varphi)\mathcal{P}_0(\varphi)\Phi(\tau,\varphi)^{-1} + \Phi(\tau,\varphi)\mathcal{P}_0(\varphi)\partial_\tau\Big(\Phi(\tau,\varphi)^{-1}\Big) \\
&= \partial_\tau \Phi(\tau,\varphi)\mathcal{P}_0(\varphi)\Phi(\tau,\varphi)^{-1} - \Phi(\tau,\varphi)\mathcal{P}_0(\varphi)\Phi(\tau,\varphi)^{-1}\partial_\tau\Phi(\tau,\varphi)\Phi(\tau,\varphi)^{-1} \\
&\overset{(2.69)}{=} B(\tau,\varphi)\Phi(\tau,\varphi)\mathcal{P}_0(\varphi)\Phi(\tau,\varphi)^{-1} \\
&\quad - \Phi(\tau,\varphi)\mathcal{P}_0(\varphi)\Phi(\tau,\varphi)^{-1}B(\tau,\varphi)\Phi(\tau,\varphi)\Phi(\tau,\varphi)^{-1} \\
&= [B(\tau,\varphi),\mathcal{P}(\tau,\varphi)]. \qquad (2.87)
\end{aligned}
$$

We look for an approximate solution of (2.86) of the form

$$
\mathcal{P}_N(\tau,\varphi) := \sum_{i=0}^{N} p_{m-i}(\tau,\varphi,x)\partial_x^{m-i} \qquad (2.88)
$$

for suitable functions $p_{m-i}(\tau,\varphi,x)$ to be determined. By (2.84)

$$
[B(\tau,\varphi),\mathcal{P}_N(\tau,\varphi)] = [B_1(\tau,\varphi),\mathcal{P}_N(\tau,\varphi)] + [B_\infty(\tau,\varphi),\mathcal{P}_N(\tau,\varphi)], \qquad (2.89)
$$

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 2.12, one has for any $i = 0,\dots,N$,

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

where the functions $g_j(b,p_{m-i})(\tau,\varphi,x)$, $j = 0,\dots,N-i$, 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 $s \geqq s_0$, (cf. Lemma 2.30-(i))

$$
\begin{aligned}
\|g_j(b,p_{m-i})\|_s^{\mathrm{Lip}(\gamma)} &\lesssim_{m,N,s} \|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}\|p_{m-i}\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} \\
&\quad + \|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)}\|p_{m-i}\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}, \qquad (2.90)
\end{aligned}
$$

and for any $s \geqq s_0$ and $\alpha \in \mathbb{N}$ (cf. Lemma 2.12-(ii))

$$
\begin{aligned}
|\mathcal{R}_N(b,p_{m-i})|_{m-N-1,s,\alpha}^{\mathrm{Lip}(\gamma)} &\lesssim_{m,N,s,\alpha} \|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}\|p_{m-i}\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} \\
&\quad + \|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)}\|p_{m-i}\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}. \qquad (2.91)
\end{aligned}
$$

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

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

$$
\begin{aligned}
&= \sum_{i=0}^{N} \big(b(p_{m-i})_x - (m-i) b_x p_{m-i}\big)\partial_x^{m-i} \\
&\quad + \sum_{k=1}^{N}\sum_{j=1}^{k} g_j(b, p_{m-k+j})\partial_x^{m-k} + \sum_{i=0}^{N} \mathcal{R}_N(b, p_{m-i}) \\
&= \big(b(p_m)_x - m b_x p_m\big)\partial_x^m \\
&\quad + \sum_{i=1}^{N} \big(b(p_{m-i})_x - (m-i) b_x p_{m-i} + \widetilde{g}_i\big)\partial_x^{m-i} + \mathcal{Q}_N,
\end{aligned} \tag{2.92}
$$

where

$$
\mathcal{Q}_N := \sum_{i=0}^{N} \mathcal{R}_N(b, p_{m-i}) \in OPS^{m-N-1}, \quad \widetilde{g}_i := \sum_{j=1}^{i} g_j(b, p_{m-i+j}), \quad \forall 1 \le i \le N. \tag{2.93}
$$

Defining for any $s \geqq 0$ and $1 \le i \le N$,

$$
\begin{aligned}
\mathtt{M}_{<i}(s) &:= \max\{\| p_{m-k}\|_s^{\mathrm{Lip}(\gamma)},\ k = 0, \ldots, i-1\}, \\
\mathtt{M}(s) &:= \max\{\| p_{m-i}\|_s^{\mathrm{Lip}(\gamma)},\ i = 0, \ldots, N\},
\end{aligned} \tag{2.94}
$$

we deduce from (2.90) and (2.91) that for any $s \geqq s_0$, $\alpha \in \mathbb{N}$, $i = 0, \ldots, N$,

$$
\begin{aligned}
&\|\widetilde{g}_i\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,N} \mathtt{M}_{<i}(s+\sigma_N)\|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} + \mathtt{M}_{<i}(s_0+\sigma_N)\|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)} \\
&|\mathcal{Q}_N|_{m-N-1,s,\alpha}^{\mathrm{Lip}(\gamma)} \lesssim_{s,N} \mathtt{M}(s+\sigma_N)\|\beta\|_{s_0+\sigma_N}^{\mathrm{Lip}(\gamma)} + \mathtt{M}(s_0+\sigma_N)\|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}.
\end{aligned} \tag{2.95}
$$

By (2.88), (2.89), and (2.92) 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

$$
\begin{aligned}
&\partial_\tau p_m = b(p_m)_x - m b_x p_m, \\
&\partial_\tau p_{m-i} = b(p_{m-i})_x - (m-i) b_x p_{m-i} + \widetilde{g}_i, \quad \forall i = 1, \ldots, N.
\end{aligned} \tag{2.96}
$$

Note that, since $\widetilde{g}_i$ only depends on $p_{m-i+1}, \ldots, p_m$, we can solve (2.96) inductively.

Determination of p_m. We solve the first equation in (2.96),

$$
\begin{aligned}
\partial_\tau p_m(\tau, \varphi, x) &= b(\tau, \varphi, x)\partial_x p_m(\tau, \varphi, x) - m b_x(\tau, \varphi, x) p_m(\tau, \varphi, x), \\
p_m(0, \varphi, x) &= a(\varphi, x).
\end{aligned}
$$

By the method of characteristics we deduce that

$$
p_m(\tau, \varphi, \gamma^{0,\tau}(\varphi, x)) = \exp\Big(-m \int_0^\tau b_x(t, \varphi, \gamma^{0,t}(\varphi, x))\, \mathrm{d}t\Big) a(\varphi, x) \tag{2.97}
$$

where $\gamma^{0,\tau}(\varphi, x)$ is given by (2.78). Differentiating the equation (2.77) 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

$$\partial_x \gamma^{\tau_0,\tau}(\varphi, x) = \exp\Big(-\int_{\tau_0}^{\tau} b_x(t, \varphi, \gamma^{\tau_0,t}(\varphi, x))\, \mathrm{d}t \Big). \tag{2.98}$$

From (2.97) and (2.98) we infer that

$$p_m(\tau, \varphi, y) = \Big(\big(\partial_x \gamma^{0,\tau}(\varphi, x)\big)^m a(\varphi, x) \Big)|_{x=\gamma^{\tau,0}(\varphi,y)}. \tag{2.99}$$

Evaluating the latter identity at $\tau = 1$ and using (2.78), we obtain (2.81).
Inductive determination of p_{m-i}. For $i = 1, \dots, N$, we solve the inhomogeneous transport equation,

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

By the method of characteristics one has

$$p_{m-i}(\tau, \varphi, y) = \int_0^\tau \exp\Big(-(m-i)\int_t^\tau b_x(s, \varphi, \gamma^{\tau,s}(\varphi, y))\, \mathrm{d}s \Big) \widetilde{g}_i(t, \varphi, \gamma^{\tau,t}(\varphi, y))\, \mathrm{d}t. \tag{2.100}$$

The functions $p_{m-i}(\varphi, y)$ in the expansion (2.80) are then given by $p_{m-i}(\varphi, y) := p_{m-i}(1, \varphi, y)$. Next we prove the estimates for p_{m-i} stated in (2.82). They follow from the following

Lemma 2.32. *There exist $\sigma_N^{(N)} > \sigma_N^{(N-1)} > \cdots > \sigma_N^{(0)} > 0$ so that for any $i \in \{1, \dots, N\}$, $\tau \in [0, 1]$, and $s \geqq s_0$,*

$$\begin{aligned} \|p_m(\tau, \cdot) - a\|_s^{\mathrm{Lip}(\gamma)} &\lesssim_s \|\beta\|_{s+\sigma_N^{(0)}}^{\mathrm{Lip}(\gamma)} + \|a\|_{s+\sigma_N^{(0)}}^{\mathrm{Lip}(\gamma)} \|\beta\|_{s_0+\sigma_N^{(0)}}^{\mathrm{Lip}(\gamma)}, \\ \|p_{m-i}(\tau, \cdot)\|_s^{\mathrm{Lip}(\gamma)} &\lesssim_s \|\beta\|_{s+\sigma_N^{(i)}}^{\mathrm{Lip}(\gamma)} + \|a\|_{s+\sigma_N^{(i)}}^{\mathrm{Lip}(\gamma)} \|\beta\|_{s_0+\sigma_N^{(i)}}. \end{aligned} \tag{2.101}$$

Proof of Lemma 2.32. We argue by induction. First we prove the claimed estimate for $p_m - a$ with p_m given by (2.99). 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. (2.78)). Since $a(\varphi, y + \tau\beta(\varphi, y)) - a(\varphi, y) = \int_0^\tau a_x(\varphi, y + t\beta(\varphi, y))\beta(\varphi, y)\mathrm{d}t$, the claimed estimate for p_m then follows by Lemmata 2.1, 2.30 and assumption (2.79). Now assume that for any $k \in \{1, \dots, i-1\}$, $1 \le i \le N$, the function p_{m-k}, given by (2.100), satisfies the estimates (2.101). The ones for p_{m-i} then follow by Lemmata 2.1, 2.3, 2.30, (2.95), (2.94), and (2.79). □

Continuing the proof of Proposition 2.31, note that in view of the definition (2.88) of $\mathcal{P}_N(\tau,\varphi)$, it follows from (2.101), Lemma 2.10, (2.22) and (2.21) that for any $\alpha \in \mathbb{N}$, $\tau \in [0,1]$, and $s \geqq s_0$,

$$|\mathcal{P}_N(\tau,\varphi)|^{\mathrm{Lip}(\gamma)}_{m,s,\alpha} \lesssim_{m,s,N,\alpha} \|a\|^{\mathrm{Lip}(\gamma)}_{s} + \|\beta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_N^{(N)}} + \|a\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_N^{(N)}}\|\beta\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_N^{(N)}}. \tag{2.102}$$

By (2.89), (2.92), and (2.96) we deduce that $\mathcal{P}_N(\tau,\varphi)$ solves

$$\begin{aligned} &\partial_\tau \mathcal{P}_N(\tau,\varphi) = [B(\tau,\varphi),\mathcal{P}_N(\tau,\varphi)] - \mathcal{Q}^{(1)}_N(\tau,\varphi), \quad \mathcal{P}_N(0,\varphi) = a\partial_x^m, \\ &\mathcal{Q}^{(1)}_N(\tau,\varphi) := \mathcal{Q}_N(\tau,\varphi) + [B_\infty(\tau,\varphi),\mathcal{P}_N(\tau,\varphi)] \in OPS^{m-N-1}, \end{aligned} \tag{2.103}$$

where $\mathcal{Q}_N$ is defined in (2.93).

Next we estimate the difference between $\mathcal{P}_N(\tau)$ and $\mathcal{P}(\tau)$. First we establish the following formula:

Lemma 2.33. *The operator $\mathcal{R}_N(\tau,\varphi) := \mathcal{P}(\tau,\varphi) - \mathcal{P}_N(\tau,\varphi)$ is given by*

$$\mathcal{R}_N(\tau,\varphi) = \int_0^\tau \Phi(\eta,\tau,\varphi)\mathcal{Q}^{(1)}_N(\eta,\varphi)\Phi(\tau,\eta,\varphi)\,d\eta. \tag{2.104}$$

Proof of Lemma 2.33. Writing $\mathcal{P}_N(\tau,\varphi) - \mathcal{P}(\tau,\varphi)$ as

$$\begin{aligned} \mathcal{P}_N(\tau,\varphi) - \mathcal{P}(\tau,\varphi) &= \mathcal{V}_N(\tau,\varphi)\Phi(\tau,\varphi)^{-1}, \\ \mathcal{V}_N(\tau,\varphi) &:= \mathcal{P}_N(\tau,\varphi)\Phi(\tau,\varphi) - \Phi(\tau,\varphi)\mathcal{P}_0(\varphi), \end{aligned} \tag{2.105}$$

one verifies by a straightforward calculation that $\mathcal{V}_N(\tau)$ solves

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

where $\mathcal{Q}^{(1)}_N$ is given in (2.103). By the variation of the constants formula,

$$\mathcal{V}_N(\tau,\varphi) = -\int_0^\tau \Phi(\tau,\varphi)\Phi(\eta,\varphi)^{-1}\mathcal{Q}^{(1)}_N(\eta,\varphi)\Phi(\eta,\varphi)\,\mathrm{d}\eta$$

and, by (2.105) and (2.71), we deduce (2.104). □

Using formula (2.104) we now prove the estimate for $\mathcal{R}_N(\tau,\varphi)$ stated in (2.83) of Proposition 2.31. First we estimate $\mathcal{Q}^{(1)}_N = \mathcal{Q}_N(\tau,\varphi) + [B_\infty(\tau,\varphi),\mathcal{P}_N(\tau,\varphi)] \in OPS^{m-N-1}$ (cf. (2.103)). The estimate of $\mathcal{Q}_N$, obtained from (2.95), (2.94), (2.101), and the one of $[B_\infty(\tau,\varphi),\mathcal{P}_N(\tau,\varphi)]$, obtained from (2.85), (2.102), Lemma 2.11, imply that there exists a constant $\aleph_N > 0$ so that for any $s \geqq s_0$, $\alpha \in \mathbb{N}$,

$$|\mathcal{Q}^{(1)}_N(\eta,\varphi)|^{\mathrm{Lip}(\gamma)}_{m-N-1,s,\alpha} \lesssim_{m,s,\alpha,N} \|\beta\|^{\mathrm{Lip}(\gamma)}_{s+\aleph_N} + \|a\|^{\mathrm{Lip}(\gamma)}_{s+\aleph_N}\|\beta\|^{\mathrm{Lip}(\gamma)}_{s_0+\aleph_N}. \tag{2.106}$$

Let $k \in \mathbb{S}_+$, λ_0 with $\lambda \leqq \lambda_0$, and $n_1, n_2 \in \mathbb{N}$ with $n_1 + n_2 + \lambda_0 + m \leqq N-1$. In view of the formula (2.104) of $\mathcal{R}_N(\tau,\varphi)$, the claimed estimate of $\langle D\rangle^{n_1}\partial^\lambda_{\varphi_k}\mathcal{R}_N(\tau,\varphi)\langle D\rangle^{n_2}$ follows from corresponding ones of

$$\langle D\rangle^{n_1}\partial^{\lambda_1}_{\varphi_k}\Phi(\eta,\tau,\varphi)\partial^{\lambda_2}_{\varphi_k}\mathcal{Q}^{(1)}_N(\eta,\varphi)\partial^{\lambda_3}_{\varphi_k}\Phi(\tau,\eta,\varphi)\langle D\rangle^{n_2}$$

($\tau, \eta \in [0,1]$ and $\lambda_1 + \lambda_2 + \lambda_3 = \lambda$) which we write as

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

By Lemma 2.28, one obtains tame constants for the operators

$$\langle D\rangle^{n_1}\partial_{\varphi_k}^{\lambda_1}\Phi(\eta,\tau,\varphi)\langle D\rangle^{-n_1-\lambda_1-1}, \quad \langle D\rangle^{-n_2-\lambda_3-1}\partial_{\varphi_k}^{\lambda_3}\Phi(\tau,\eta,\varphi)\langle D\rangle^{n_2},$$

and by the estimates (2.106), (2.21), and Lemmata 2.10, 2.18 a tame constant for

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

allowing us to deduce that the composition of these three operators satisfies the bound (2.83) (using also Lemma 2.16 together with the assumption (2.79)). This proves the bound (2.83) for $\mathcal{R}_N$.

Item 4. of Proposition 2.31 can be shown by similar arguments. □

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 (2.69). A straightforward calculation shows that

$$\Phi(\varphi)\circ(\omega\cdot\partial_\varphi)\circ\Phi(\varphi)^{-1} = \omega\cdot\partial_\varphi + \Phi(\varphi)\circ\omega\cdot\partial_\varphi(\Phi(\varphi)^{-1}),$$

where, according to Definition 2.4-4, for any φ-dependent family of linear operators $A(\varphi)$, the operator $\omega\cdot\partial_\varphi A(\varphi)$ is defined as

$$\omega\cdot\partial_\varphi A(\varphi) = \sum_{m\in\mathbb{S}_+}\omega_m\partial_{\varphi_m}A(\varphi) = \sum_{m\in\mathbb{S}_+}\omega_m[\partial_{\varphi_m}, A(\varphi)].$$

We now show that the operator $\Phi(\varphi)\circ\omega\cdot\partial_\varphi(\Phi(\varphi)^{-1})$ is a pseudo-differential operator of order one, admitting an expansion in decreasing symbols. More precisely, the following holds:

Proposition 2.34. (Conjugation of $\omega\cdot\partial_\varphi$) *Let $N, \lambda_0 \in \mathbb{N}$ and $S > s_0$ and assume that $\beta(\cdot,\cdot;\omega)$ 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 $\sigma_N(\lambda_0)$, $\sigma_N > 0$, $\delta(S,N,\lambda_0) \in (0,1)$, and $C_0 > 0$ so that the following holds: if*

$$\|\beta\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma_N(\lambda_0)} \leqq \delta, \tag{2.107}$$

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

$$\Psi(\varphi) = \sum_{i=0}^{N} p_{1-i}(\varphi,x;\omega)\partial_x^{1-i} + \mathcal{R}_N(\varphi) \tag{2.108}$$

with the following properties:

1. *For any* $i = 0, \dots, N$ *and* $s \geqq s_0$, $\|p_{1-i}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,N} \|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}$.
2. *For any* $k \in \mathbb{S}_+$, *any* $\lambda \in \mathbb{N}$ *with* $\lambda \leqq \lambda_0$, *and any* $n_1, n_2 \in \mathbb{N}$ *with* $n_1+n_2+\lambda_0 \leqq N-1-m$, *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*

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

3. *Let* $s_0 < s_1 < S$ *and assume that* $\|\beta_j\|_{s_1+\sigma_N(\lambda_0)} \leqq \delta$, $j = 1, 2$. *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 $\lambda \leqq \lambda_0$, $n_1, n_2 \in \mathbb{N}$ *with* $n_1 + n_2 + \lambda_0 \leqq N - 2$, *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 $\Delta_{12}\beta = \beta_2 - \beta_1$, $\Delta_{12} p_{1-i} = p_{1-i}^{(2)} - p_{1-i}^{(1)}$, *and* $\Delta_{12}\mathcal{R}_N = \mathcal{R}_N^{(2)} - \mathcal{R}_N^{(1)}$. *Here* $p_{1-i}^{(j)}$ *and* $\mathcal{R}_N^{(j)}$ *denote the coefficient* p_{m-i} *and the remainder* $\mathcal{R}_N$ *in the expansion* (2.108)*, corresponding to* β_j *for* $j = 1, 2$.

Proof. We have that $\Psi(\varphi) = \Psi(1, \varphi)$ where $\Psi(\tau, \varphi) := \Phi(\tau, \varphi)\circ\omega\cdot\partial_\varphi(\Phi(\tau, \varphi)^{-1})$. Arguing as in (2.87), one sees that the operator $\Psi(\tau, \varphi)$ 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 (2.86) by looking for approximate solutions of the form of a pseudo-differential operator of order 1, admitting an expansion of the form (2.108) (cf. (2.88)). The proof then proceeds in the same way as the one for Proposition 2.31 and hence is omitted. □

We finish this section by the following application of Proposition 2.31 to Fourier multipliers.

Lemma 2.35. *Let* $N, \lambda_0 \in \mathbb{N}$, $m \in \mathbb{Z}$, *and* $S > s_0$ *and assume that* $\mathcal{Q}$ *is a Lipschitz family of Fourier multipliers with an expansion of the form*

$$\mathcal{Q} = \sum_{i=0}^{N} c_{m-i}(\omega)\partial_x^{m-i} + \mathcal{Q}_N(\omega), \quad \mathcal{Q}_N(\omega) \in \mathcal{B}(H^s, H^{s+N+1-m}), \quad \forall s \geqq 0. \tag{2.109}$$

Then there exist $\sigma_N(\lambda_0)$, $\sigma_N > 0$, *and* $\delta = \delta(S, N, \lambda_0) \in (0, 1)$ *so that the following holds: if*

$$\|\beta\|_{s_0+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)} \leqq \delta, \tag{2.110}$$

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. $\mathcal{Q}_\Phi(\varphi) = \sum_{i=0}^N \alpha_{m-i}(\varphi, x; \omega)\partial_x^{m-i}$ *where for any* $s \geqq s_0$,

$$\|\alpha_{m-i}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,N} \|\beta\|_{s+\sigma_N}^{\mathrm{Lip}(\gamma)}, \quad i = 0, \dots, N. \tag{2.111}$$

2. *For any* $k \in \mathbb{S}_+$, $\lambda \in \mathbb{N}$ *with* $\lambda \leqq \lambda_0$, *and* $n_1, n_2 \in \mathbb{N}$ *with* $n_1 + n_2 + \lambda_0 \leqq N - m - 2$, *the operator* $\langle D\rangle^{n_1}\partial_{\varphi_k}^\lambda \mathcal{R}_N \langle D\rangle^{n_2}$ *is* $\mathrm{Lip}(\gamma)$*-tame with a tame constant satisfying*

$$\mathfrak{M}_{\langle D\rangle^{n_1}\partial_{\varphi_k}^\lambda \mathcal{R}_N\langle D\rangle^{n_2}}(s) \lesssim_{S,N,\lambda_0} \|\beta\|_{s+\sigma_N(\lambda_0)}^{\mathrm{Lip}(\gamma)}, \quad \forall s_0 \leqq s \leqq S. \tag{2.112}$$

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

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

and, for any $k \in \mathbb{S}_+$, $\lambda \leqq \lambda_0$, *and* $n_1, n_2 \in \mathbb{N}$ *with* $n_1 + n_2 + \lambda_0 \leqq N - m - 2$,

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

where $\Delta_{12}\beta = \beta_2 - \beta_1$, $\Delta_{12}\alpha_{m-i} = \alpha_{m-i}^{(2)} - \alpha_{m-i}^{(1)}$, *and* $\Delta_{12}\mathcal{R}_N = \mathcal{R}_N^{(2)} - \mathcal{R}_N^{(1)}$. *Here* $\alpha_{m-i}^{(j)}$ *and* $\mathcal{R}_N^{(j)}$ *denote the coefficient* α_{m-i} *and, respectively, remainder* $\mathcal{R}_N$, *corresponding to* β_j *for* $j = 1, 2$.

Proof. Applying Proposition 2.31 to $\Phi(\varphi)\partial_x^{m-i}\Phi(\varphi)^{-1}$ for $i = 0, \dots, N$, we get

$$\Phi(\varphi)\Big(\sum_{i=0}^N c_{m-i}(\omega)\partial_x^{m-i}\Big)\Phi(\varphi)^{-1} = \sum_{i=0}^N c_{m-i}(\omega)\partial_x^{m-i} + \mathcal{Q}_\Phi(\varphi) + \mathcal{R}_N^{(1)}(\varphi)$$

where $\mathcal{Q}_\Phi(\varphi) = \sum_{i=0}^N \alpha_{m-i}(\varphi, x; \omega)\partial_x^{m-i}$ with α_{m-i} satisfying (2.111) and the remainder $\mathcal{R}_N^{(1)}(\varphi)$ satisfying (2.112). 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 2.31. Using Lemma 2.28 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 (2.112), and $\mathcal{R}_N(\varphi) = \mathcal{R}_N^{(1)}(\varphi) + \mathcal{R}_N^{(2)}(\varphi)$ satisfies (2.112) as well. Item 3. follows by similar arguments. □

3. Integrable Features of KdV

In this section we discuss the canonical coordinates which are used to prove the existence of quasi-periodic solutions of the perturbed KdV equation (1.6) close to the finite gap manifold $\mathcal{M}_{\mathbb{S}_+}$, cf. (1.10), via a Nash–Moser iterative scheme. These coordinates were constructed in [20] specifically for this purpose. Their main properties were described in broad terms in the Introduction (cf. $(P1)$–$(P6)$). We discuss their features in detail in Sections 3.2–3.3. In Section 3.4 we record properties of the KdV-frequencies, which will be needed in particular to derive the measure estimates of Section 8.2.

3.1. Birkhoff Coordinates

According to [21], the KdV equation (1.1) on the torus admits global canonical coordinates, called Birkhoff coordinates, so that equation (1.1) can be solved by quadrature, cf. Theorem 3.1. We use them to describe the $\mathbb{S}_+$-gap potentials. Unfortunately they are not suited for implementing a Nash–Moser iteration scheme for the search of quasi-periodic solutions of (1.6), since they do not seem to possess an expansion in terms of pseudo-differential operators.

The Birkhoff coordinates z_k, $k \neq 0$, take values in the sequence space h_0^0 (cf. (1.23)), which we endow with the standard Poisson bracket defined by $\{z_n, z_k\} = \mathrm{i}2\pi n\, \delta_{k,-n}$ for any $n, k \in \mathbb{Z} \setminus \{0\}$.

Theorem 3.1. (Birkhoff coordinates, [21]) *There exists a real analytic diffeomorphism $\Psi^{kdv} : h_0^0 \to H_0^0(\mathbb{T}_1)$ so that the following holds:*

- (i) *for any $s \in \mathbb{Z}_{\geqq 0}$, $\Psi^{kdv}(h_0^s) \subseteq H_0^s(\mathbb{T}_1)$ 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 $\mathcal{F}^{-1}$.*

By Theorem 3.1, the KdV equation, expressed in the Birkhoff coordinates $(z_n)_{n\neq 0}$, reads

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

and its solutions are given by $z(t) := (z_n(t))_{n\neq 0}$ where

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

Let us consider a finite set $\mathbb{S}_+ \subset \mathbb{N}_+ = \{1, 2, \ldots\}$ and define

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

In Birkhoff coordinates, a $\mathbb{S}_+-$gap solution of the KdV equation, also referred to as $\mathbb{S}_+-$gap solution, is a solution of the form

$$z_n(t) = \exp\big(\mathrm{i}\omega_n^{kdv}(\nu, 0)t\big) z_n(0), \ z_n(0) \neq 0, \ \forall n \in \mathbb{S}, \quad z_n(t) = 0, \ \forall n \in \mathbb{S}^\perp, \tag{3.1}$$

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

$$\omega_n^{kdv}(I, (I_k)_{k\in\mathbb{S}_+^\perp}) = \omega_n^{kdv}((I_k)_{k\geqq 1}), \quad I := (I_k)_{k\in\mathbb{S}_+} \in \mathbb{R}_{>0}^{\mathbb{S}_+}. \tag{3.2}$$

Such solutions are quasi-periodic in time with frequency vector (cf. (1.11))

$$\omega^{kdv}(\nu) = \big(\omega_n^{kdv}(\nu, 0)\big)_{n\in\mathbb{S}_+} \in \mathbb{R}^{\mathbb{S}_+}, \tag{3.3}$$

parametrized by $\nu \in \mathbb{R}^{\mathbb{S}_+}_{>0}$. The map $\nu \mapsto \omega^{kdv}(\nu)$ is a local analytic diffeomorphism, see Remark 3.10.

When written in action-angle coordinates $\theta = (\theta_n)_{n \in \mathbb{S}_+} \in \mathbb{T}^{\mathbb{S}_+}$, $I = (I_n)_{n \in \mathbb{S}_+} \in \mathbb{R}^{\mathbb{S}_+}_{>0}$, which are related to the complex Birkhoff coordinates $z_n = z_n(\theta, I)$, $n \neq 0$ by

$$z_{\pm n}(\theta, I) := \sqrt{2\pi n I_n} e^{\mp i\theta_n}, \quad \forall n \in \mathbb{S}_+, \quad z_n(\theta, I) = 0, \quad \forall n \in \mathbb{S}^\perp, \tag{3.4}$$

the $\mathbb{S}_+$-gap solution (3.1) reads

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

Furthermore, we introduce the map $\Psi_{\mathbb{S}_+} : \mathbb{T}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \to \mathcal{M}_{\mathbb{S}_+} \subset \cap_{s \geq 0} H^s_0(\mathbb{T}_1)$, which coordinatizes the manifold $\mathcal{M}_{\mathbb{S}_+}$ of $\mathbb{S}_+$-gap potentials (cf. (1.10)),

$$\Psi_{\mathbb{S}_+}(\theta, I) := \Psi^{kdv}((z_n(\theta, I))_{n \in \mathbb{Z} \setminus \{0\}}) \tag{3.5}$$

where $z_n(\theta, I)$, $n \neq 0$, are given by (3.4).

3.2. Normal Form Coordinates for the KdV Equation

Theorem 3.2 below rephrases Theorem 1.1 in [20], in a form taylored to our needs. A key property of the normal form coordinates is stated in Theorem 3.2-(**AE1**), saying that they admit an expansion in terms of pseudo-differential operators. This property, together with the additional Corollaries 3.3 and 3.4 below, allow to prove, in Section 3.3, that the linearized Hamiltonian vector field $\partial_x d_\perp \nabla_w \mathcal{H}_\varepsilon$ admits an expansion in terms of classical pseudo-differential operators, up to smoothing remainders which satisfy tame estimates. These key results are needed for implementing our diagonalization procedure of the linearized operator carried out in Sections 6–7.

We consider an open bounded set $\Xi \subset \mathbb{R}^{\mathbb{S}_+}_{>0}$ so that (1.14) 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 (1.27) and the spaces $\mathcal{E}_s$ and E_s are given by (1.25). The elements in $\mathcal{E}_s$ are denoted by $\mathfrak{x} = (\theta, y, w)$ whereas the ones in E_s by $\widehat{\mathfrak{x}} = (\widehat{\theta}, \widehat{y}, \widehat{w})$. The space $\mathcal{V}(\delta) \cap \mathcal{E}_s$ is endowed with the symplectic form

$$\mathcal{W} := \Big(\sum\nolimits_{j \in \mathbb{S}_+} \mathrm{d}y_j \wedge \mathrm{d}\theta_j \Big) \oplus \mathcal{W}_\perp, \tag{3.6}$$

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 (1.9). 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

$$\mathcal{J} : E_s \to E_s, \quad (\widehat{\theta}, \widehat{y}, \widehat{w}) \mapsto (-\widehat{y}, \widehat{\theta}, \partial_x \widehat{w}), \tag{3.7}$$

where $\langle\,,\,\rangle$ is the inner product (1.26).

Theorem 3.2. (Normal form KdV coordinates with pseudo-differential expansion, [20]) *Let $\mathbb{S}_+ \subseteq \mathbb{N}$ be finite, Ξ an open bounded subset of $\mathbb{R}^{\mathbb{S}_+}_{>0}$ so that (1.14) 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 extends $\Psi_{\mathbb{S}_+}$, introduced in (3.5), namely*

$$\Psi_\nu(\theta, y, 0) = \Psi_{\mathbb{S}_+}(\theta, \nu + y), \quad \forall (\theta, y, 0) \in \mathcal{V}(\delta), \quad \forall \nu \in \Xi, \tag{3.8}$$

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 that $\Psi_\nu : \mathcal{V}(\delta) \cap \mathcal{E}_s \to H^s_0(\mathbb{T}_1)$ is a $\mathcal{C}^\infty-$diffeomorphism onto its image, so that the following holds:

*(**AE1**) (Asymptotic expansion of Ψ_ν) 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*

$$\Psi_\nu(\theta, y, w) = \Psi_{\mathbb{S}_+}(\theta, \nu + y) + w + \sum_{k=1}^{M} a^{\Psi}_{-k}(\mathfrak{x}; \nu)\, \partial_x^{-k} w + \mathcal{R}^{\Psi}_M(\mathfrak{x}; \nu) \tag{3.9}$$

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

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

are $\mathcal{C}^\infty$.
*(**AE2**) (Asymptotic expansion of $d\Psi_\nu^\top$) For any $\mathfrak{x} \in \mathcal{V}^1(\delta)$ (cf. definition (1.27)), $\nu \in \Xi$, 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*

$$\begin{aligned} d\Psi_\nu(\mathfrak{x})^\top[\widehat{q}] &= \Big(0, 0,\ \Pi_\perp \widehat{q} + \Pi_\perp \sum_{k=1}^{M} a^{d\Psi^\top}_{-k}(\mathfrak{x}; \nu) \partial_x^{-k} \widehat{q} + \Pi_\perp \sum_{k=1}^{M} (\partial_x^{-k} w)\, \mathcal{A}^{d\Psi^\top}_{-k}(\mathfrak{x}; \nu)[\widehat{q}]\Big) \\ &\quad + \mathcal{R}^{d\Psi^\top}_M(\mathfrak{x}; \nu)[\widehat{q}] \end{aligned} \tag{3.10}$$

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,

$$a^{d\Psi^\top}_{-1}(\mathfrak{x}; \nu) = -a^{\Psi}_{-1}(\mathfrak{x}; \nu). \tag{3.11}$$

*(**AE3**) (Normal form) For any $\nu \in \Xi$, the Hamiltonian $\mathcal{H}^{kdv}(\cdot\,;\nu) := H^{kdv} \circ \Psi_\nu : \mathcal{V}^1(\delta) \to \mathbb{R}$ (cf. definition (1.27)) is in normal form up to order three, meaning that*

$$\begin{aligned}\mathcal{H}^{kdv}(\theta, y, w; \nu) &= \omega^{kdv}(\nu) \cdot y + \frac{1}{2}\big(\Omega^{kdv}(D;\nu)w, w\big)_{L^2_x} \\ &\quad + \frac{1}{2}\Omega^{kdv}_{\mathbb{S}_+}(\nu)[y]\cdot y + \mathcal{R}^{kdv}(\theta, y, w; \nu) \end{aligned} \tag{3.12}$$

where $\omega^{kdv}(\nu) = (\omega_n^{kdv}(\nu, 0))_{n\in\mathbb{S}_+}$,

$$\begin{aligned}&\Omega^{kdv}_{\mathbb{S}_+}(\nu) := (\partial_{I_j}\omega_k^{kdv}(\nu,0))_{j,k\in\mathbb{S}_+}, \qquad \Omega^{kdv}(D;\nu)w := \sum_{n\in\mathbb{S}^\perp}\Omega_n^{kdv}(\nu)w_n e^{\mathrm{i}2\pi n x}, \\ &\Omega_n^{kdv}(\nu) := \frac{1}{2\pi n}\omega_n^{kdv}(\nu,0),\ \forall n\in\mathbb{S}^\perp, \quad w = \sum_{n\in\mathbb{S}^\perp} w_n e^{\mathrm{i}2\pi n x},\end{aligned} \tag{3.13}$$

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

$$\mathcal{R}^{kdv}(\theta, y, w; \nu) = O\big((\|y\| + \|w\|_{H^1_x})^3 \big), \tag{3.14}$$

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

$$\nabla\mathcal{R}^{kdv}(\theta,0,0;\nu) = 0, \quad d_\perp\nabla\mathcal{R}^{kdv}(\theta,0,0;\nu) = 0, \quad \partial_y\nabla\mathcal{R}^{kdv}(\theta,0,0;\nu) = 0. \tag{3.15}$$

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

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

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

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

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

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

Similarly, for any $\nu \in \Xi$, $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\mathfrak{x} \in \mathcal{V}^1(\delta) \cap \mathcal{E}_s$, $\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l \in E_s$, $\widehat{q} \in H_0^s$, $s \geq 1$,

$$\|\partial_\nu^\alpha \mathcal{R}_M^{d\Psi^\top}(\mathfrak{x}; \nu)[\widehat{q}]\|_{E_{s+M+1}} \lesssim_{s,M,\alpha} \|\widehat{q}\|_{H_x^s} + \|w\|_{H_x^s} \|\widehat{q}\|_{H_x^1},$$

$$\|d^l\big(\partial_\nu^\alpha \mathcal{R}_M^{d\Psi^\top}(\mathfrak{x}; \nu)[\widehat{q}]\big)[\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l]\|_{E_{s+M+1}} \lesssim_{s,M,l,\alpha} \|\widehat{q}\|_{H_x^s} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_1}$$

$$+ \|\widehat{q}\|_{H_x^1} \sum_{j=1}^{l} \Big(\|\widehat{\mathfrak{x}}_j\|_{E_s} \prod_{i \neq j} \|\widehat{\mathfrak{x}}_i\|_{E_1} \Big) + \|\widehat{q}\|_{H_x^1} \|w\|_{H_x^s} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_1}.$$

We now apply Theorem 3.2 to obtain in the two corollaries below novel results concerning the extensions of $d\Psi_\nu(\mathfrak{x})^\top$ and $d\Psi_\nu(\mathfrak{x})$ to Sobolev spaces of negative order. These results will be used to deduce Lemmata 3.5 and 3.7, which allow to verify the assumptions A and B of Lemma 2.26 for the remainders, as explained in Remark 2.27. The spaces $\mathcal{E}_s$, E_s for negative s are defined as in (1.25).

Corollary 3.3. (Extension of $d\Psi_\nu(\mathfrak{x})^\top$ and its asymptotic expansion) *Let* $M \geqq 1$. *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*

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

with the following properties:

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

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

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 3.2-(**AE2**)) and for any* $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, *and* $(\mathfrak{x}, \nu) \in \mathcal{V}^{\sigma_M}(\delta) \times \Xi$,

$$\|\partial_\nu^\alpha a_{-k}^{ext}(\mathfrak{x}; \nu; d\Psi^\top)\|_{H_x^s} \lesssim_{s,M,\alpha} 1,$$
$$\|\partial_\nu^\alpha d^l a_{-k}^{ext}(\mathfrak{x}; \nu; d\Psi^\top)[\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l]\|_{H_x^s} \lesssim_{s,M,l,\alpha} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}. \tag{3.17}$$

(ii) *For any* $-1 \leqq s \leqq M+1$, *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* $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, $\widehat{q} \in H_0^{-s}(\mathbb{T}_1)$, *and* $(\mathfrak{x}, \nu) \in \mathcal{V}^{\sigma_M}(\delta) \times \Xi$,

$$\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x}; \nu; d\Psi^\top)[\widehat{q}]\|_{E_{M+1-s}} \lesssim_{M,\alpha} \|\widehat{q}\|_{H_x^{-s}},$$

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

(3.18)

(iii) *For any* $s \geqq 1$, *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* $(\mathfrak{x},\nu)\in\big(\mathcal{V}^{\sigma_M}(\delta)\cap\mathcal{E}_{s+\sigma_M}\big)\times\Xi$,

$$\begin{aligned}
&\|\partial_\nu^\alpha\mathcal{R}_M^{ext}(\mathfrak{x};\nu;d\Psi^\top)[\widehat{q}]\|_{E_{M+1+s}}\lesssim_{s,M,\alpha}\|\widehat{q}\|_{H_x^s}+\|\mathfrak{x}\|_{s+\sigma_M}\|\widehat{q}\|_{H_x^1},\\
&\|\partial_\nu^\alpha d^l\mathcal{R}_M^{ext}(\mathfrak{x};\nu;d\Psi^\top)[\widehat{\mathfrak{x}}_1,\ldots,\widehat{\mathfrak{x}}_l][\widehat{q}]\|_{E_{M+1+s}}\lesssim_{s,M,l,\alpha}\|\widehat{q}\|_{H_x^s}\prod_{j=1}^l\|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\\
&\qquad+\|\widehat{q}\|_{H_x^1}\Big(\sum_{j=1}^l\|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}}\prod_{i\neq j}\|\widehat{\mathfrak{x}}_i\|_{E_{\sigma_M}}+\|\mathfrak{x}\|_{E_{s+\sigma_M}}\prod_{j=1}^l\|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}\Big).
\end{aligned}\tag{3.19}$$

Proof. By Theorem 3.2, 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 $M\geqq 1$, differentiating (3.9), $d\Psi_\nu(\mathfrak{x})[\widehat{\mathfrak{x}}]$ admits the expansion for any $\widehat{\mathfrak{x}}=(\widehat{\theta},\widehat{y},\widehat{w})\in E_0$ of the form

$$\mathrm{d}\Psi_\nu(\mathfrak{x})[\widehat{\mathfrak{x}}]=\widehat{w}+\sum_{k=1}^M a_{-k}^\Psi(\mathfrak{x};\nu)\partial_x^{-k}\widehat{w}+\mathcal{R}_M^{(1)}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}],\tag{3.20}$$

$$\mathcal{R}_M^{(1)}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}]:=\sum_{k=1}^M(\partial_x^{-k}w)\mathrm{d}a_{-k}^\Psi(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}]+\mathrm{d}\mathcal{R}_M^\Psi(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}]+d_{\theta,y}\Psi_{\mathbb{S}_+}(\theta,\nu+y)[\widehat{\theta},\widehat{y}].$$

For $\sigma_M\geqq M$, 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 3.2-(**Est1**), for any $\alpha\in\mathbb{N}^{\mathbb{S}_+}$, $l\geqq 1$,

$$\begin{aligned}
&\|\partial_\nu^\alpha\mathcal{R}_M^{(1)}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}]\|_{H_x^{M+1}}\lesssim_{M,\alpha}\|\widehat{\mathfrak{x}}\|_{E_0},\\
&\|\partial_\nu^\alpha d^l\mathcal{R}_M^{(1)}(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}_1,\ldots,\widehat{\mathfrak{x}}_l][\widehat{\mathfrak{x}}]\|_{H_x^{M+1}}\lesssim_{M,l,\alpha}\|\widehat{\mathfrak{x}}\|_{E_0}\prod_{j=1}^l\|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}.
\end{aligned}\tag{3.21}$$

Now consider the transpose operator $d\Psi_\nu(\mathfrak{x})^\top:L_0^2(\mathbb{T}_1)\to E_0$. By (3.20), for any $\widehat{q}\in L_0^2(\mathbb{T}_1)$, one has

$$d\Psi_\nu(\mathfrak{x})^\top[\widehat{q}]=\Big(0,0,\Pi_\perp\widehat{q}+\Pi_\perp\sum_{k=1}^M(-1)^k\partial_x^{-k}\big(a_{-k}^\Psi(\mathfrak{x};\nu)\,\widehat{q}\big)\Big)+\mathcal{R}_M^{(1)}(\mathfrak{x};\nu)^\top[\widehat{q}].\tag{3.22}$$

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 (3.22) 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 (2.12), the expansion (3.22) yields one of the form (3.16) where by (3.21) and Theorem 3.2-(**Est1**), the remainder $\mathcal{R}_M^{ext}(\mathfrak{x}; \nu; \mathrm{d}\Psi^\top)$ satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, and $\widehat{q} \in H_0^{-M-1}(\mathbb{T}_1)$

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x}; \nu; \mathrm{d}\Psi^\top)[\widehat{q}]\|_{E_0} \lesssim_{M,\alpha} \|\widehat{q}\|_{H_x^{-M-1}}, \\ &\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x}; \nu; d\Psi^\top)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l][\widehat{q}]\|_{E_0} \lesssim_{M,l,\alpha} \|\widehat{q}\|_{H_x^{-M-1}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}. \end{aligned} \tag{3.23}$$

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 (3.10) and, by the uniqueness of an expansion of this form,

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

The claimed estimates (3.17) and (3.19) then follow by Theorem 3.2-(**Est2**). In particular we have, for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, $\widehat{q} \in H_0^1(\mathbb{T}_1)$,

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

Finally the estimates (3.18) follow by interpolation between (3.23) and (3.24). □

Corollary 3.4. (Extension of $d_\perp \Psi_\nu(\mathfrak{x})$ and its asymptotic expansion) *Let $M \geqq 1$. 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*

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

with the following properties:

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

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

*are $\mathcal{C}^\infty$. They satisfy $a_{-1}^{ext}(\mathfrak{x}; \nu; d_\perp \Psi) = a_{-1}^{\Psi}(\mathfrak{x}; \nu)$ (cf. Theorem 3.2-(**AE1**)) and for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, and $(\mathfrak{x}, \nu) \in \mathcal{V}^{\sigma_M}(\delta) \times \Xi$,*

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

(ii) *For any $0 \leqq s \leqq M+2$, 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 $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, $\widehat{w} \in H_\perp^{-s}(\mathbb{T}_1)$, and $(\mathfrak{x}, \nu) \in \mathcal{V}^{\sigma_M}(\delta) \times \Xi$,

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

(iii) *For any $s \geqq 0$, 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 $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{s+\sigma_M}$, $\widehat{w} \in H_\perp^{s}(\mathbb{T}_1)$, and $(\mathfrak{x}, \nu) \in \big(\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}\big) \times \Xi$,

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x}; \nu; d_\perp \Psi)[\widehat{w}]\|_{H_x^{M+1+s}} \lesssim_{s,M,\alpha} \|\widehat{w}\|_{H_x^s} + \|\mathfrak{x}\|_{E_{s+\sigma_M}} \|\widehat{w}\|_{L_x^2}, \\ &\|\partial_\nu^\alpha d^l \big(\mathcal{R}_M^{ext}(\mathfrak{x}; \nu; d_\perp \Psi)[\widehat{w}]\big)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{M+1+s}} \lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H_x^s} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} \\ &\quad + \|\widehat{w}\|_{L_x^2} \Big(\sum_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{i \neq j} \|\widehat{\mathfrak{x}}_i\|_{E_{\sigma_M}} + \|\mathfrak{x}\|_{E_{s+\sigma_M}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} \Big). \end{aligned} \tag{3.28}$$

Proof. By Theorem 3.2-(**AE2**), for any $(\mathfrak{x}, \nu) \in \mathcal{V}^1(\delta) \times \Xi$, the operator $d_\perp \Psi_\nu(\mathfrak{x})^\top : H_0^1(\mathbb{T}_1) \to H_\perp^1(\mathbb{T}_1)$ is bounded and for any $M \geqq 1$ and $\widehat{q} \in H_0^1(\mathbb{T}_1)$, $d_\perp \Psi_\nu(\mathfrak{x})^\top[\widehat{q}\,]$ admits the expansion of the form

$$\begin{aligned} &d_\perp \Psi_\nu(\mathfrak{x})^\top[\widehat{q}\,] = \Pi_\perp \widehat{q} + \Pi_\perp \sum_{k=1}^{M} a_{-k}^{d\Psi^\top}(\mathfrak{x}; \nu) \partial_x^{-k} \widehat{q} + \mathcal{R}_M^{(2)}(\mathfrak{x}; \nu)[\widehat{q}\,], \\ &\mathcal{R}_M^{(2)}(\mathfrak{x}; \nu)[\widehat{q}\,] := \Pi_\perp \sum_{k=1}^{M} (\partial_x^{-k} w)\, \mathcal{A}_{-k}^{d\Psi^\top}(\mathfrak{x}; \nu)[\widehat{q}\,] + \mathcal{R}_M^{d\Psi^\top}(\mathfrak{x}; \nu)[\widehat{q}\,]. \end{aligned} \tag{3.29}$$

For $\sigma_M \geqq M+1$, the map $\mathcal{R}_M^{(2)} : \mathcal{V}^{\sigma_M}(\delta) \times \Xi \to \mathcal{B}(H_0^1(\mathbb{T}_1), H_\perp^{M+2}(\mathbb{T}_1))$ is $\mathcal{C}^\infty$ and by Theorem 3.2-(**Est2**), satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$ and $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M^{(2)}(\mathfrak{x}; \nu)[\widehat{q}\,]\|_{H_x^{M+2}} \lesssim_{M,\alpha} \|\widehat{q}\|_{H_x^1}, \\ &\|\partial_\nu^\alpha d^l \mathcal{R}_M^{(2)}(\mathfrak{x}; \nu)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l][\widehat{q}\,]\|_{H_x^{M+2}} \lesssim_{M,l,\alpha} \|\widehat{q}\|_{H_x^1} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}. \end{aligned} \tag{3.30}$$

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 (3.29), for any $\widehat{w} \in H_\perp^{-1}(\mathbb{T}_1)$, one has

$$d_\perp \Psi_\nu(\mathfrak{x})[\widehat{w}] = \widehat{w} + \sum_{k=1}^{M} (-1)^k \partial_x^{-k}\big(a_{-k}^{d\Psi^\top}(\mathfrak{x};\nu)\widehat{w}\big) + \mathcal{R}_M^{(2)}(\mathfrak{x};\nu)^\top[\widehat{w}]. \quad (3.31)$$

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 (3.31) 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 (2.12), the expansion (3.31) yields one of the form (3.25) where by (3.30) and Theorem 3.2-(**Est2**), the remainder $\mathcal{R}_M^{ext}(\mathfrak{x};\nu; d\Psi^\top)$ satisfies for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, and $\widehat{w} \in H_0^{-M-2}(\mathbb{T}_1)$

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x};\nu; d_\perp\Psi)[\widehat{w}]\|_{H_x^{-1}} \lesssim_{M,\alpha} \|\widehat{w}\|_{H_x^{-M-2}}, \\ &\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x};\nu; d_\perp\Psi)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l][\widehat{w}]\|_{H_x^{-1}} \lesssim_{M,l,\alpha} \|\widehat{w}\|_{H_x^{-M-2}} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}. \end{aligned} \quad (3.32)$$

The restriction of the expansion (3.31) to $L_\perp^2(\mathbb{T}_1)$ coincides with the one of $d_\perp \Psi_\nu(\mathfrak{x})[\widehat{w}]$, obtained by differentiating (3.9) (see (3.20)). It then follows from the uniqueness of an expansion of this form that

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

The claimed estimates (3.26) and (3.28) thus follow by Theorem 3.2-(**Est1**). In particular, for any $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l \in E_{\sigma_M}$, and $\widehat{w} \in L_\perp^2(\mathbb{T}_1)$,

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M^{ext}(\mathfrak{x};\nu; d_\perp\Psi)[\widehat{w}]\|_{H_x^{M+1}} \lesssim_{M,\alpha} \|\widehat{w}\|_{L_x^2}, \\ &\|\partial_\nu^\alpha d^l \mathcal{R}_M^{ext}(\mathfrak{x};\nu; d_\perp\Psi)[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l][\widehat{w}]\|_{H_x^{M+1}} \lesssim_{M,l,\alpha} \|\widehat{w}\|_{L_x^2} \prod_{j=1}^{l} \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}. \end{aligned} \quad (3.33)$$

The claimed estimates (3.27) are then obtained by interpolating between (3.32) and (3.33). □

3.3. Expansions of Linearized Hamiltonian Vector Fields

In this section, we apply Theorem 3.2 and Corollaries 3.3–3.4, to obtain asymptotic expansions of the linearized Hamiltonian vector fields $\partial_x d_\perp \nabla_w \mathcal{P}$ and $\partial_x d_\perp \nabla_w \mathcal{H}^{kdv}$. These expansions are key for implementing the KAM reduction procedure of Sections 6–7 to obtain an approximate right inverse of $\omega \cdot \partial_\varphi - dX_{\mathcal{H}_\varepsilon}$.

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

$$f : \mathbb{T}_1 \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}, \ (x, \zeta_0, \zeta_1) \mapsto f(x, \zeta_0, \zeta_1), \tag{3.34}$$

define

$$\mathcal{P} := P \circ \Psi_\nu, \quad \mathcal{P}(\mathfrak{x}; \nu) := P(\Psi_\nu(\mathfrak{x})), \quad \mathfrak{x} = (\theta, y, w), \tag{3.35}$$

where Ψ_ν is the coordinate transformation of Theorem 3.2. As a first result, we provide an expansion of the linearized Hamiltonian vector field $\partial_x d_\perp \nabla_w \mathcal{P}$.

Lemma 3.5. (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* $\sigma_M \geq M + 1$ *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*

$$\begin{aligned} \partial_x d_\perp \nabla_w \mathcal{P}(\mathfrak{x}; \nu)[\cdot] \ &= \Pi_\perp \sum_{k=0}^{M+3} a_{3-k}(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})\, \partial_x^{3-k}[\cdot] \\ &\quad + \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})[\cdot] \end{aligned} \tag{3.36}$$

with the following properties:

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

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

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

$$\|\partial_\nu^\alpha a_{3-k}(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})\|_{H^s_x} \lesssim_{s,M,\alpha} 1 + \|w\|_{H_x^{s+\sigma_M}}, \tag{3.37}$$

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

2. *For any* $0 \leqq s \leqq M + 1$, *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* $\widehat{w} \in H_\perp^{-s}(\mathbb{T}_1)$,

$$\|\partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})[\widehat{w}]\|_{H_x^{M+1-s}} \lesssim_{s,M,\alpha} \|\widehat{w}\|_{H_x^{-s}}, \tag{3.38}$$

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

3. *For any $s \geqq 0$, 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 $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l \in E_{s+\sigma_M}$, $(\mathfrak{x}, \nu) \in (\mathcal{V}^{\sigma_M}(\delta) \cap \mathcal{E}_{s+\sigma_M}) \times \Xi$, and $\widehat{w} \in H_\perp^s(\mathbb{T}_1)$,

$$\begin{aligned}
&\|\partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})[\widehat{w}]\|_{H_x^{s+M+1}} \lesssim_{s,M,\alpha} \|\widehat{w}\|_{H_x^s} + \|w\|_{H_x^{s+\sigma_M}} \|\widehat{w}\|_{L_x^2}, \\
&\|\partial_\nu^\alpha d^l \big(\mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{P})[\widehat{w}]\big)[\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l]\|_{H_x^{s+M+1}} \lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H_x^s} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} \\
&\quad + \|\widehat{w}\|_{L_x^2} \big(\|w\|_{H_x^{s+\sigma_M}} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} + \sum_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{i \neq j} \|\widehat{\mathfrak{x}}_i\|_{E_{\sigma_M}}\big). \quad (3.39)
\end{aligned}$$

Remark 3.6. The coefficient a_3 in (3.36) 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 (3.35) we have that

$$\nabla \mathcal{P}(\mathfrak{x}; \nu) = (d\Psi_\nu(\mathfrak{x}))^\top \big[\nabla P(\Psi_\nu(\mathfrak{x}))\big], \tag{3.40}$$

where by (3.34),

$$\nabla P(u) = \Pi_0^\perp \big[(\partial_{\zeta_0} f)(x, u, u_x) - \big((\partial_{\zeta_1} f)(x, u, u_x)\big)_x\big] \tag{3.41}$$

and $\Pi_0^\perp$ is the L^2-orthogonal projector of $L^2(\mathbb{T}_1)$ onto $L_0^2(\mathbb{T}_1)$. By (3.40), 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

$$\begin{aligned}
d_\perp \nabla_w \mathcal{P}(\mathfrak{x}; \nu)[\widehat{w}] &= (d_\perp \Psi_\nu(\mathfrak{x}))^\top \big[d\nabla P(\Psi_\nu(\mathfrak{x}))\big[d_\perp \Psi_\nu(\mathfrak{x})[\widehat{w}]\big]\big] \\
&\quad + \big(d_\perp (d_\perp \Psi_\nu(\mathfrak{x}))^\top [\widehat{w}]\big)\big[\nabla P(\Psi_\nu(\mathfrak{x}))\big].
\end{aligned} \tag{3.42}$$

Analysis of the first term on the right hand side of (3.42): Evaluating the differential $d\nabla P(u)$ of (3.41) at $u = \Psi_\nu(\mathfrak{x})$, one gets

$$\begin{aligned}
&\mathrm{d}(\nabla P)(\Psi_\nu(\mathfrak{x}))[h] = \Pi_0^\perp \big(b_2(\mathfrak{x}; \nu)\partial_x^2 h + b_1(\mathfrak{x}; \nu)\partial_x h + b_0(\mathfrak{x}; \nu) h\big), \\
&b_2(\mathfrak{x}; \nu) := -\,\partial_{\zeta_1}^2 f(x, u, u_x)\Big|_{u=\Psi_\nu(\mathfrak{x})}, \quad b_1(\mathfrak{x}; \nu) := (b_2(\mathfrak{x}; \nu))_x, \\
&b_0(\mathfrak{x}; \nu) := \big((\partial_{\zeta_0}^2 f)(x, u, u_x) - \big((\partial_{\zeta_0 \zeta_1}^2 f)(x, u, u_x)\big)_x\big)\big|_{u=\Psi_\nu(\mathfrak{x})}.
\end{aligned} \tag{3.43}$$

By Lemma 2.3 and Theorem 3.2 one infers that for any $s \geqq 0$, 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 $\alpha \in \mathbb{N}^{\mathbb{S}_+}$, $\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l \in E_{s+3}$, and $(\mathfrak{x}, \nu) \in \big(\mathcal{V}^3(\delta) \cap \mathcal{E}_{s+3}\big) \times \Xi$,

$$\begin{aligned}
&\|\partial_\nu^\alpha b_i(\mathfrak{x}; \nu)\|_{H_x^s} \lesssim_{s,\alpha} 1 + \|w\|_{H_x^{s+3}}, \\
&\|\partial_\nu^\alpha d^l b_i(\mathfrak{x}; \nu)[\widehat{\mathfrak{x}}_1, \dots, \widehat{\mathfrak{x}}_l]\|_{H_x^s} \lesssim_{s,l,\alpha} \sum_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{s+3}} \prod_{i \neq j} \|\widehat{\mathfrak{x}}_i\|_{E_3} + \|w\|_{H_x^{s+3}} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_3}.
\end{aligned} \tag{3.44}$$

By Corollary 3.3 (expansion of $(d_\perp\Psi_\nu)^\top$), Corollary 3.4 (expansion of $d_\perp\Psi_\nu$), (3.44) (estimates of b_i), (3.43) (formula for $d(\nabla P)(\Psi_\nu(\mathfrak{x}))$), and Lemma 2.12 (composition), one obtains the expansion

$$\begin{aligned}\partial_x (d_\perp\Psi_\nu(\mathfrak{x}))^\top\big[\mathrm{d}\nabla P(\Psi_\nu(\mathfrak{x}))\big[d_\perp\Psi_\nu(\mathfrak{x})[\cdot]\big]\big]\\ = \Pi_\perp \sum_{k=0}^{M+3} a^{(1)}_{3-k}(\mathfrak{x};\nu)\partial_x^{3-k} + R_1(\mathfrak{x};\nu),\end{aligned} \tag{3.45}$$

where $a^{(1)}_3(\mathfrak{x};\nu) = b_2(\mathfrak{x};\nu)$, 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 (3.37)–(3.39).

Analysis of the second term on the right hand side of (3.42): Since $d\Psi_\nu(\mathfrak{x})$ is symplectic, $d\Psi_\nu(\mathfrak{x})^\top = \mathcal{J}^{-1}\mathrm{d}\Psi_\nu(\mathfrak{x})^{-1}\partial_x$ where $\mathcal{J}$ is the Poisson operator defined in (3.7), implying that for any $\widehat{w} \in H^1_\perp(\mathbb{T}_1)$,

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

By this identity we get

$$\begin{aligned}&\partial_x\big(d_\perp(d_\perp\Psi_\nu(\mathfrak{x}))^\top[\cdot]\big)\big[\nabla P(\Psi_\nu(\mathfrak{x}))\big]\\ &\quad = -\partial_x d\Psi_\nu(\mathfrak{x})^\top\partial_x^{-1}\mathrm{d}\big(d_\perp\Psi_\nu(\mathfrak{x})[\cdot]\big)\big[\mathcal{J}\mathrm{d}\Psi_\nu(\mathfrak{x})^\top\nabla P(\Psi_\nu(\mathfrak{x}))\big].\end{aligned} \tag{3.46}$$

Arguing as for the first term on the right hand side of (3.42) (cf. (3.45)) one gets an expansion of the form

$$\partial_x\big(d_\perp(d_\perp\Psi_\nu(\mathfrak{x}))^\top[\cdot]\big)\big[\nabla P(\Psi_\nu(\mathfrak{x}))\big] = \Pi_\perp\sum_{k=3}^{M+3} a^{(2)}_{3-k}(\mathfrak{x};\nu)\partial_x^{3-k} + R_2(\mathfrak{x};\nu), \tag{3.47}$$

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 (3.37)–(3.39).

Conclusion: By (3.42) and the above analysis of the expansions (3.45) and (3.47), the lemma and Remark 3.6 follow. □

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 $\mathcal{H}^{kdv}(\mathfrak{x};\nu) = H^{kdv}(\Psi_\nu(\mathfrak{x}))$ (cf. Theorem 3.2-(**AE3**)). We recall that the family of Fourier multipliers $\Omega^{kdv}(D;\nu)$, $\nu\in\Xi$, is defined in (3.13).

Lemma 3.7. (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*

$$\begin{aligned}&\partial_x d_\perp\nabla_w\mathcal{H}^{kdv}(\mathfrak{x};\nu)[\cdot] = \partial_x\Omega^{kdv}(D;\nu)[\cdot] + \partial_x d_\perp\nabla_w\mathcal{R}^{kdv}(\mathfrak{x};\nu)[\cdot],\\ &\partial_x d_\perp\nabla_w\mathcal{R}^{kdv}(\mathfrak{x};\nu)[\cdot] = \Pi_\perp\sum_{k=0}^{M+1} a_{1-k}(\mathfrak{x};\nu;\partial_x d_\perp\nabla_w\mathcal{R}^{kdv})\,\partial_x^{1-k}[\cdot]\\ &\qquad + \mathcal{R}_M(\mathfrak{x};\nu;\partial_x d_\perp\nabla_w\mathcal{R}^{kdv})[\cdot],\end{aligned} \tag{3.48}$$

with the following properties:

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

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

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

$$\begin{aligned} &\|\partial_\nu^\alpha a_{1-k}(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{R}^{kdv})\|_{H^s_x} \lesssim_{s,k,\alpha} \|y\| + \|w\|_{H_x^{s+\sigma_M}}, \\ &\|d^l \partial_\nu^\alpha a_{1-k}(\mathfrak{x}; \nu; \partial_x d_\perp \nabla_w \mathcal{R}^{kdv})[\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H^s_x} \lesssim_{s,k,l,\alpha} \textstyle\sum_{j=1}^l \big(\|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{n \neq j} \|\widehat{\mathfrak{x}}_n\|_{E_{\sigma_M}}\big) \\ &\qquad + (\|y\| + \|w\|_{H_x^{s+\sigma_M}}) \textstyle\prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}. \end{aligned} \tag{3.49}$$

2. *For any* $0 \leqq s \leqq M+1$*, 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* $\widehat{w} \in H_\perp^{-s}(\mathbb{T}_1)$*,*

$$\|\partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla \mathcal{R}^{kdv})[\widehat{w}]\|_{H_x^{M+1-s}} \lesssim_{s,M,\alpha} (\|y\| + \|w\|_{H_x^{\sigma_M}}) \|\widehat{w}\|_{H_x^{-s}}, \tag{3.50}$$

$$\|d^l \partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla \mathcal{R}^{kdv})[\widehat{w}][\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{M+1-s}} \lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H_x^{-s}} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}. \tag{3.51}$$

3. *For any* $s \geqq 0$*, 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* $\widehat{w} \in H_\perp^s(\mathbb{T}_1)$*,*

$$\begin{aligned} &\|\partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla \mathcal{R}^{kdv})[\widehat{w}]\|_{H_x^{s+M+1}} \\ &\lesssim_{s,M,\alpha} (\|y\| + \|w\|_{H_x^{s+\sigma_M}}) \|\widehat{w}\|_{L_x^2} + (\|y\| + \|w\|_{H_x^{\sigma_M}}) \|\widehat{w}\|_{H_x^s}, \end{aligned} \tag{3.52}$$

$$\begin{aligned} &\|d^l \partial_\nu^\alpha \mathcal{R}_M(\mathfrak{x}; \nu; \partial_x d_\perp \nabla \mathcal{R}^{kdv})[\widehat{w}][\widehat{\mathfrak{x}}_1, \ldots, \widehat{\mathfrak{x}}_l]\|_{H_x^{s+M+1}} \lesssim_{s,M,l,\alpha} \|\widehat{w}\|_{H_x^s} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}} \\ &\quad + \|\widehat{w}\|_{L_x^2} \sum_{j=1}^l \Big(\|\widehat{\mathfrak{x}}_j\|_{E_{s+\sigma_M}} \prod_{n \neq j} \|\widehat{\mathfrak{x}}_n\|_{E_{\sigma_M}}\Big) + \|\widehat{w}\|_{L_x^2} \|w\|_{H_x^{s+\sigma_M}} \prod_{j=1}^l \|\widehat{\mathfrak{x}}_j\|_{E_{\sigma_M}}. \end{aligned} \tag{3.53}$$

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

$$\nabla_w \mathcal{H}^{kdv}(\mathfrak{x}; \nu) = (d_\perp \Psi_\nu(\mathfrak{x}))^\top \big[\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big] \tag{3.54}$$

where by (1.2),

$$\nabla H^{kdv}(u) = \Pi_0^{\perp}\big(3u^2 - u_{xx}\big) \tag{3.55}$$

and where $\Pi_0^{\perp}$ is the L^2-orthogonal projector onto $L_0^2(\mathbb{T}_1)$. Differentiating (3.54) with respect to w in direction $\widehat{w}$ we get

$$\begin{aligned} d_{\perp}\nabla_w \mathcal{H}^{kdv}(\mathfrak{x};\nu)[\widehat{w}] &= (d_{\perp}\Psi_\nu(\mathfrak{x}))^{\top}\big[\mathrm{d}\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))[d_{\perp}\Psi_\nu(\mathfrak{x})[\widehat{w}]]\big] \\ &\quad + \big(d_{\perp}(d_{\perp}\Psi_\nu(\mathfrak{x}))^{\top}[\widehat{w}]\big)\big[\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big]. \end{aligned} \tag{3.56}$$

On the other hand, by (3.12)

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

and by (3.15) $d_{\perp}\nabla_w \mathcal{R}^{kdv}(\theta,0,0;\nu) = 0$, implying that

$$\begin{aligned} &d_{\perp}\nabla_w \mathcal{H}^{kdv}(\theta,0,0;\nu) = \Omega^{kdv}(D;\nu), \\ &d_{\perp}\nabla_w \mathcal{R}^{kdv}(\mathfrak{x};\nu) = d_{\perp}\nabla_w \mathcal{H}^{kdv}(\theta,y,w;\nu) - d_{\perp}\nabla_w \mathcal{H}^{kdv}(\theta,0,0;\nu). \end{aligned} \tag{3.57}$$

In order to obtain the expansion (3.48) 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 (3.56).
Analysis of the first term on the right hand side of (3.56): Evaluating the differential $d\nabla H^{kdv}(u)$ at $u = \Psi_\nu(\mathfrak{x})$, one gets

$$\mathrm{d}(\nabla H^{kdv})(\Psi_\nu(\mathfrak{x}))[h] = \Pi_0^{\perp}\big(-\partial_x^2 h + b_0(\mathfrak{x};\nu)h\big), \quad b_0(\mathfrak{x};\nu) := 6\Psi_\nu(\mathfrak{x}). \tag{3.58}$$

By Theorem 3.2-(**AE1**) and the estimates (**Est1**), the function $b_0(\mathfrak{x};\nu)$ satisfies, for any $s \geqq 0$,

$$\begin{aligned} &\|\partial_\nu^\alpha b_0(\mathfrak{x};\nu)\|_{H_x^s} \lesssim_{s,\alpha} 1 + \|w\|_{H_x^{s+1}}, \\ &\|\partial_\nu^\alpha d^l b_0(\mathfrak{x};\nu)[\widehat{\mathfrak{x}}_1,\ldots,\widehat{\mathfrak{x}}_l]\|_{H_x^s} \lesssim_{s,l,\alpha} \sum_{j=1}^{l}\|\widehat{\mathfrak{x}}_j\|_{E_{s+1}}\prod_{i\neq j}\|\widehat{\mathfrak{x}}_i\|_{E_1} + \|w\|_{H_x^{s+1}}\prod_{j=1}^{l}\|\widehat{\mathfrak{x}}_j\|_{E_1}. \end{aligned} \tag{3.59}$$

By Corollary 3.3 (expansion of $(d_{\perp}\Psi_\nu)^{\top}$), Corollary 3.4 (expansion of $d_{\perp}\Psi_\nu$), (3.59) (estimates of b_0), (3.58) (formula for $d(\nabla H^{kdv})(\Psi_\nu(\mathfrak{x}))$), and Lemma 2.12 (composition), one obtains the expansion

$$\begin{aligned} &\partial_x(d_{\perp}\Psi_\nu(\mathfrak{x}))^{\top}\big[d\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big[d_{\perp}\Psi_\nu(\mathfrak{x})[\cdot]\big]\big] \\ &= \Pi_{\perp}\Big(-\partial_x^3 - (a_{-1}^{\Psi}(\mathfrak{x};\nu) + a_{-1}^{d\Psi^{\top}}(\mathfrak{x};\nu))\partial_x^2 + \sum_{k=0}^{M+1} a_{1-k}^{(1)}(\mathfrak{x};\nu)\partial_x^{1-k}\Big) + R_1(\mathfrak{x};\nu) \\ &\overset{(3.11)}{=} \Pi_{\perp}\Big(-\partial_x^3 + \sum_{k=0}^{M+1} a_{1-k}^{(1)}(\mathfrak{x};\nu)\partial_x^{1-k}\Big) + R_1(\mathfrak{x};\nu), \end{aligned} \tag{3.60}$$

where the functions $a_{1-k}^{(1)}(\mathfrak{x};\nu)$, $k = 0,\ldots,M+1$ and the remainder $R_1(\mathfrak{x};\nu)$ satisfy the properties 1–3 stated in Lemma 3.5, in particular (3.37)–(3.39).

Analysis of the second term on the right hand side of (3.56): By (3.46) one has

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

Arguing as for the first term on the right hand side of (3.56) one obtains an expansion of the form

$$\partial_x \big(d_\perp (d_\perp \Psi_\nu(\mathfrak{x}))^\top[\cdot]\big)\big[\nabla H^{kdv}(\Psi_\nu(\mathfrak{x}))\big] = \Pi_\perp \sum_{k=0}^{M+1} a^{(2)}_{1-k}(\mathfrak{x};\nu)\partial_x^{1-k} + R_2(\mathfrak{x};\nu), \tag{3.61}$$

where $a^{(2)}_1(\mathfrak{x};\nu) = 0$ (cf. (3.16)) and where the functions $a^{(2)}_{1-k}(\mathfrak{x};\nu)$, $k = 1, \dots, M+1$ and the remainder $R_2(\mathfrak{x};\nu)$ satisfy the properties 1-3 of Lemma 3.5, in particular (3.37)–(3.39).

Conclusion: Combining (3.56), (3.57), (3.60), and (3.61) one obtains the claimed expansion (3.48) with

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

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 3.5, in particular (3.37)–(3.39), the claimed estimates (3.49)–(3.53) then follow by the mean value theorem. □

3.4. *Frequencies of KdV*

In this section we record properties of the KdV frequencies ω_n^{kdv}, used in particular for the measure estimates in Section 8.2, and discuss an expansion of $\partial_x \Omega^{kdv}(D;\nu)$ (cf. (3.13)) needed in Section 6.

Recall that the family of operators $\Omega^{kdv}(D;\nu)$, introduced in (3.13), is defined for $\nu \in \Xi \subset \mathbb{R}^{\mathbb{S}_+}_{>0}$. Actually, it is defined on all of $\mathbb{R}^{\mathbb{S}_+}_{>0}$ (cf. (3.2)) and according to [20, Lemma 4.1], $\partial_x \Omega^{kdv}(D;I)$ can be written as

$$\partial_x \Omega^{kdv}(D;I) = -\partial_x^3 + Q^{kdv}_{-1}(D;I), \tag{3.62}$$

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 3.8. *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*

$$Q^{kdv}_{-1}(D;I) = \Omega^{kdv}_{-1}(D;I) + \mathcal{R}_M(D;I;Q^{kdv}_{-1}),$$
$$\Omega^{kdv}_{-1}(\xi;I) = \sum_{k=1}^{M} a_{-k}(I;\Omega^{kdv}_{-1})\chi_0(\xi)(\mathrm{i}2\pi\xi)^{-k}, \tag{3.63}$$

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

$$\mathcal{R}_M(n; I; Q_{-1}^{kdv}) = \frac{\mathcal{R}_M^{\omega_n}(I)}{(2\pi n)^{M+1}}, \quad \mathcal{R}_M(-n; I; Q_{-1}^{kdv}) = -\mathcal{R}_M(n; I; Q_{-1}^{kdv}), \quad \forall n \in \mathbb{S}_+^{\perp}, \tag{3.64}$$

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

$$\sup_{n \in \mathbb{S}^{\perp}} |\mathcal{R}_M^{\omega_n}(I)| \leqq C_M, \quad \sup_{n \in \mathbb{S}^{\perp}} |\partial_{I_j}^{\beta} \mathcal{R}_M^{\omega_n}(I)| \leqq C_{M,\beta}, \tag{3.65}$$

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

Proof. The result follows from [20, Lemma C.7]. □

Lemma 3.9. (Non-degeneracy of KdV frequencies, [21, Proposition 15.5]) *For any finite subset $\mathbb{S}_+ \subset \mathbb{N}$ the following holds on $\mathbb{R}_{>0}^{\mathbb{S}_+}$:*

(i) *The map $I \mapsto \det\big((\partial_{I_k} \omega_n^{kdv}(I, 0))_{k,n \in \mathbb{S}_+}\big)$ 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:*

$$\sum_{n \in \mathbb{S}_+} \ell_n \omega_n^{kdv} + \omega_j^{kdv} \neq 0, \quad \sum_{n \in \mathbb{S}_+} \ell_n \omega_n^{kdv} + \omega_j^{kdv} - \omega_k^{kdv} \neq 0. \tag{3.66}$$

Remark 3.10. It was shown in [12] that for any $I \in \mathbb{R}_{>0}^{\mathbb{S}_+}$, $\det\big((\partial_{I_k} \omega_n^{kdv}(I, 0))_{k,n \in \mathbb{S}_+}\big) \neq 0$.

Finally, we record the following asymptotics of the KdV frequencies, used in Section 7,

$$\omega_n^{kdv}(I, 0) - (2\pi n)^3 = O(n^{-1}), \quad n\, \partial_I \omega_n^{kdv}(I, 0) = O(1), \tag{3.67}$$

uniformly on compact sets of actions $I \in \mathbb{R}_{>0}^{\mathbb{S}_+}$ (cf. for example [22, Proposition 8.1]).

4. Nash–Moser Theorem

The purpose of this short section is to state Theorem 4.1 which reformulates Theorem 1.1 in the canonical coordinates described in the previous section. In Section 8.3, we derive Theorem 1.1 from Theorem 4.1.

In the canonical coordinates $\mathfrak{x} = (\theta, y, w) \in \mathcal{V}(\delta) \cap \mathcal{E}_s$, defined by Theorem 3.2, with symplectic 2-form given by (3.6), the Hamiltonian equation (1.6) reads as

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

where $\mathcal{H}_\varepsilon := H_\varepsilon \circ \Psi_\nu$ and H_ε is given by (1.7). More explicitly,

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

where $\mathcal{H}^{kdv}(\mathfrak{x};\nu)$ has the normal form expansion (3.12). 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 $\omega^{kdv}(\nu) := (\omega_n^{kdv}(\nu, 0))_{n\in\mathbb{S}_+}$ 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

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

To keep the notation simpler, 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,

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

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 $\iota(\varphi) = (\Theta(\varphi), y(\varphi), w(\varphi))$ and where $\Theta(\varphi) := \theta(\varphi) - \varphi$ is $(2\pi\mathbb{Z})^{\mathbb{S}_+}$ periodic. Often we will refer to the torus embedding $\breve{\iota}$ simply as torus. 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, in particular (5.35). 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 $X_{\mathcal{H}_\varepsilon}$ due to (5.5). Note 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 defined by

$$\|\iota\|_s^{\mathrm{Lip}(\gamma)} := \|\Theta\|_s^{\mathrm{Lip}(\gamma)} + \|y\|_s^{\mathrm{Lip}(\gamma)} + \|w\|_s^{\mathrm{Lip}(\gamma)} \tag{4.8}$$

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

$$\|\Theta\|_s^{\mathrm{Lip}(\gamma)} \equiv \|\Theta\|_{H^s_\varphi}^{\mathrm{Lip}(\gamma)} := \|\Theta\|_{H^s(\mathbb{T}^{\mathbb{S}_+},\mathbb{R}^{\mathbb{S}_+})}^{\mathrm{Lip}(\gamma)}, \quad \|y\|_s^{\mathrm{Lip}(\gamma)} \equiv \|y\|_{H^s_\varphi}^{\mathrm{Lip}(\gamma)} := \|y\|_{H^s(\mathbb{T}^{\mathbb{S}_+},\mathbb{R}^{\mathbb{S}_+})}^{\mathrm{Lip}(\gamma)}. \tag{4.9}$$

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

$$\lim_{\varepsilon \to 0} \frac{\text{meas}(\Omega_\varepsilon)}{\text{meas}(\Omega)} = 1 \tag{4.10}$$

and for any $\omega \in \Omega_\varepsilon$, *there exists a torus embedding with lift* $\breve{\iota}_\omega : \mathbb{R}^{\mathbb{S}_+} \to \mathbb{R}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times H_\perp^{\bar{s}}(\mathbb{T}_1)$ *(cf. (4.5)) which satisfies the estimate*

$$\|\breve{\iota}_\omega - (\varphi, 0, 0)\|_{\bar{s}}^{\text{Lip}(\gamma)} = O(\varepsilon\gamma^{-2}), \quad \gamma = \varepsilon^{\mathfrak{a}}, \ 0 < \mathfrak{a} \ll 1, \tag{4.11}$$

and solves

$$\omega \cdot \partial_\varphi \breve{\iota}_\omega(\varphi) - X_{\mathcal{H}_\varepsilon(\cdot;\nu)}(\breve{\iota}_\omega(\varphi)) = 0, \qquad \nu = (\omega^{kdv})^{-1}(-\omega). \tag{4.12}$$

As a consequence, the embedded torus $\breve{\iota}_\omega(\mathbb{T}^{\mathbb{S}_+})$ *is invariant under the flow of the Hamiltonian vector field* $X_{\mathcal{H}_\varepsilon(\cdot;\nu)}$ *and is filled by quasi-periodic solutions of* (4.1) *with frequency vector* $\omega \in \Omega_\varepsilon$. *Furthermore, the quasi-periodic solution* $\breve{\iota}_\omega(\omega t) = (\omega t, 0, 0) + \iota_\omega(\omega t)$ *is linearly stable.*

Remark 4.2. Up to the end of Section 7, $\gamma \in (0, 1)$ is assumed to be a constant independent of ε with $\varepsilon\gamma^{-3} \ll 1$. Only in Section 8, γ and ε are required to be related by $\gamma = \varepsilon^{\mathfrak{a}}$ for some $0 < \mathfrak{a} \ll 1$.

Theorem 4.1 is proved in Section 8 and is applied to deduce Theorem 1.1 (cf. Section 8.3 for details). At the core of the proof of Theorem 4.1 is 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 5–7.

Along the proof we shall use the following tame estimates of the Hamiltonian vector field $X_{\mathcal{H}_\varepsilon}$ with respect to the norm $\|\cdot\|_s^{\text{Lip}(\gamma)}$ in (4.8). Using the expansion (3.12) provided in Theorem 3.2, and the definition of $\mathcal{P}$ in (3.35), we decompose the Hamiltonian $\mathcal{H}_\varepsilon$, defined in (4.2), as

$$\begin{aligned} \mathcal{H}_\varepsilon &= \mathcal{N} + \mathcal{P}_\varepsilon \quad \text{where} \\ \mathcal{N}(y, w; \nu) &:= \omega^{kdv}(\nu) \cdot y + \frac{1}{2}\Omega_{\mathbb{S}_+}^{kdv}(\nu)[y] \cdot y + \frac{1}{2}\big(\Omega^{kdv}(D; \nu)w, \ w\big)_{L_x^2}, \\ \mathcal{P}_\varepsilon &:= \mathcal{R}^{kdv} + \varepsilon\mathcal{P}. \end{aligned} \tag{4.13}$$

The Hamiltonian vector field of $\mathcal{P}_\varepsilon$ and $\mathcal{H}_\varepsilon$ satisfy the following tame estimates:

Lemma 4.3. *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* $\|\iota\|_{s_0+\sigma_1}^{\text{Lip}(\gamma)} \leqq \delta$, *and any maps* $\widehat{\iota}, \widehat{\iota}_1, \widehat{\iota}_2 : \mathbb{T}^{\mathbb{S}_+} \to E_s$, *the following tame estimates hold:*

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

Proof. By (4.13), one has $X_{\mathcal{P}_\varepsilon} = \varepsilon X_{\mathcal{P}} + X_{\mathcal{R}^{kdv}}$ and $d^2 X_{\mathcal{H}_\varepsilon} = d^2 X_{\mathcal{N}} + d^2 X_{\mathcal{P}_\varepsilon}$. The claimed estimate for $dX_{\mathcal{P}_\varepsilon}(\breve{\iota})[\widehat{\iota}]$ follows noting that, by Lemmata 3.5, 2.25, 2.26, $\|\mathrm{d}X_{\mathcal{P}}(\breve{\iota})[\widehat{\iota}]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\widehat{\iota}\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)}$ and, by Lemmata 3.7, 2.25, 2.26,

$$\|\mathrm{d}X_{\mathcal{R}^{kdv}}(\breve{\iota})[\widehat{\iota}]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\iota\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)} \|\widehat{\iota}\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)}.$$

The one for $\|X_{\mathcal{P}_\varepsilon}(\breve{\iota})\|_s^{\mathrm{Lip}(\gamma)}$ is obtained by the mean value theorem. Indeed, one has $\|X_{\mathcal{P}}(\breve{\iota})\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s 1 + (\|\iota\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)} + \|\iota\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)} \|\iota\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)})$, and taking into account (3.15), $\|X_{\mathcal{R}^{kdv}}(\breve{\iota})\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\iota\|_{s_0+\sigma_1}^{\mathrm{Lip}(\gamma)} \|\iota\|_{s+\sigma_1}^{\mathrm{Lip}(\gamma)}$. Finally, the estimate for $d^2 X_{\mathcal{H}_\varepsilon}(\breve{\iota})[\widehat{\iota}_1, \widehat{\iota}_2]$ is verified using again Lemmata 3.5, 3.7, 2.25, 2.26. □

5. Approximate Inverse

In order to prove Theorem 4.1 we implement a Nash–Moser iteration scheme that leads to a solution of $\mathcal{F}_\omega(\iota, \zeta) = 0$ (cf. (4.6)). For this purpose we construct an *almost-approximate right inverse* 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 (4.13). Note that the perturbation $\mathcal{P}_\varepsilon$ and the differential $d_{\iota,\zeta}\mathcal{F}_\omega(\iota, \zeta)$ are independent of ζ. Thus, in the sequel, we often write $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$ instead of $d_{\iota,\zeta}\mathcal{F}_\omega(\iota, \zeta)$. The construction of an almost-approximate right inverse of $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$ is the main result of this section, stated in Theorem 5.7. It is proved under the assumption **A-I**, introduced below (cf. (5.29)–(5.32)). In Theorem 7.11, these assumptions are stated as a theorem, and its proof is given in Section 7.2.

Throughout this section we assume that the following ansatz holds:

- **Ansatz.** *The maps* $\omega \mapsto \iota(\omega) := \breve{\iota}(\varphi; \omega) - (\varphi, 0, 0)$, *and* $\omega \mapsto \zeta(\omega)$ *are Lipschitz continuous with respect to* $\omega \in \Omega$, *and for* $0 < \gamma < 1$

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

where Z, *referred to as error function, is defined by*

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

We already mention that at each step of the Nash–Moser scheme of Theorem 8.1, the above 'Ansatz' is shown to hold, with the constant μ_0 specified in Theorem 5.7, depending on $|\mathbb{S}_+|$ and τ (given in Section 8).

Let us first give an outline of the proof of Theorem 5.7. Since the θ-, y-, and w-components of the linear operator $d_\iota X_{\mathcal{H}_\varepsilon}(\breve{\iota})$ form a coupled system, it turns out to be difficult to invert the operator $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$ in (5.1). To overcome this difficulty, we use the approach developed in [3,8–10], consisting in transforming $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$ into approximately triangular form, see (5.33). Let us describe in broad terms how to achieve this: If the error function Z, defined in (5.3), vanishes, then the torus $\breve{\iota}$ is invariant for the Hamiltonian $\mathcal{H}_{\varepsilon,\zeta}$ defined in (4.7). Furthermore, by (5.5) below, also ζ vanishes in this case, implying that $\breve{\iota}$ is invariant for the Hamiltonian $\mathcal{H}_\varepsilon$. Hence the invariant torus $\breve{\iota}$ is isotropic (cf. [8]) (where $\breve{\iota}$ being isotropic means that the pullback of the symplectic form by $\breve{\iota}$ vanishes) and there exist symplectic coordinates in a neighborhood of this torus (cf. (5.16)) so that when expressed in these coordinates, the linearized equations form a triangular system as described in (5.33). In general, the torus $\breve{\iota}$ is only "approximately" invariant up to order $O(Z)$ and the linearized equations can only be approximately conjugated to a triangular system as in (5.33). Taking all this into account, we proceed as follows: given an approximately invariant torus $\breve{\iota}(\varphi) = (\theta(\varphi), y(\varphi), w(\varphi))$ satisfying (5.2), we first construct an isotropic torus $\breve{\iota}_\delta(\varphi) = (\theta(\varphi), y_\delta(\varphi), w(\varphi))$ which is close to $\breve{\iota}$ (cf. Lemma 5.4). Note that by (5.13), $\mathcal{F}(\breve{\iota}_\delta, \zeta)$ is also of the order $O(Z)$. Since $\breve{\iota}_\delta$ is isotropic, the diffeomorphism $(\phi, \eta, v) \mapsto (\theta, y, w) = G_\delta(\phi, \eta, v)$ defined in (5.16) is symplectic. In these coordinates, the torus $\breve{\iota}_\delta$ reads $\varphi \mapsto (\phi, \eta, v) = (\varphi, 0, 0)$, and the transformed Hamiltonian $\mathcal{K} := \mathcal{H}_{\varepsilon,\zeta} \circ G_\delta$ takes the form (5.18), where the terms $\partial_\phi \mathcal{K}_{00}$, $\mathcal{K}_{10} + \omega$, and $\mathcal{K}_{01}$ are of the order $O(Z)$ (cf. Lemma 5.5). Neglecting terms of the order $O(Z)$, the problem of finding an approximate right inverse of the operator $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$ is reduced to the task of inverting the operator $\mathbb{D}$ in (5.33). The system (5.34) is solved in a *triangular* fashion as follows. First we solve the second equation in the system (5.34), cf. (5.35)–(5.36). Then we solve the third equation in (5.34) using the assumption **A-I**, cf. (5.37). Finally, to determine $\widehat{\phi}$, we solve the equation (5.38), cf. (5.41)–(5.42). In conclusion, we prove that the operator (5.44) is an almost-approximate right inverse of the operator $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$ which satisfies tame estimates – see Theorem 5.7 for details.

We start our construction by noting that the 2-form $\mathcal{W}$ given by (3.6) is exact:

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

where Λ is the Liouville 1-form

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

The pullbacks $\breve{\iota}^*\Lambda$ and $\breve{\iota}^*\mathcal{W}$ of Λ and respectively $\mathcal{W}$ by a torus embedding $\breve{\iota}$ are related by $\breve{\iota}^*\mathcal{W} = \mathrm{d}\breve{\iota}^*\Lambda$. Recall that the embedding $\breve{\iota}$ is said to be isotropic if $\breve{\iota}^*\mathcal{W} = 0$.

First, we provide an estimate of ζ in terms of the error function $Z(\varphi)$ defined in (5.3). We recall that by the ansatz (5.2), $Z(\varphi)$ and ζ are Lipschitz continuous with respect to the parameter $\omega \in \Omega$.

Lemma 5.1. *One has*

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

Proof. We follow the arguments in [8, Lemma 3] and [3, Lemma 6.1]. Since the Hamiltonian $\mathcal{H}_\varepsilon$ is autonomous, the "restricted" action functional

$$G : \mathbb{T}^{\mathbb{S}_+} \to \mathbb{R}, \quad \psi \mapsto G(\psi) := \int_{\mathbb{T}^{\mathbb{S}_+}} \Big(- \Lambda_{\breve{\iota}^{(\psi)}}(\omega \cdot \partial_\varphi \breve{\iota}^{(\psi)}(\varphi)) - \mathcal{H}_\varepsilon(\breve{\iota}^{(\psi)}(\varphi))\Big)\, \mathrm{d}\varphi$$

is constant, where $\breve{\imath}^{(\psi)}(\varphi) := \breve{\imath}(\psi + \varphi)$ and $\Lambda_{\breve{\imath}(\psi+\varphi)}$ is the canonical one form Λ defined in (5.4), evaluated at $\breve{\imath}(\psi + \varphi)$. Using that $\partial_\psi G(0) = 0$, a direct calculation shows that ζ can be expressed in terms of $Z(\varphi) = (Z_\theta(\varphi), Z_y(\varphi), Z_w(\varphi))$ as

$$\zeta = \frac{1}{(2\pi)^{|\mathbb{S}_+|}} \int_{\mathbb{T}^{\mathbb{S}_+}} \Big([\partial_\varphi y(\varphi)]^\top Z_\theta(\varphi) - [\partial_\varphi \theta(\varphi)]^\top Z_y(\varphi) - [\partial_\varphi w(\varphi)]^\top \partial_x^{-1} Z_w(\varphi) \Big)\, \mathrm{d}\varphi.$$

The latter formula together with the tame estimates (2.5) and the ansatz (5.2) imply (5.5). □

For an approximately invariant torus embedding $\breve{\imath} = (\theta(\varphi), y(\varphi), w(\varphi))$, the 1-form

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

is only "approximately closed", in the sense that

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

is of the order $O(Z)$. Here $\underline{e}_k$, $k \in \mathbb{S}_+$, denotes the standard basis of $\mathbb{R}^{\mathbb{S}_+}$. More precisely, the following lemma holds:

Lemma 5.2. *Let $\omega \in \mathtt{DC}(\gamma, \tau)$ (cf. (4.4)). Then for any $k, j \in \mathbb{S}_+$, the coefficient A_{kj} in (5.7) satisfies, for some $\sigma = \sigma(\tau, \mathbb{S}_+) > 0$, $\forall\, s \geq s_0$,*

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

Remark 5.3. In the sequel the constant $\sigma = \sigma(\tau, \mathbb{S}_+)$, referred to as loss of derivatives, will be tacitly increased in the course of our arguments if needed.

Proof. For any $j, k \in \mathbb{S}_+$, the coefficient A_{kj} satisfies the identity $\omega \cdot \partial_\varphi A_{kj} = \mathcal{W}\big(\partial_\varphi Z(\varphi) \underline{e}_k, \partial_\varphi \breve{\imath}(\varphi) \underline{e}_j\big) + \mathcal{W}\big(\partial_\varphi \breve{\imath}(\varphi) \underline{e}_k, \partial_\varphi Z(\varphi) \underline{e}_j\big)$ (cf. [8, Lemma 5]). The estimate (5.8) then follows by (5.2) and (2.10). □

As in [3,8] we first modify the approximate torus $\breve{\imath}$ to obtain an isotropic torus $\breve{\imath}_\delta$ which is still approximately invariant. Let $\Delta_\varphi := \sum_{k \in \mathbb{S}_+} \partial_{\varphi_k}^2$.

Lemma 5.4. (Isotropic torus) *Let $\omega \in \mathtt{DC}(\gamma, \tau)$. The torus $\breve{\imath}_\delta(\varphi) := (\theta(\varphi), y_\delta(\varphi), w(\varphi))$ defined by*

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

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

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

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

$$\| d_\iota \iota_\delta [\widehat{\iota}] \|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \| \widehat{\iota} \|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \| \iota \|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \| \widehat{\iota} \|_{s_0}^{\mathrm{Lip}(\gamma)}, \tag{5.12}$$

$$\| \mathcal{F}_\omega(\iota_\delta, \zeta) \|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \gamma^{-1} \big(\| Z \|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \| \iota \|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \| Z \|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} \big). \tag{5.13}$$

Proof. The isotropy of the modified torus $\breve{\iota}_\delta$ is proved in [8]. By a standard Neumann series argument and using the ansatz (5.2), it follows that for any $s \geqq s_0$, there is a constant $C(s) > 0$ so that

$$\|[\partial_\varphi \theta]^{-\top} - \mathrm{Id}\|_s^{\mathrm{Lip}(\gamma)} \leqq C(s)\|\iota\|_{s+1}^{\mathrm{Lip}(\gamma)}, \quad \|[\partial_\varphi \theta]^{-\top}\|_s^{\mathrm{Lip}(\gamma)} \leqq 1 + C(s)\|\iota\|_{s+1}^{\mathrm{Lip}(\gamma)}. \tag{5.14}$$

Furthermore, by the estimate (5.8) for the coefficients A_{kj} and the definition (5.9) of ρ, one gets

$$\|\rho\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \gamma^{-1}\big(\|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\big). \tag{5.15}$$

The estimates (5.10), (5.11) then follow by using (5.14), (5.15), the interpolation estimate (2.5), the ansatz (5.2), $\varepsilon\gamma^{-2} \ll 1$, and the definition of a_k, $k \in \mathbb{S}_+$, in (5.6). The estimate (5.12) follows by similar arguments. To prove (5.13), it suffices to estimate $\mathcal{F}_\omega(\iota_\delta, \zeta) - \mathcal{F}_\omega(\iota, \zeta)$. One computes

$$\mathcal{F}_\omega(\iota_\delta, \zeta) - \mathcal{F}_\omega(\iota, \zeta) = \begin{pmatrix} \Omega_{\mathbb{S}_+}(\nu)[y_\delta - y] \\ \omega \cdot \partial_\varphi(y_\delta - y_0) \\ 0 \end{pmatrix} + X_{\mathcal{P}_\varepsilon}(\iota_\delta) - X_{\mathcal{P}_\varepsilon}(\iota).$$

The estimate (5.13) then follows by using the mean value theorem to bound $X_{\mathcal{P}_\varepsilon}(\iota_\delta) - X_{\mathcal{P}_\varepsilon}(\iota)$, together with Lemma 4.3 and by applying the estimate (5.11) on $y_\delta - y$ (using also 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 diffeomorphism $G_\delta : (\phi, \eta, v) \mapsto (\theta, y, w)$ of the phase space $\mathbb{T}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times L^2_\perp(\mathbb{T}_1)$, defined by

$$\begin{pmatrix} \theta \\ y \\ w \end{pmatrix} := G_\delta \begin{pmatrix} \phi \\ \eta \\ v \end{pmatrix} := \begin{matrix} \theta(\phi) \\ y_\delta(\phi) + [\partial_\phi \theta(\phi)]^{-\top}\eta - \big[(\partial_\theta \tilde{w})(\theta(\phi))\big]^\top (\partial_x^{-1} v) \\ w(\phi) + v \end{matrix} \tag{5.16}$$

where $\tilde{w} := w \circ \theta^{-1}$. Since $\breve{\iota}_\delta$ is an isotropic torus embedding (cf. Lemma 5.4), G_δ is symplectic by [8, Lemma 2]. In the new coordinates, $\breve{\iota}_\delta$ is the trivial embedded torus $\varphi \mapsto (\phi, \eta, v) = (\varphi, 0, 0)$ and the Hamiltonian vector field $X_{\mathcal{H}_{\varepsilon,\zeta}}$ (with $\mathcal{H}_{\varepsilon,\zeta}$ defined in (4.7)) is given by

$$X_{\mathcal{K}} = (dG_\delta)^{-1} X_{\mathcal{H}_{\varepsilon,\zeta}} \circ G_\delta, \quad \mathcal{K} \equiv \mathcal{K}_{\varepsilon,\zeta} := \mathcal{H}_{\varepsilon,\zeta} \circ G_\delta. \tag{5.17}$$

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

$$\begin{aligned} \mathcal{K}(\phi, \eta, v, \zeta) = {} & \theta(\phi) \cdot \zeta + \mathcal{K}_{00}(\phi) + \mathcal{K}_{10}(\phi) \cdot \eta + (\mathcal{K}_{01}(\phi), v)_{L^2_x} + \frac{1}{2}\mathcal{K}_{20}(\phi)\eta \cdot \eta \\ & + \big(\mathcal{K}_{11}(\phi)\eta, v\big)_{L^2_x} + \frac{1}{2}\big(\mathcal{K}_{02}(\phi)v, v\big)_{L^2_x} + \mathcal{K}_{\geqq 3}(\phi, \eta, v) \end{aligned} \tag{5.18}$$

where $\mathcal{K}_{\geqq 3}$ comprises all the terms which are at least cubic in the variables (η, v), and where $\mathcal{K}_{00}(\phi) \in \mathbb{R}$, $\mathcal{K}_{10}(\phi) \in \mathbb{R}^{\mathbb{S}_+}$, $\mathcal{K}_{01}(\phi) \in L^2_\perp(\mathbb{T}_1)$, $\mathcal{K}_{20}(\phi)$ 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 $\mathcal{K}_{00} = \text{const}$, $\mathcal{K}_{10} = -\omega$, $\mathcal{K}_{01} = 0$.

Denote by $\text{Id}_\perp$ 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

$$dG_\delta|_{(\varphi,0,0)} \begin{pmatrix} \widehat{\phi} \\ \widehat{\eta} \\ \widehat{v} \end{pmatrix} := \begin{pmatrix} \partial_\phi \theta(\varphi) & 0 & 0 \\ \partial_\phi y_\delta(\varphi) & [\partial_\phi \theta(\varphi)]^{-\top} & -[(\partial_\theta \tilde{w})(\theta(\varphi))]^\top \partial_x^{-1} \\ \partial_\phi w(\varphi) & 0 & \text{Id}_\perp \end{pmatrix} \begin{pmatrix} \widehat{\phi} \\ \widehat{\eta} \\ \widehat{v} \end{pmatrix}. \tag{5.19}$$

It approximately transforms the linearized operator $d_{\iota,\zeta}\mathcal{F}_\omega(\iota_\delta)$ (see the proof of Theorem 5.7) into the one obtained when the Hamiltonian system with Hamiltonian $\mathcal{K}_{\varepsilon,\zeta}$ (cf. (5.17)) is linearized at $(\phi, \eta, v) = (\varphi, 0, 0)$, differentiated also with respect to ζ, and ∂_t is replaced by $\omega \cdot \partial_\varphi$,

$$\begin{pmatrix} \widehat{\phi} \\ \widehat{\eta} \\ \widehat{v} \\ \widehat{\zeta} \end{pmatrix} \mapsto \begin{pmatrix} \omega \cdot \partial_\varphi \widehat{\phi} + \partial_\phi \mathcal{K}_{10}(\varphi)[\widehat{\phi}] + \mathcal{K}_{20}(\varphi)\widehat{\eta} + \mathcal{K}_{11}^\top(\varphi)\widehat{v} \\ \omega \cdot \partial_\varphi \widehat{\eta} - \big(\partial_\phi \theta(\varphi)\big)^\top [\widehat{\zeta}] - \partial_\phi \big(\partial_\phi \theta(\varphi)^\top [\zeta]\big)[\widehat{\phi}] - \partial_{\phi\phi}\mathcal{K}_{00}(\varphi)[\widehat{\phi}] - [\partial_\phi \mathcal{K}_{10}(\varphi)]^\top \widehat{\eta} - [\partial_\phi \mathcal{K}_{01}(\varphi)]^\top \widehat{v} \\ \omega \cdot \partial_\varphi \widehat{v} - \partial_x \{\partial_\phi \mathcal{K}_{01}(\varphi)[\widehat{\phi}] + \mathcal{K}_{11}(\varphi)\widehat{\eta} + \mathcal{K}_{02}(\varphi)\widehat{v}\} \end{pmatrix}. \tag{5.20}$$

Using (5.2) and (5.10), one shows as in [3] that the operator $\widehat{\iota} = (\widehat{\phi}, \widehat{\eta}, \widehat{v}) \mapsto dG_\delta[\widehat{\iota}]$ satisfies for any $s \geq s_0$

$$\begin{aligned} &\|dG_\delta(\varphi, 0, 0)[\widehat{\iota}]\|_s^{\text{Lip}(\gamma)}, \ \|dG_\delta(\varphi, 0, 0)^{-1}[\widehat{\iota}]\|_s^{\text{Lip}(\gamma)} \\ &\quad \lesssim_s \|\widehat{\iota}\|_s^{\text{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\text{Lip}(\gamma)} \|\widehat{\iota}\|_{s_0}^{\text{Lip}(\gamma)}, \end{aligned} \tag{5.21}$$

$$\begin{aligned} &\|d^2 G_\delta(\varphi, 0, 0)[\widehat{\iota}_1, \widehat{\iota}_2]\|_s^{\text{Lip}(\gamma)} \\ &\quad \lesssim_s \|\widehat{\iota}_1\|_s^{\text{Lip}(\gamma)} \|\widehat{\iota}_2\|_{s_0}^{\text{Lip}(\gamma)} + \|\widehat{\iota}_1\|_{s_0}^{\text{Lip}(\gamma)} \|\widehat{\iota}_2\|_s^{\text{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\text{Lip}(\gamma)} \|\widehat{\iota}_1\|_{s_0}^{\text{Lip}(\gamma)} \|\widehat{\iota}_2\|_{s_0}^{\text{Lip}(\gamma)}. \end{aligned} \tag{5.22}$$

The next lemma provides estimates for the coefficients of the Taylor expansion (5.18) of the Hamiltonian $\mathcal{K}$.

Lemma 5.5. *There exists $\sigma := \sigma(\tau, \mathbb{S}_+) > 0$ so that for any $s \geq s_0$*

$$\begin{aligned} &\|\partial_\phi \mathcal{K}_{00}\|_s^{\text{Lip}(\gamma)} + \|\mathcal{K}_{10} + \omega\|_s^{\text{Lip}(\gamma)} + \|\mathcal{K}_{01}\|_s^{\text{Lip}(\gamma)} \lesssim_s \gamma^{-1}\big(\|Z\|_{s+\sigma}^{\text{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\text{Lip}(\gamma)} \|Z\|_{s_0+\sigma}^{\text{Lip}(\gamma)}\big). \\ &\|\mathcal{K}_{20} - \Omega_{\mathbb{S}_+}^{kdv}(\nu)\|_s^{\text{Lip}(\gamma)} \lesssim_s \varepsilon + \|\iota\|_{s+\sigma}^{\text{Lip}(\gamma)}, \\ &\|\mathcal{K}_{11}\eta\|_s^{\text{Lip}(\gamma)} \lesssim_s \varepsilon\gamma^{-2} \|\eta\|_{s+\sigma}^{\text{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\text{Lip}(\gamma)} \|\eta\|_{s_0+\sigma}^{\text{Lip}(\gamma)}, \\ &\|\mathcal{K}_{11}^\top v\|_s^{\text{Lip}(\gamma)} \lesssim_s \varepsilon\gamma^{-2} \|v\|_{s+\sigma}^{\text{Lip}(\gamma)} + \|\iota\|_{s+\sigma}^{\text{Lip}(\gamma)} \|v\|_{s_0+\sigma}^{\text{Lip}(\gamma)}. \end{aligned} \tag{5.23}$$

Proof. First we prove the claimed estimates for $\mathcal{K}_{00}, \mathcal{K}_{10}, \mathcal{K}_{01}$ and then the ones for $\mathcal{K}_{20}, \mathcal{K}_{11}, \mathcal{K}_{11}^\top$.

Estimates of $\mathcal{K}_{00}, \mathcal{K}_{10}, \mathcal{K}_{01}$: The Hamiltonian vector field associated to the Hamiltonian $\mathcal{K}$ in (5.17) is given by $X_\mathcal{K} := (-\nabla_\eta \mathcal{K}, \ \nabla_\phi \mathcal{K} + (\partial_\phi \theta)^\top \zeta, \ \partial_x \nabla_v \mathcal{K})$. Furthermore, since $\breve{\iota}_\delta(\varphi) = G_\delta(\varphi, 0, 0)$, the directional derivative $\omega \cdot \partial_\varphi \breve{\iota}_\delta(\varphi)$ equals $dG_\delta(\varphi, 0, 0)[(\omega, 0, 0)]$. Using the transformation law of vector fields we get that

$$\begin{aligned} \mathcal{F}_\omega(\iota_\delta, \zeta)(\varphi) &= \omega \cdot \partial_\varphi \breve{\iota}_\delta(\varphi) - X_{\mathcal{H}_{\varepsilon,\zeta}}(\breve{\iota}_\delta(\varphi)) \\ &= \text{d}G_\delta(\varphi, 0, 0)[(\omega, 0, 0)] - \text{d}G_\delta(\varphi, 0, 0) X_\mathcal{K}(\varphi, 0, 0), \end{aligned}$$

or

$$X_{\mathcal{K}}(\varphi, 0, 0) = (\omega, 0, 0) - (dG_\delta(\varphi, 0, 0))^{-1}\mathcal{F}_\omega(\iota_\delta, \zeta). \tag{5.24}$$

Furthermore, by (5.18), one computes

$$X_{\mathcal{K}}(\varphi, 0, 0) = \big(- \mathcal{K}_{10}(\varphi),\ \partial_\varphi \theta(\varphi)[\zeta] + \nabla_\varphi \mathcal{K}_{00}(\varphi),\ \partial_x \mathcal{K}_{01}(\varphi)\big). \tag{5.25}$$

By comparing the two expressions (5.24) and (5.25), and by using the estimate (5.5) of ζ, the estimate (5.13) of $\mathcal{F}_\omega(\iota_\delta, \zeta)$, the estimate (5.21) of dG_δ, the ansatz (5.2) and that $\|[\partial_\varphi\theta]^\top\|_s^{\mathrm{Lip}(\gamma)} \leqq 1 + C(s)\|\iota\|_s^{\mathrm{Lip}(\gamma)}$, one gets the first estimate in (5.23).

ESTIMATES OF $\mathcal{K}_{20}, \mathcal{K}_{11}, \mathcal{K}_{11}^\top$: We prove the claimed bound for $\mathcal{K}_{20}$ and $\mathcal{K}_{11}$. The estimate for $\mathcal{K}_{11}^\top$ can be proved arguing similarly. By (5.18), one has $\mathcal{K}_{20}(\varphi) = \partial_\eta \nabla_\eta \mathcal{K}(\varphi, 0, 0)$ and $\mathcal{K}_{11}(\varphi) = \partial_\eta \nabla_v \mathcal{K}(\varphi, 0, 0)$. Furthermore, taking into account the formulae (4.7), (5.16), (5.17), one then infers that

$$\begin{aligned}
\mathcal{K}_{20}(\varphi) &= [\partial_\varphi \theta(\varphi)]^{-1} \partial_y \nabla_y \mathcal{H}_\varepsilon(\breve{\iota}(\varphi))[\partial_\varphi\theta(\varphi)]^{-\top} \\
&\overset{(4.13)}{=} [\partial_\varphi \theta(\varphi)]^{-1} \Omega^{kdv}_{\mathbb{S}_+}(\nu)[\partial_\varphi\theta(\varphi)]^{-\top} + [\partial_\varphi \theta(\varphi)]^{-1} \partial_y \nabla_y \mathcal{P}_\varepsilon(\breve{\iota}(\varphi))[\partial_\varphi\theta(\varphi)]^{-\top}, \\
\mathcal{K}_{11}(\varphi) &= -[(\partial_\theta \tilde{w})(\theta(\varphi))]\partial_y \nabla_y \mathcal{H}_\varepsilon(\breve{\iota}_\delta(\varphi))[\partial_\varphi\theta(\varphi)]^{-\top} + \partial_y \nabla_w \mathcal{H}_\varepsilon(\breve{\iota}_\delta(\varphi))[\partial_\varphi\theta(\varphi)]^{-\top} \\
&\overset{(4.13)}{=} -[(\partial_\theta \tilde{w})(\theta(\varphi))]\Omega^{kdv}_{\mathbb{S}_+}(\nu)[\partial_\varphi\theta(\varphi)]^{-\top} - [(\partial_\theta \tilde{w})(\theta(\varphi))]\partial_y \nabla_y \mathcal{P}_\varepsilon(\breve{\iota}_\delta(\varphi))[\partial_\varphi\theta(\varphi)]^{-\top} \\
&\quad + \partial_y \nabla_w \mathcal{P}_\varepsilon(\breve{\iota}_\delta(\varphi))[\partial_\varphi\theta(\varphi)]^{-\top}.
\end{aligned}$$

By $\|[\partial_\varphi\theta(\varphi)]^{-1} - \mathrm{Id}\|_s^{\mathrm{Lip}(\gamma)} + \|[\partial_\varphi\theta(\varphi)]^{-\top} - \mathrm{Id}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\iota\|_{s+1}^{\mathrm{Lip}(\gamma)}$ (cf. (5.14)), the estimates for $\partial_y \nabla_y \mathcal{P}_\varepsilon(\breve{\iota}_\delta)$, $\partial_y \nabla_w \mathcal{P}_\varepsilon(\breve{\iota}_\delta)$ of Lemma 4.3, the interpolation estimate (2.5), and (5.2), one obtains the claimed bounds. □

In order to construct an almost-approximate right inverse of (5.20), we need that

$$\mathcal{L}_\omega := \Pi_\perp \big(\omega \cdot \partial_\varphi - \partial_x \mathcal{K}_{02}(\varphi)\big)_{|L^2_\perp} \tag{5.26}$$

is "almost-invertible", that is invertible up to a remainder of order $O(N_{n-1}^{-\mathtt{a}})$, where

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

and

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

are the scales used in the nonlinear Nash–Moser iteration in Section 8.1. The constants $\mathtt{a}$, K_0 are given in the almost-invertibility assumption **A-I** below.

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 subset Ω_o of the set Ω of frequency vectors ω, for which $\mathcal{L}_\omega$ can be shown to admit an almost-approximate right inverse. Recall that $\mathtt{DC}(\gamma, \tau)$ is the set of diophantine frequencies in Ω, defined in (4.4).

A-I 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* (5.26) *admits a decomposition*

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

with the following properties: there exist constants K_0, N_0, σ, τ_1, $\mu(\mathtt{b})$, $\mathtt{a}$, p, $\mathfrak{s}_M > 0$, *so that for any* $S > \mathfrak{s}_M$, $\mathfrak{s}_M \leqq s \leqq S$ *and* $\omega \in \Omega_o$ *the following holds:*
(i) *The operators* $\mathcal{R}_\omega$, $\mathcal{R}_\omega^\perp$ *satisfy the estimates*

$$\|\mathcal{R}_\omega h\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \varepsilon\gamma^{-2} N_{n-1}^{-\mathtt{a}} \big(\|h\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1}\|\iota\|_{s+\mu(\mathtt{b})+\sigma}^{\mathrm{Lip}(\gamma)} \|h\|_{\mathfrak{s}_M+\sigma}^{\mathrm{Lip}(\gamma)}\big), \tag{5.30}$$

$$\|\mathcal{R}_\omega^\perp h\|_{\mathfrak{s}_M}^{\mathrm{Lip}(\gamma)} \lesssim_{S,b} K_n^{-b} \big(\|h\|_{\mathfrak{s}_M+b+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1}\|\iota\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\sigma+b}^{\mathrm{Lip}(\gamma)} \|h\|_{\mathfrak{s}_M+\sigma}^{\mathrm{Lip}(\gamma)}\big), \quad \forall b > 0. \tag{5.31}$$

(ii) *The operator* $\mathcal{L}_\omega^<$ *admits a right inverse. More precisely, for any* $g \in H_\perp^{s+\sigma}(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1)$, *there is a solution* $h \in H_\perp^s(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1)$ *of the linear equation* $\mathcal{L}_\omega^< h = g$, *denoted by* $(\mathcal{L}_\omega^<)^{-1} g$, *satisfying the tame estimates*

$$\|(\mathcal{L}_\omega^<)^{-1} g\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \gamma^{-1}\big(\|g\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1}\|\iota\|_{s+\mu(\mathtt{b})+\sigma}^{\mathrm{Lip}(\gamma)} \|g\|_{\mathfrak{s}_M+\sigma}^{\mathrm{Lip}(\gamma)}\big). \tag{5.32}$$

In order to find an almost-approximate inverse of the linear operator (5.20) and hence of $d_{\iota,\zeta}\mathcal{F}_\omega(\iota_\delta)$, note that the remainder $\mathcal{L}_\omega - \mathcal{L}_\omega^< = \mathcal{R}_\omega + \mathcal{R}_\omega^\perp$ is small (cf. (5.30)–(5.31) in **A-I**) and that by Lemma 5.5 and by the estimate (5.5) of ζ, the terms $\partial_\phi\mathcal{K}_{10}$, $\partial_{\phi\phi}\mathcal{K}_{00}$, $\partial_\phi\mathcal{K}_{01}$ and $\partial_\phi\big(\partial_\phi\theta(\varphi)^\top[\zeta]\big)$ in (5.20) are of the order $O(Z)$. Therefore, it suffices to invert the operator

$$\mathbb{D}[\widehat{\phi}, \widehat{\eta}, \widehat{v}, \widehat{\zeta}] := \begin{pmatrix} \omega\cdot\partial_\varphi\widehat{\phi} + \mathcal{K}_{20}(\varphi)\widehat{\eta} + \mathcal{K}_{11}(\varphi)^\top\widehat{v} \\ \omega\cdot\partial_\varphi\widehat{\eta} - \partial_\phi\theta(\varphi)^\top\widehat{\zeta} \\ \mathcal{L}_\omega^<\widehat{v} - \partial_x\mathcal{K}_{11}(\varphi)\widehat{\eta} \end{pmatrix} \tag{5.33}$$

obtained by neglecting in (5.20) the terms $\partial_\phi\mathcal{K}_{10}$, $\partial_{\phi\phi}\mathcal{K}_{00}$, $\partial_\phi\mathcal{K}_{01}$, $\partial_\phi\big(\partial_\phi\theta(\varphi)^\top[\zeta]\big)$ and by replacing $\mathcal{L}_\omega$ by $\mathcal{L}_\omega^<$ (cf. (5.29)). We look for a right inverse of $\mathbb{D}$ by solving the system

$$\mathbb{D}[\widehat{\phi}, \widehat{\eta}, \widehat{v}, \widehat{\zeta}] = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix}. \tag{5.34}$$

We first consider the second equation in (5.34), $\omega\cdot\partial_\varphi\widehat{\eta} = g_2 + \partial_\phi\theta(\varphi)^\top\widehat{\zeta}$. Since $\partial_\varphi\theta(\varphi) = \mathrm{Id} + \partial_\varphi\Theta(\varphi)$, the average $\langle\partial_\varphi\theta^\top\rangle_\varphi = \frac{1}{(2\pi)^{|\mathbb{S}_+|}}\int_{\mathbb{T}^{\mathbb{S}_+}}\partial_\varphi\theta^\top(\varphi)\mathrm{d}\varphi$ equals the identity matrix Id of $\mathbb{R}^{\mathbb{S}_+}$. We then define

$$\widehat{\zeta} := -\langle g_2\rangle_\varphi \tag{5.35}$$

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

$$\widehat{\eta} := \widehat{\eta}_0 + \widehat{\eta}_1, \quad \widehat{\eta}_1 := (\omega\cdot\partial_\varphi)^{-1}\big(g_2 + \partial_\phi\theta(\varphi)^\top\widehat{\zeta}\big), \tag{5.36}$$

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 (5.34). Next we consider the third equation in (5.34),

$(\mathcal{L}_\omega^<)\widehat{v} = g_3 + \partial_x \mathcal{K}_{11}(\varphi)\widehat{\eta}$, which, by assumption (5.32) on the invertibility of $\mathcal{L}_\omega^<$, has the solution

$$\widehat{v} = (\mathcal{L}_\omega^<)^{-1}\big(g_3 + \partial_x \mathcal{K}_{11}(\varphi)\widehat{\eta}_1\big) + (\mathcal{L}_\omega^<)^{-1}\partial_x \mathcal{K}_{11}(\varphi)\widehat{\eta}_0. \tag{5.37}$$

Finally, we solve the first equation in (5.34). After substituting the solutions $\widehat{\zeta}$, $\widehat{\eta}$ (cf. (5.35), (5.36)), and $\widehat{v}$ (cf. (5.37)), this equation becomes

$$\omega \cdot \partial_\varphi \widehat{\phi} = g_1 + M_1\widehat{\eta}_0 + M_2 g_2 + M_3 g_3 - M_4\langle g_2\rangle_\varphi, \tag{5.38}$$

where $\widehat{\phi}$ and $\widehat{\eta}_0$ are the unknowns and where $M_j : \varphi \mapsto M_j(\varphi)$, $1 \le j \le 4$, are defined as

$$M_1(\varphi) := -\mathcal{K}_{20}(\varphi) - \mathcal{K}_{11}(\varphi)^\top (\mathcal{L}_\omega^<)^{-1}\partial_x \mathcal{K}_{11}(\varphi), \tag{5.39}$$

$$M_2(\varphi) := M_1(\varphi)[\omega\cdot\partial_\varphi]^{-1}, \quad M_3(\varphi) := -\mathcal{K}_{11}(\varphi)^\top(\mathcal{L}_\omega^<)^{-1}, \quad M_4(\varphi) := M_2(\varphi)\partial_\phi\theta(\varphi)^\top. \tag{5.40}$$

In order to solve equation (5.38) we have to choose $\widehat{\eta}_0$ so that the average of the right hand side vanishes. By Lemma 5.5, by the ansatz (5.2) and by the tame estimates (5.32), the φ-averaged matrix is $\langle M_1\rangle_\varphi = -\Omega^{kdv}_{\mathbb{S}_+}(\nu) + O(\varepsilon\gamma^{-2})$. Since the matrix $\Omega^{kdv}_{\mathbb{S}_+}(\nu) = (\partial_{I_k}\omega_n^{kdv}(\nu, 0))_{k,n\in\mathbb{S}_+}$ is invertible (cf. Lemma 3.9-(i), Remark 3.10), $\langle M_1\rangle_\varphi$ 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

$$\widehat{\eta}_0 := -\langle M_1\rangle_\varphi^{-1}\Big(\langle g_1\rangle_\varphi + \langle M_2 g_2\rangle_\varphi + \langle M_3 g_3\rangle_\varphi - \langle M_4\rangle_\varphi\langle g_2\rangle_\varphi\Big). \tag{5.41}$$

With this choice of $\widehat{\eta}_0$, the equation (5.38) has the solution

$$\widehat{\phi} = (\omega\cdot\partial_\varphi)^{-1}\big(g_1 + M_1\widehat{\eta}_0 + M_2 g_2 + M_3 g_3 - M_4\langle g_2\rangle_\varphi\big). \tag{5.42}$$

Altogether, we have obtained a solution $(\widehat{\phi}, \widehat{\eta}, \widehat{v}, \widehat{\zeta})$ of the linear system (5.34).

Proposition 5.6. *Assume (5.2) (Ansatz) with $\mu_0 = \mu(\mathtt{b}) + \sigma$ and that the estimates (5.32) (item (ii) of* **A-I***) hold. Then, for any $\omega \in \Omega_o$ and any $g := (g_1, g_2, g_3)$ with $g_1, g_2 \in H^{s+\sigma}(\mathbb{T}^{\mathbb{S}_+}, \mathbb{R}^{\mathbb{S}_+})$, $g_3 \in H_\perp^{s+\sigma}(\mathbb{T}^{\mathbb{S}_+}\times\mathbb{T}_1)$, and $\mathfrak{s}_M \leqq s \leqq S$, the system (5.34) has a solution $(\widehat{\phi}, \widehat{\eta}, \widehat{v}, \widehat{\zeta})$, where $\widehat{\phi}$, $\widehat{\eta}$, $\widehat{v}$, $\widehat{\zeta}$ are defined in (5.35)–(5.37), (5.41)–(5.42). We denote $(\widehat{\phi}, \widehat{\eta}, \widehat{v}, \widehat{\zeta})$ by $\mathbb{D}^{-1}g$. It satisfies the tame estimates*

$$\|\mathbb{D}^{-1}g\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \gamma^{-2}\big(\|g\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1}\|\iota\|_{s+\mu(\mathtt{b})+\sigma}^{\mathrm{Lip}(\gamma)}\|g\|_{\mathfrak{s}_M+\sigma}^{\mathrm{Lip}(\gamma)}\big). \tag{5.43}$$

Proof. The proposition follows by the definitions of $\widehat{\zeta}$ (cf. (5.35)), $\widehat{\eta}_1$ (cf. (5.36)), $\widehat{v}$ (cf. (5.37)), $\widehat{\eta}_0$ (cf. (5.41)), $\widehat{\phi}$ (cf. (5.42)), the definitions of M_j, $1 \le j \le 4$, in (5.39)–(5.40), by the estimates of Lemma 5.5, and the ansatz (5.2) as well as the estimates (5.32) for $(\mathcal{L}_\omega^<)^{-1}$ (item (ii) in **A-I**). □

Let $\widetilde{G}_\delta : (\phi, \eta, v, \zeta) \mapsto \big(G_\delta(\phi, \eta, v), \zeta\big)$ and note that its differential $\mathrm{d}\widetilde{G}_\delta(\phi, \eta, v, \zeta)$ 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

$$\mathbf{T}_0 = \mathbf{T}_0(\iota) := \mathrm{d}\widetilde{G}_\delta|_{(\varphi,0,0)} \circ \mathbb{D}^{-1} \circ \big(dG_\delta|_{(\varphi,0,0)}\big)^{-1} \tag{5.44}$$

is an almost-approximate right inverse for $d_{\iota,\zeta}\mathcal{F}_\omega(\iota)$ meaning that $d_{\iota,\zeta}\mathcal{F}_\omega(\iota) \circ \mathbf{T}_0(\iota) - \mathrm{Id}$ can be estimated in terms of the error function $Z = \mathcal{F}_\omega(\iota)$ ('approximate', cf. (5.47)) and of terms which are small ('almost', cf. (5.48), (5.49)). Let $\|(\phi, \eta, v, \zeta)\|_s^{\mathrm{Lip}(\gamma)} := \max\{\|(\phi, \eta, v)\|_s^{\mathrm{Lip}(\gamma)}, |\zeta|^{\mathrm{Lip}(\gamma)}\}$.

Theorem 5.7. (Almost-approximate inverse) *Assume the almost-invertibility assumption* **A-I** *of $\mathcal{L}_\omega$. Then there exists $\sigma_2 = \sigma_2(\tau, \mathbb{S}_+) > 0$ so that, if (5.2) (Ansatz) holds with $\mu_0 \geqq \mathfrak{s}_M + \mu(\mathrm{b}) + \sigma_2$, then for any $\omega \in \Omega_o$ and any $g = (g_1, g_2, g_3)$ with $g_1, g_2 \in H^{s+\sigma}(\mathbb{T}^{\mathbb{S}_+}, \mathbb{R}^{\mathbb{S}_+})$, $g_3 \in H_\perp^{s+\sigma}(\mathbb{T}^{\mathbb{S}_+} \times \mathbb{T}_1)$, and $\mathfrak{s}_M \leqq s \leqq S$, $\mathbf{T}_0(\iota)g$, defined by (5.44), satisfies*

$$\|\mathbf{T}_0(\iota)g\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \gamma^{-2}\Big(\|g\|_{s+\sigma_2}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1}\|\iota\|_{s+\mu(\mathrm{b})+\sigma_2}^{\mathrm{Lip}(\gamma)}\|g\|_{\mathfrak{s}_M+\sigma_2}^{\mathrm{Lip}(\gamma)}\Big). \tag{5.45}$$

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

$$d_{\iota,\zeta}\mathcal{F}_\omega(\iota) \circ \mathbf{T}_0(\iota) - \mathrm{Id} = \mathcal{P} + \mathcal{P}_\omega + \mathcal{P}_\omega^\perp, \tag{5.46}$$

where the operators $\mathcal{P}$, $\mathcal{P}_\omega$, $\mathcal{P}_\omega^\perp$ are defined in the course of the proof and satisfy the following estimates:

$$\|\mathcal{P}g\|_{\mathfrak{s}_M}^{\mathrm{Lip}(\gamma)} \lesssim_S \gamma^{-3}\|\mathcal{F}_\omega(\iota, \zeta)\|_{\mathfrak{s}_M+\sigma_2}^{\mathrm{Lip}(\gamma)}\Big(1 + N_0^{\tau_1}\gamma^{-1}\|\iota\|_{\mathfrak{s}_M+\mu(\mathrm{b})+\sigma_2}^{\mathrm{Lip}(\gamma)}\Big)\|g\|_{\mathfrak{s}_M+\sigma_2}^{\mathrm{Lip}(\gamma)}, \tag{5.47}$$

$$\|\mathcal{P}_\omega g\|_{\mathfrak{s}_M}^{\mathrm{Lip}(\gamma)} \lesssim_S \varepsilon\gamma^{-4}N_{n-1}^{-\mathrm{a}}\big(1 + N_0^{\tau_1}\gamma^{-1}\|\iota\|_{\mathfrak{s}_M+\mu(\mathrm{b})+\sigma_2}^{\mathrm{Lip}(\gamma)}\big)\|g\|_{\mathfrak{s}_M+\sigma_2}^{\mathrm{Lip}(\gamma)}, \tag{5.48}$$

$$\|\mathcal{P}_\omega^\perp g\|_{\mathfrak{s}_M}^{\mathrm{Lip}(\gamma)} \lesssim_{S,b} \gamma^{-2}K_n^{-b}\big(\|g\|_{\mathfrak{s}_M+\sigma_2+b}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1}\gamma^{-1}\|\iota\|_{\mathfrak{s}_M+\mu(\mathrm{b})+\sigma_2+b}^{\mathrm{Lip}(\gamma)}\|g\|_{\mathfrak{s}_M+\sigma_2}^{\mathrm{Lip}(\gamma)}\big), \quad \forall b > 0. \tag{5.49}$$

Proof. The bound (5.45) follows from the definition of $\mathbf{T}_0(\iota)$ in (5.44), the estimates (5.43) of $\|\mathbb{D}^{-1}g\|_s^{\mathrm{Lip}(\gamma)}$, and the ones of $dG_\delta(\varphi, 0, 0)$ and of its inverse in (5.21). It remains to estimate $d_{\iota,\zeta}\mathcal{F}_\omega(\iota) \circ \mathbf{T}_0(\iota) - \mathrm{Id}$. The operators $\mathcal{P}$, $\mathcal{P}_\omega$, $\mathcal{P}_\omega^\perp$ in (5.46) are defined as follows: by the 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)), one has $d_{\iota,\zeta}\mathcal{F}_\omega(\iota) = d_{\iota,\zeta}\mathcal{F}_\omega(\iota_\delta) + \mathcal{E}_0$ where, by the mean value theorem, $\mathcal{E}_0$ can be written as

$$\mathcal{E}_0[\widehat{\iota}, \widehat{\zeta}] = \int_0^1 \partial_y\big(d_\iota X_{\mathcal{H}_\varepsilon}(\theta, y_\delta + s(y - y_\delta), w)[\widehat{\iota}]\big)\mathrm{d}s\,[y - y_\delta]. \tag{5.50}$$

Denote by $\kappa := (\phi, \eta, v)$ the symplectic coordinates defined by G_δ. Let also $\breve{\kappa}(\varphi) = (\varphi, 0, 0) + \kappa(\varphi)$ the torus embedding defined by $\breve{\iota}(\varphi) = G_\delta(\breve{\kappa}(\varphi))$ where G_δ is the symplectic transformation given by (5.16). The nonlinear operator $\mathcal{F}_\omega$ (cf. (4.6)) is transformed under G_δ into

$$\mathcal{F}_\omega(\iota, \zeta)(\varphi) = \mathrm{d}G_\delta(\breve{\kappa}(\varphi))[\omega \cdot \partial_\varphi\breve{\kappa}(\varphi) - X_{\mathcal{K}}(\breve{\kappa}(\varphi))] \tag{5.51}$$

where $\mathcal{K} = \mathcal{H}_{\varepsilon,\zeta} \circ G_\delta$ (cf. (5.17)). Differentiating (5.51) at the trivial torus $\breve{\kappa}_0(\varphi) = G_\delta^{-1}(\breve{\iota}_\delta)(\varphi) = (\varphi, 0, 0)$, we get

$$d_{\iota,\zeta}\mathcal{F}_\omega(\iota_\delta) = dG_\delta(\breve{\kappa}_0)\big(\omega\cdot\partial_\varphi - d_{\kappa,\zeta}X_{\mathcal{K}}(\breve{\kappa}_0)\big)\mathrm{d}\widetilde{G}_\delta(\breve{\kappa}_0)^{-1} + \mathcal{E}_1, \tag{5.52}$$

$$\mathcal{E}_1 := d^2 G_\delta(\breve{\kappa}_0)\big[dG_\delta(\breve{\kappa}_0)^{-1}\mathcal{F}_\omega(\iota_\delta,\zeta),\ \mathrm{d}G_\delta(\breve{\kappa}_0)^{-1}\Pi[\,\cdot\,]\big], \tag{5.53}$$

where Π denotes the projection $\Pi : (\widehat{\iota}, \widehat{\zeta}) \mapsto \widehat{\iota}$. Let us consider the operator $\omega\cdot\partial_\varphi - d_{\kappa,\zeta}X_{\mathcal{K}}(\breve{\kappa}_0)$ in more detail. By the definition (5.33) of $\mathbb{D}$ and the discussion following it, we decompose $\omega\cdot\partial_\varphi - d_{\kappa,\zeta}X_{\mathcal{K}}(\breve{\kappa}_0)$ as

$$\omega\cdot\partial_\varphi - d_{\kappa,\zeta}X_{\mathcal{K}}(\breve{\kappa}_0) = \mathbb{D} + R_Z + \mathbb{R}_\omega + \mathbb{R}_\omega^\perp \tag{5.54}$$

where in view of (5.20),

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

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

with $\mathcal{R}_\omega$ and $\mathcal{R}_\omega^\perp$ given by (5.29). By (5.50) and (5.52)–(5.54) we get the decomposition

$$d_{\iota,\zeta}\mathcal{F}_\omega(\iota) = dG_\delta(\breve{\kappa}_0)\circ\mathbb{D}\circ\big(d\widetilde{G}_\delta(\breve{\kappa}_0)\big)^{-1} + \mathcal{E} + \mathcal{E}_\omega + \mathcal{E}_\omega^\perp \tag{5.55}$$

where

$$\mathcal{E} := \mathcal{E}_0 + \mathcal{E}_1 + dG_\delta(\breve{\kappa}_0)R_Z\big(d\widetilde{G}_\delta(\breve{\kappa}_0)\big)^{-1}, \tag{5.56}$$

$$\mathcal{E}_\omega := \mathrm{d}G_\delta(\breve{\kappa}_0)\mathbb{R}_\omega\big(d\widetilde{G}_\delta(\breve{\kappa}_0)\big)^{-1}, \quad \mathcal{E}_\omega^\perp := \mathrm{d}G_\delta(\breve{\kappa}_0)\mathbb{R}_\omega^\perp\big(\mathrm{d}\widetilde{G}_\delta(\breve{\kappa}_0)\big)^{-1}. \tag{5.57}$$

Letting the operator $\mathbf{T}_0 = \mathbf{T}_0(\iota)$ (cf. (5.44)) act from the right to both sides of the identity (5.55) and taking into accout that $\breve{\kappa}_0(\varphi) = (\varphi, 0, 0)$, one obtains

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

To obtain the claimed estimate for $\mathcal{P}$ we first need to estimate $\mathcal{E}$. By (5.2) (Ansatz), (5.5) (estimate for ζ), Lemma 5.5 (estimates of the components of R_Z), (5.10)–(5.13) (estimates related to ι_δ), and (5.21)–(5.22) (estimates of $dG_\delta(\breve{\kappa}_0)$ and its inverse) one infers that

$$\|\mathcal{E}[\widehat{\iota},\widehat{\zeta}\,]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \gamma^{-1}\Big(\|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|\widehat{\iota}\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|\widehat{\iota}\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|\widehat{\iota}\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\Big), \tag{5.58}$$

for some $\sigma > 0$, where Z is the error function, $Z = \mathcal{F}_\omega(\iota,\zeta)$ (cf. (5.3)). The claimed estimate (5.47) for $\mathcal{P}$ then follows from (5.58), the estimate (5.45) of $\mathbf{T}_0$, and the ansatz (5.2). The claimed estimates (5.48), (5.49) for $\mathcal{P}_\omega$ and, respectively, $\mathcal{P}_\omega^\perp$ follow by the estimates (5.30)–(5.31) of $\mathcal{R}_\omega$ and $\mathcal{R}_\omega^\perp$ (cf. **A-I**), the estimate (5.45) of $\mathbf{T}_0$, the estimate (5.21) of $dG_\delta(\breve{\kappa}_0)$ 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 (5.26), satisfies the almost-invertibility **A-I**, including the tame estimates (5.30)–(5.32).

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 (5.26), to a differential operator of order three with constant coefficients, up to an operator of order zero – see $\mathcal{L}_\omega^{(4)}$ defined in (6.69). It is the starting point for the KAM reduction scheme, implemented in Section 7, which will reduce $\mathcal{L}_\omega^{(4)}$ to a diagonal operator with constant coefficients. The main result of this section is Proposition 6.7.

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

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

where $\mu_0 := \mu_0(\tau, \mathbb{S}_+) > s_0$ and $S > \mu_0$ are sufficiently large, $0 < \delta(S) < 1$ is sufficiently small, and $0 < \gamma < 1$. The index S of the Sobolev space $H_\perp^S$ will be fixed in (8.4), along the Nash Moser iteration scheme of Section 8.1. 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).

Notation. 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 (5.26).

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 $s \geqq s_0$

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

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

$$\|\Delta_{12} g_j\|_{s_1} + \|\Delta_{12}\chi_j\|_{s_1} \lesssim_{s_1} \|\iota_2 - \iota_1\|_{s_1+\sigma}. \tag{6.5}$$

Proof. The operator $\mathcal{K}_{02}(\varphi)$ is defined by $\mathcal{K}_{02}(\varphi) = d_\perp \nabla_v \mathcal{K}(\varphi, 0, 0) = d_\perp \nabla_v (\mathcal{H}_\varepsilon \circ G_\delta)(\varphi, 0, 0)$. Differentiating the Hamiltonian $(\mathcal{H}_\varepsilon \circ G_\delta)(\varphi, \eta, v) = \mathcal{H}_\varepsilon\big(\theta(\varphi), y_\delta(\varphi) + L_1(\varphi)\eta + L_2(\varphi)w, w(\varphi) + v\big)$ with respect to v, we get $\nabla_v (\mathcal{H}_\varepsilon \circ G_\delta)(\varphi, \eta, v) = L_2(\varphi)^\top \partial_y \mathcal{H}_\varepsilon(G_\delta(\varphi, \eta, v)) + \nabla_w \mathcal{H}_\varepsilon(G_\delta(\varphi, \eta, v))$, where we used that by (5.16) $L_1(\varphi) := [\partial_\phi \theta_0(\varphi)]^{-\top}$ and $L_2(\varphi) := -[\partial_\theta \tilde{w}(\theta(\varphi))]^\top \partial_x^{-1}$. Since $G_\delta(\varphi, 0, 0) = \breve{\iota}_\delta(\varphi)$, it then follows that

$$\partial_x d_\perp \nabla_v (\mathcal{H}_\varepsilon \circ G_\delta)(\varphi, 0, 0) = \partial_x d_\perp \nabla_w \mathcal{H}_\varepsilon(\breve{\iota}_\delta(\varphi)) + R(\varphi)$$

where $R(\varphi) := R_1(\varphi) + R_2(\varphi) + R_3(\varphi)$ with $R_1(\varphi) := \partial_x L_2(\varphi)^\top \partial_{yy} \mathcal{H}_\varepsilon(\breve{\iota}_\delta(\varphi)) L_2(\phi)$ and

$$R_2(\varphi) := \partial_x L_2(\varphi)^\top d_\perp \partial_y \mathcal{H}_\varepsilon(\breve{\iota}_\delta(\varphi)), \quad R_3(\varphi) := \partial_x \partial_y \nabla_w \mathcal{H}_\varepsilon(\breve{\iota}_\delta(\varphi)) L_2(\varphi).$$

Each of the linear operators R_1, R_2, R_3 is of the form (6.3) since it is the composition of linear operators, at least one of which has finite rank. For example, expressing the linear operator $L_2(\phi) : L^2_\perp(\mathbb{T}_1) \to \mathbb{R}^{\mathbb{S}_+}$ in terms of the canonical basis $\underline{e}_j$, $j \in \mathbb{S}_+$, $L_2(\phi)[h] = \sum_{j \in \mathbb{S}_+} \big(h,\ L_2(\phi)^\top [\underline{e}_j]\big)_{L^2_x} \underline{e}_j$, $\forall\, h \in L^2_\perp$, one obtains

$$R_1(\varphi)[h] = \sum\nolimits_{j \in \mathbb{S}_+} \big(h,\ L_2(\varphi)^\top [\underline{e}_j]\big)_{L^2_x} A_1(\varphi)[\underline{e}_j], \quad A_1(\varphi) := \partial_x L_2(\varphi)^\top \partial_{yy} \mathcal{H}_\varepsilon(\breve{\iota}_\delta(\varphi)),$$

showing that it has the form (6.3). By similar arguments one concludes the same for R_2 and R_3. Let us prove that R_1 satisfies the estimates (6.4). By the explicit form of $L_2(\varphi)$, Lemma 2.1-(ii) and (5.2), one gets

$$\|L_2(\varphi)[\underline{e}_j]\|_s^{\mathrm{Lip}(\gamma)},\ \|L_2(\varphi)^\top [\underline{e}_j]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \quad \text{for some} \quad \sigma > 0. \tag{6.6}$$

Furthermore, since $\mathcal{H}_\varepsilon = \mathcal{N} + \mathcal{P}_\varepsilon$ with $\mathcal{N}$ and $\mathcal{P}_\varepsilon$ given by (4.13), one has $\partial_{yy} \mathcal{H}_\varepsilon = \Omega^{kdv}_{\mathbb{S}_+} + \partial_{yy} \mathcal{P}_\varepsilon$, $\partial_y \nabla_w \mathcal{H}_\varepsilon = \partial_y \nabla_w \mathcal{P}_\varepsilon$ and $d_\perp \partial_y \mathcal{H}_\varepsilon = d_\perp \partial_y \mathcal{P}_\varepsilon$. Using the estimates in Lemma 4.3 and (5.2), we then infer the bound $\|A_1[\underline{e}_j]\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\iota\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}$ for some $\sigma > 0$, implying together with (6.6) the claimed estimate in (6.4). By similar arguments, one obtains the ones for R_2 and R_3, as well as the estimate (6.5). □

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

$$\mathcal{L}_\omega = \mathcal{L}_\omega^{(0)} - R, \quad \mathcal{L}_\omega^{(0)} := \omega \cdot \partial_\varphi - \Pi_\perp \partial_x (d_\perp \nabla_w \mathcal{H}_\varepsilon)(\breve{\iota}_\delta(\varphi)), \tag{6.7}$$

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. The operator $\mathcal{L}_\omega$ is defined for any $\omega \in \Omega$. In a next step we prove in Lemmata 6.2 and 6.3 below that the Hamiltonian operator $\mathcal{L}_\omega^{(0)}$, acting on $L^2_\perp(\mathbb{T}_1)$, is the sum of a pseudo-differential operator of order three, a Fourier multiplier with $\varphi-$independent coefficients and a small smoothing remainder.

First note that, 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. (3.12)), we have

$$\mathcal{L}_\omega^{(0)} = \omega \cdot \partial_\varphi + \partial_x^3 - \Pi_\perp Q_{-1}^{kdv}(D; \omega) - \Pi_\perp \partial_x d_\perp \nabla_w \mathcal{R}^{kdv}(\breve{\iota}_\delta) - \varepsilon \Pi_\perp \partial_x d_\perp \nabla_w \mathcal{P}(\breve{\iota}_\delta) \tag{6.8}$$

where we write ∂_x^3 instead of $\partial_x^3|_{L^2_\perp}$ and where $Q_{-1}^{kdv}(D; \omega)$ is given by (cf. (3.62))

$$Q_{-1}^{kdv}(D; \omega) \equiv Q_{-1}^{kdv}(D; \nu(\omega)) = \partial_x \Omega^{kdv}(D; \nu(\omega)) + \partial_x^3 \tag{6.9}$$

with $\nu(\omega)$ defined in (4.3).

Lemma 6.2. (Asymptotic expansion of $\mathcal{L}_\omega^{(0)}$) *For any $M \in \mathbb{N}$, the Hamiltonian operator $\mathcal{L}_\omega^{(0)}$, $\omega \in \Omega$, acting on $L^2_\perp(\mathbb{T}_1)$, defined in (6.7), admits an expansion of the form*

$$\begin{aligned} \mathcal{L}_\omega^{(0)} \; &:= \omega \cdot \partial_\varphi - \Pi_\perp \Big(a_3^{(0)} \partial_x^3 + 2(a_3^{(0)})_x \partial_x^2 + a_1^{(0)} \partial_x + \mathrm{Op}(r_0^{(0)}) + Q_{-1}^{kdv}(D; \omega) \Big) \\ &\quad + \mathcal{R}_M^{(0)}(\breve{\iota}_\delta(\varphi); \omega), \end{aligned} \tag{6.10}$$

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 $s \geq s_0$

$$\| a_3^{(0)} + 1 \|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon (1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}), \quad \| a_1^{(0)} \|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)} \tag{6.11}$$

for some $\sigma_M > 0$ and where the pseudo-differential symbol $r_0^{(0)} := r_0^{(0)}(\varphi, x, \xi; \omega)$ has an expansion in homogeneous components

$$r_0^{(0)}(\varphi, x, \xi; \omega) = \sum_{k=0}^{M} a_{-k}^{(0)}(\varphi, x; \omega) (\mathrm{i} 2\pi \xi)^{-k} \chi_0(\xi) \tag{6.12}$$

(with χ_0 defined in (2.18)) where the coefficients $a_{-k}^{(0)} := a_{-k}^{(0)}(\varphi, x; \omega)$ satisfy

$$\sup_{k=0,\dots,M} \| a_{-k}^{(0)} \|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}, \quad \forall s \geqq s_0. \tag{6.13}$$

Furthermore, the remainder is defined by

$$\begin{aligned} \mathcal{R}_M^{(0)}(\breve{\iota}_\delta(\varphi); \omega) \; &:= - \mathcal{R}_M(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{R}^{kdv}) \\ &\quad - \varepsilon \mathcal{R}_M(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{P}) \end{aligned} \tag{6.14}$$

where the latter two remainders are given by (3.48) and (3.36) with $\nu(\omega) = (\omega^{kdv})^{-1}(-\omega)$.

Finally, for any $s_1 \geqq s_0$ and any torus embeddings $\breve{\iota}_1, \breve{\iota}_2$ satisfying (6.1) with $\mu_0 \geqq s_1 + \sigma_M$ it follows that for any $0 \leq k \leq M + 1$,

$$\|\Delta_{12} a_3^{(0)}\|_{s_1} \lesssim_{s_1, M} \varepsilon \|\iota_1 - \iota_2\|_{s_1+\sigma_M}, \quad \|\Delta_{12} a_{1-k}^{(0)}\|_{s_1} \lesssim_{s_1, M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}. \tag{6.15}$$

Proof. By the definition (6.8) of $\mathcal{L}_\omega^{(0)}$, the expansion (3.48) of $\partial_x d_\perp \nabla_w \mathcal{R}^{kdv}$, the expansion (3.36) of $\partial_x d_\perp \nabla_w \mathcal{P}$, and the formula for the coefficient of ∂_x^2, described in Lemma 2.7, one obtains (6.10) with

$$a_3^{(0)}(\varphi, x; \omega) := -1 + \varepsilon a_3(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{P}),$$
$$a_1^{(0)}(\varphi, x; \omega) := a_1(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{R}^{kdv}) + \varepsilon a_1(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{P}),$$
$$a_{-k}^{(0)}(\varphi, x; \omega) := a_{-k}(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{R}^{kdv}) + \varepsilon a_{-k}(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{P}), \quad k = 0, \dots, M,$$

and $\nu(\omega) = (\omega^{kdv})^{-1}(-\omega)$. By Lemma 3.7-1, the functions $a_{1-k}(\mathfrak{x}; \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{R}^{kdv})$, $0 \le k \le M+1$, satisfy the hypothesis of Lemma 2.25-(ii). In view of (5.10) one then infers that for any $s \geqq s_0$

$$\|a_{1-k}(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{R}^{kdv})\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}$$

for some $\sigma_M > 0$. Similarly, by the first item of Lemma 3.5, the functions $a_{3-k}(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{P})$, $0 \le k \le M+3$, satisfy the hypothesis of Lemma 2.25-(i), implying that for any $s \geqq s_0$,

$$\|a_{3-k}(\breve{\iota}_\delta(\varphi); \nu(\omega); \partial_x d_\perp \nabla_w \mathcal{P})\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} 1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}$$

for some $\sigma_M > 0$, proving (6.11), (6.13). The estimates (6.15) follow by similar arguments. □

We remark that in the finitely many steps of our reduction procedure, described in the subsequent sections, the *loss of derivatives* $\sigma_M = \sigma_M(\tau, \mathbb{S}_+) > 0$ might have to be increased, but the notation will not be changed.

We finish this section by showing that the operator $Q_{-1}^{kdv}(D; \omega)$, which is a Fourier multiplier with $\varphi-$independent coefficients, admits an expansion as described in the following lemma:

Lemma 6.3. *For any $M \in \mathbb{N}$,*

$$Q_{-1}^{kdv}(D; \omega) = \sum_{k=1}^{M} c_{-k}^{kdv}(\omega) \partial_x^{-k} + \mathcal{R}_M(Q_{-1}^{kdv}; \omega) \tag{6.16}$$

where for any $1 \le k \le M$, the function $\Omega \to \mathbb{R}$, $\omega \mapsto c_{-k}^{kdv}(\omega)$ is Lipschitz and where $\mathcal{R}_M(Q_{-1}^{kdv}; \omega) : L_\perp^2(\mathbb{T}_1) \to L_\perp^2(\mathbb{T}_1)$ is a Lipschitz family of diagonal operators of order $-M-1$. Furthermore, for any $n_1, n_2 \in \mathbb{N}$, $n_1 + n_2 \leqq M+1$, the operator $\langle D\rangle^{n_1} \mathcal{R}_M(Q_{-1}^{kdv}; \omega) \langle D\rangle^{n_2}$ is Lip(γ)*-tame with a tame constant satisfying $\mathfrak{M}_{\langle D\rangle^{n_1} \mathcal{R}_M(Q_{-1}^{kdv}; \omega)\langle D\rangle^{n_2}}(s) \leqq C(s, M)$ for any $s \geqq s_0$ and $C(s, M) > 0$.*

Proof. By Lemma 3.8, $Q^{kdv}_{-1}(D;\omega) = \Omega^{kdv}_{-1}(D;\omega) + \mathcal{R}_M(D;\omega;Q^{kdv}_{-1})$ where

$$\Omega^{kdv}_{-1}(D;\omega)h = \sum_{k=1}^{M}\sum_{\xi\in\mathbb{S}^\perp} a_{-k}(\omega;\Omega^{kdv}_{-1})\chi_0(\xi)(\mathrm{i}2\pi\xi)^{-k}h_\xi(\varphi)e^{\mathrm{i}2\pi x\xi},$$

$$\mathcal{R}_M(D;\omega;Q^{kdv}_{-1})[h] = \sum_{\xi\in\mathbb{S}^\perp}\frac{\mathcal{R}^{\omega_\xi}_M(\omega)}{(2\pi\xi)^{M+1}}h_\xi(\varphi)e^{\mathrm{i}2\pi x\xi}.$$

Hence (6.16) holds with $c^{kdv}_{-k}(\omega) := a_{-k}(\omega;\Omega^{kdv}_{-1}), k = 1,\dots,M$ and $\mathcal{R}_M(Q^{kdv}_{-1};\omega) = \mathcal{R}_M(D;\omega;Q^{kdv}_{-1})$. For any integer n_1, n_2 such that $n_1 + n_2 \leqq M + 1$, one has that

$$\langle D\rangle^{n_1}\mathcal{R}_M(Q^{kdv}_{-1};\omega)\langle D\rangle^{n_2}h = \sum_{\xi\in\mathbb{S}^\perp}\frac{\mathcal{R}^{\omega_\xi}_M(\omega)\langle\xi\rangle^{n_1+n_2}}{(2\pi\xi)^{M+1}}h_\xi(\varphi)e^{\mathrm{i}2\pi x\xi},$$

where, by (3.65), $\left|\frac{\mathcal{R}^{\omega_\xi}_M(\omega)\langle\xi\rangle^{n_1+n_2}}{(2\pi\xi)^{M+1}}\right| \leqq C_M$. Therefore $\|\langle D\rangle^{n_1}\mathcal{R}_M(Q^{kdv}_{-1};\omega)\langle D\rangle^{n_2}h\|_s \lesssim_M \|h\|_s$ for any $s \geqq 0$. The correpsonding Lipschitz estimate is proved in a similar way. □

6.2. *Quasi-periodic Reparametrization of Time*

The goal of this section is to conjugate the operator $\mathcal{L}_\omega = \mathcal{L}^{(0)}_\omega - R$ in (6.7) to the operator $\mathcal{L}^{(1)}_\omega$, defined in (6.22), which admits an expansion of the form (6.23) with the property that its highest order coefficient $a^{(1)}_3$ satisfies (6.25). This property will allow us in Section 6.3 to conjugate $\mathcal{L}^{(1)}_\omega$ to an operator with constant highest order coefficient (cf. (6.42)).

The operator $\Phi^{(1)}$, by which $\mathcal{L}_\omega$ is conjugated, is induced by the change of variable φ, defined by the quasi-periodic reparametrization of time,

$$\vartheta = \varphi + \alpha^{(1)}(\varphi)\omega \quad \text{or equivalently} \quad \varphi = \vartheta + \breve{\alpha}^{(1)}(\vartheta)\omega$$

where $\alpha^{(1)} : \mathbb{T}^{\mathbb{S}_+} \to \mathbb{R}$, is a small, real valued function chosen below (cf. (6.19)). In more detail, $\Phi^{(1)}$ and its inverse $(\Phi^{(1)})^{-1}$ are given by

$$(\Phi^{(1)}h)(\varphi,x) := h(\varphi + \alpha^{(1)}(\varphi)\omega, x), \quad ((\Phi^{(1)})^{-1}h)(\vartheta,x) := h(\vartheta + \breve{\alpha}^{(1)}(\vartheta)\omega, x). \tag{6.17}$$

First recall that the coefficient $a^{(0)}_3$ in the expansion (6.10) satisfies $a^{(0)}_3 = -1 + O(\varepsilon)$ (cf. (6.11)). Hence the the cube root $(a^{(0)}_3(\varphi,x))^{\frac{1}{3}}$ is smooth.

Lemma 6.4. *Let m_3 be the constant*

$$m_3(\omega) := \frac{1}{(2\pi)^{|\mathbb{S}_+|}}\int_{\mathbb{T}^{\mathbb{S}_+}}\left(\int_{\mathbb{T}_1}\frac{dx}{(a^{(0)}_3(\vartheta,x;\omega))^{\frac{1}{3}}}\right)^{-3}\mathrm{d}\vartheta, \quad \forall\omega\in\Omega, \tag{6.18}$$

and define, for $\omega \in \mathtt{DC}(\gamma,\tau)$, the function

$$\breve{\alpha}^{(1)}(\vartheta;\omega) := (\omega\cdot\partial_\varphi)^{-1}\Big[\frac{1}{m_3}\Big(\int_{\mathbb{T}_1}\frac{dx}{(a_3^{(0)}(\vartheta,x;\omega))^{\frac{1}{3}}}\Big)^{-3} - 1\Big]. \tag{6.19}$$

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*

$$|m_3 + 1|^{\mathrm{Lip}(\gamma)} \lesssim_M \varepsilon \tag{6.20}$$

and for any $s \geqq s_0$, $\alpha^{(1)}, \breve{\alpha}^{(1)}$ satisfy

$$\|\alpha^{(1)}\|_s^{\mathrm{Lip}(\gamma)}, \|\breve{\alpha}^{(1)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon\gamma^{-1}(1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}). \tag{6.21}$$

(ii) *The Hamiltonian operator*

$$\mathcal{L}_\omega^{(1)} := \frac{1}{\rho}\Phi^{(1)}\mathcal{L}_\omega(\Phi^{(1)})^{-1},$$
$$\rho(\vartheta) :=, \Phi^{(1)}(1 + \omega\cdot\partial_\vartheta\breve{\alpha}^{(1)}) = 1 + \Phi^{(1)}(\omega\cdot\partial_\vartheta\breve{\alpha}^{(1)}), \tag{6.22}$$

admits an expansion of the form

$$\mathcal{L}_\omega^{(1)} = \omega\cdot\partial_\vartheta - \Big(a_3^{(1)}\partial_x^3 + 2(a_3^{(1)})_x\partial_x^2 + a_1^{(1)}\partial_x + \mathrm{Op}(r_0^{(1)}) + \mathcal{Q}_{-1}^{kdv}(D;\omega)\Big) + \mathcal{R}_M^{(1)} \tag{6.23}$$

where the coefficients $a_3^{(1)} = a_3^{(1)}(\vartheta,x;\omega)$ and $a_1^{(1)} = a_1^{(1)}(\vartheta,x;\omega)$ are real valued and satisfy

$$\|a_3^{(1)} + 1\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon(1 + \|\iota\|_{s+\sigma_M}), \quad \|a_1^{(1)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}, \quad \forall s \geqq s_0, \tag{6.24}$$

and

$$\int_{\mathbb{T}_1}\frac{dx}{(a_3^{(1)}(\vartheta,x;\omega))^{\frac{1}{3}}} = m_3^{-\frac{1}{3}}, \quad \forall\vartheta\in\mathbb{T}^{\mathbb{S}_+}. \tag{6.25}$$

The function $r_0^{(1)} \equiv r_0^{(1)}(\vartheta,x,\xi;\omega)$ in (6.23) is a pseudo-differential symbol in the symbol class S^0 (cf. (2.8)) and admits an expansion of the form

$$r_0^{(1)}(\vartheta,x,\xi;\omega) = \sum_{k=0}^{M} a_{-k}^{(1)}(\vartheta,x;\omega)(\mathrm{i}2\pi\xi)^{-k}\chi_0(\xi), \tag{6.26}$$

where χ_0 is defined in (2.18) and where for any $0 \le k \le M$, $s \geqq s_0$,

$$\|a_{-k}^{(1)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}. \tag{6.27}$$

Furthermore, the function ρ appearing in (6.22) satisfies

$$\|\rho - 1\|_s^{\mathrm{Lip}(\gamma)}, \ \|\rho^{-1} - 1\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}. \tag{6.28}$$

Let $s_1 \geqq s_0$ *and let* $\breve{\iota}_1, \breve{\iota}_2$ *be torus embeddings, satisfying* (6.1) *with* $\mu_0 \geqq s_1 + \sigma_M$. *Then*

$$\begin{aligned}&|\Delta_{12} m_3|, \quad \|\Delta_{12}\alpha^{(1)}\|_{s_1}, \ \|\Delta_{12}\breve{\alpha}^{(1)}\|_{s_1}, \ \|\Delta_{12} a_1^{(1)}\|_{s_1}, \\ &\quad \|\Delta_{12}\rho^{\pm 1}\|_{s_1} \lesssim_{s_1, M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}, \\ &\|\Delta_{12} a_{-k}^{(1)}\|_{s_1} \lesssim_{s_1, M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}, \quad \forall k = 0, \ldots, M.\end{aligned} \tag{6.29}$$

(iii) *Let* $S > \mathfrak{s}_M$ *where* $\mathfrak{s}_M$ *is defined in* (2.55)*. Then the maps* $(\Phi^{(1)})^{\pm 1}$ *are* Lip(γ)*-1-tame operators with tame constants satisfying*

$$\mathfrak{M}_{(\Phi^{(1)})^{\pm 1}}(s) \lesssim_{S,M} 1 + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M}, \quad \forall s_0 + 1 \leqq s \leqq S. \tag{6.30}$$

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* $\lambda \leqq \lambda_0$ *and* $n_1 + n_2 + \lambda_0 \leqq M + 1$*, the operator* $\partial^{\lambda}_{\varphi_m}\langle D\rangle^{n_1}\mathcal{R}_M^{(1)}\langle D\rangle^{n_2}$ *is* Lip(γ)*-tame with a tame constant satisfying*

$$\mathfrak{M}_{\partial^{\lambda}_{\varphi_m}\langle D\rangle^{n_1}\mathcal{R}_M^{(1)}\langle D\rangle^{n_2}}(s) \lesssim_{S,M} \varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M(\lambda_0)}, \quad \forall \mathfrak{s}_M \leqq s \leqq S. \tag{6.31}$$

If in addition, $s_1 \geqq \mathfrak{s}_M$ *and* $\breve{\iota}_1, \breve{\iota}_2$ *are torus embeddings, satisfying* (6.1) *with* $\mu_0 \geqq s_1 + \sigma_M(\lambda_0)$*, then*

$$\|\partial^{\lambda}_{\varphi_m}\langle D\rangle^{n_1}\Delta_{12}\mathcal{R}_M^{(1)}\langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1, M, \lambda_0} \|\iota_1 - \iota_2\|_{s_1+\sigma_M(\lambda_0)}. \tag{6.32}$$

Proof. Writing $\Pi_\perp$ as Id $+ (\Pi_\perp - \mathrm{Id})$, the operator $\mathcal{L}_\omega^{(0)}$ (cf. (6.10)) becomes

$$\begin{aligned}\mathcal{L}_\omega^{(0)} &= \omega \cdot \partial_\varphi - \Big(a_3^{(0)}\partial_x^3 + 2(a_3^{(0)})_x \partial_x^2 + a_1^{(0)}\partial_x + \mathrm{Op}(r_0^{(0)}) + Q_{-1}^{kdv}(D;\omega)\Big) \\ &\quad + \mathcal{R}_M^{(I)}(\breve{\iota}_\delta(\varphi);\omega) + \mathcal{R}_M^{(0)}(\breve{\iota}_\delta(\varphi);\omega),\end{aligned}$$

where $\mathcal{R}_M^{(I)}(\breve{\iota}_\delta(\varphi);\omega) := (\mathrm{Id} - \Pi_\perp)\big(a_3^{(0)}\partial_x^3 + 2(a_3^{(0)})_x\partial_x^2 + a_1^{(0)}\partial_x + \mathrm{Op}(r_0^{(0)})\big)$. Since $(\mathrm{Id} - \Pi_\perp)\partial_x^3 h = 0$ for any $h \in H^s_\perp$, the operator $\mathcal{R}_M^{(I)} = \mathcal{R}_M^{(I)}(\breve{\iota}_\delta(\varphi);\omega)$ can be written as

$$\mathcal{R}_M^{(I)} = (\mathrm{Id} - \Pi_\perp)\Big((a_3^{(0)} + 1)\partial_x^3 + 2(a_3^{(0)})_x\partial_x^2 + a_1^{(0)}\partial_x + \mathrm{Op}(r_0^{(0)})\Big), \tag{6.33}$$

and is a finite rank operator of the form (6.3) with functions $g_j, \chi_j \in H^s_\perp$ satisfying (6.4) (use (6.11), (6.13)).

The estimate (6.30) follows by Lemma 2.1-(iii) and (6.21). Note 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, that is, $\Phi^{(1)} g(D)(\Phi^{(1)})^{-1} = g(D)$. Using (6.7) and (6.10), we obtain (6.23) where

$$a_3^{(1)} := \Phi^{(1)}\Big(\frac{a_3^{(0)}}{1 + \omega \cdot \partial_\varphi \breve{\alpha}^{(1)}}\Big), \tag{6.34}$$

$a_1^{(1)} := \frac{1}{\rho}\Phi^{(1)}(a_1^{(0)})$, $r_0^{(1)}$ is of the form (6.26) with $a_{-k}^{(1)} := \frac{1}{\rho}\Phi^{(1)}(a_{-k}^{(0)})$, and the remainder $\mathcal{R}_M^{(1)}$ is given by

$$\begin{aligned}\mathcal{R}_M^{(1)} &= \frac{1}{\rho}\Phi^{(1)}\mathcal{R}_M^{(I)}(\Phi^{(1)})^{-1} + \frac{1}{\rho}\Phi^{(1)}\mathcal{R}_M^{(0)}(\breve{\imath}_\delta(\varphi))(\Phi^{(1)})^{-1} \\ &\quad - \frac{1}{\rho}\Phi^{(1)}R(\varphi)(\Phi^{(1)})^{-1}. \end{aligned} \tag{6.35}$$

We choose $\breve{\alpha}^{(1)}$ such that (6.25) holds, yielding (6.18), (6.19). We now verify the estimates, stated in items (i) and (ii). Recall that we assume throughout that (6.1) holds. The estimates (6.20)–(6.21) follow by (6.18), (6.19), (6.11), and by using Lemma 2.1-(iii), Lemma 2.3. The estimate (6.28) on ρ follows by the definition (6.22), (6.19), and by applying Lemma 2.1-(iii), Lemma 2.3. Hence, by Lemma 2.1 and the estimates (6.11), (6.13), and (6.28), we deduce (6.27). The estimates (6.29) are obtained by similar arguments. Let us now prove item (iii). The estimate (6.30) follows from Lemma 2.1-(iii). Since $(\Phi^{(1)})^{\pm 1}$ commutes with every Fourier multiplier, we get

$$\frac{1}{\rho}\langle D\rangle^{n_1}\Phi^{(1)}\mathcal{R}_M^{(0)}(\breve{\imath}_\delta(\varphi))(\Phi^{(1)})^{-1}\langle D\rangle^{n_2} = \frac{1}{\rho}\langle D\rangle^{n_1}\mathcal{R}_M^{(0)}(\breve{\imath}_{\delta,\alpha}(\varphi))\langle D\rangle^{n_2} \tag{6.36}$$

where $\breve{\imath}_{\delta,\alpha}(\varphi) := \breve{\imath}_\delta(\varphi + \alpha^{(1)}(\varphi)\omega)$. By Lemma 2.1, (5.10), and (6.21) one has $\|\iota_{\delta,\alpha}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}$. Moreover, by (6.3), we have

$$\frac{1}{\rho}\Phi^{(1)}R(\varphi)(\Phi^{(1)})^{-1}h = \sum\nolimits_{j\in\mathbb{S}_+}\big(h,\ (\Phi^{(1)}g_j)\big)_{L_x^2}\frac{1}{\rho}(\Phi^{(1)}\chi_j),\ \forall h \in L_\perp^2, \tag{6.37}$$

and by (6.33), the conjugated operator $\frac{1}{\rho}\Phi^{(1)}\mathcal{R}_M^{(I)}(\Phi^{(1)})^{-1}h$ has the same form. The estimates (6.31) are then obtained by using (6.36), (6.14), and Lemmata 3.5, 3.7, 2.26 to estimate the first term on the right hand side of (6.35) and by (6.37), (6.30), (6.4) and Lemma 2.24, to estimate the second and third term in (6.35). The estimates (6.32) are proved by similar arguments. □

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 (6.22)–(6.23), where we rename ϑ to φ. Actually this step will at the same time also remove the coefficient of ∂_x^2 thanks to the Hamiltonian nature of the subprincipal operator of order 2, described in Lemma 2.7. We achieve these goals by conjugating the operator $\mathcal{L}_\omega^{(1)}$ by the flow $\Phi^{(2)}(\tau, \varphi)$, acting on $L_\perp^2(\mathbb{T}_1)$, defined by the transport equation

$$\partial_\tau\Phi^{(2)}(\tau,\varphi) = \Pi_\perp\partial_x\big(b^{(2)}(\tau,\varphi,x)\Phi^{(2)}(\tau,\varphi)\big), \quad \Phi^{(2)}(0,\varphi) = \mathrm{Id}_\perp, \tag{6.38}$$

where

$$b^{(2)} \equiv b^{(2)}(\tau,\varphi,x) := \frac{\beta^{(2)}(\varphi,x)}{1+\tau\beta_x^{(2)}(\varphi,x)},$$

and $\beta^{(2)}(\varphi, x)$ is a small, real valued periodic function chosen in (6.40) below. The flow $\Phi^{(2)}(\tau, \varphi)$ is well defined for $0 \leq \tau \leq 1$ and satisfies the tame estimates provided in Lemma 2.28. 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 $\frac{1}{2}\int_{\mathbb{T}_1} b^{(2)}h^2\,\mathrm{d}x$), each $\Phi^{(2)}(\tau, \varphi)$, $0 \leq \tau \leq 1$, $\varphi \in \mathbb{T}^{\mathbb{S}_+}$ is a symplectic linear isomorphism of $H^s_\perp(\mathbb{T}_1)$. Therefore the time one conjugated operator,

$$\mathcal{L}^{(2)}_\omega := \Phi^{(2)}\mathcal{L}^{(1)}_\omega\big(\Phi^{(2)}\big)^{-1}, \quad \Phi^{(2)} := \Phi^{(2)}(1, \varphi), \tag{6.39}$$

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. (Reduction to constant coefficients of the third order term) *Let $\breve{\beta}^{(2)}(\varphi, y; \omega) \equiv \breve{\beta}^{(2)}(1, \varphi, y; \omega)$ be the real valued, periodic function*

$$\breve{\beta}^{(2)}(\varphi, y; \omega) := \partial_y^{-1}\Big(\frac{m_3^{1/3}}{(a_3^{(1)}(\varphi, y; \omega))^{1/3}} - 1\Big) \tag{6.40}$$

(which is well defined by (6.25)) and let $M \in \mathbb{N}$. Then there exists $\sigma_M > 0$ so that the following holds:

(i) *For any $s \geqq s_0$*

$$\|\beta^{(2)}\|_s^{\mathrm{Lip}(\gamma)}, \ \|\breve{\beta}^{(2)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon\big(1 + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M}\big). \tag{6.41}$$

(ii) *The Hamiltonian operator $\mathcal{L}^{(2)}_\omega$ in (6.39) admits an expansion of the form*

$$\mathcal{L}^{(2)}_\omega = \omega \cdot \partial_\varphi - \big(m_3\partial_x^3 + a_1^{(2)}\partial_x + \mathrm{Op}(r_0^{(2)}) + Q^{kdv}_{-1}(D; \omega)\big) + \mathcal{R}^{(2)}_M \tag{6.42}$$

where $a_1^{(2)} := a_1^{(2)}(\varphi, x; \omega)$ is a real valued, periodic function, satisfying

$$\|a_1^{(2)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M}. \tag{6.43}$$

The pseudo-differential symbol $r_0^{(2)} \equiv r_0^{(2)}(\varphi, x, \xi; \omega)$ is in S^0 and satisfies, for any $s \geqq s_0$, the estimate

$$|\mathrm{Op}(r_0^{(2)})|^{\mathrm{Lip}(\gamma)}_{0,s,0} \lesssim_{s,M} \varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M}. \tag{6.44}$$

Let $s_1 \geqq s_0$ and let $\breve{\iota}_1, \breve{\iota}_2$ be torus embeddings satisfying (6.1) with $\mu_0 \geqq s_1 + \sigma_M$. Then

$$\begin{aligned}&\|\Delta_{12}\beta^{(2)}\|_{s_1}, \ \|\Delta_{12}\breve{\beta}^{(2)}\|_{s_1}, \ \|\Delta_{12}a_1^{(2)}\|_{s_1}, \ |\Delta_{12}\mathrm{Op}(r_0^{(2)})|_{0,s_1,0} \\ &\quad \lesssim_{s_1,M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}.\end{aligned} \tag{6.45}$$

(iii) *Let $S > \mathfrak{s}_M$. Then the symplectic maps $(\Phi^{(2)})^{\pm 1}$ are* Lip(γ)*-1 tame operators with tame constants satisfying*

$$\mathfrak{M}_{(\Phi^{(2)})^{\pm 1}}(s) \lesssim_{S,M} 1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}, \quad \forall s_0 + 1 \leqq s \leqq S. \tag{6.46}$$

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 $\lambda \leqq \lambda_0$ and $n_1 + n_2 + \lambda_0 \leqq M - 1$, 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 Lip(γ)*-tame with a tame constant satisfying*

$$\mathfrak{M}_{\partial_{\varphi_m}^{\lambda} \langle D\rangle^{n_1} \mathcal{R}_M^{(2)} \langle D\rangle^{n_2}}(s) \lesssim_{S,M,\lambda_0} \varepsilon + \|\iota\|_{s+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)}, \quad \forall \mathfrak{s}_M \leqq s \leqq S. \tag{6.47}$$

Let $s_1 \geqq \mathfrak{s}_M$ and $\breve{\iota}_1$, $\breve{\iota}_2$ be torus embeddings satisfying (6.1) with $\mu_0 \geqq s_1 + \sigma_M(\lambda_0)$. Then

$$\|\partial_{\varphi_m}^{\lambda} \langle D\rangle^{n_1} \Delta_{12} \mathcal{R}_M^{(2)} \langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,M,\lambda_0} \|\iota_1 - \iota_2\|_{s_1+\sigma_M(\lambda_0)}. \tag{6.48}$$

Proof. We use the Egorov type results proved in Section 2.5. According to (6.23), (6.26), the conjugated operator is given by

$$\begin{aligned}
\mathcal{L}_\omega^{(2)} &= \Phi^{(2)} \mathcal{L}_\omega^{(1)} (\Phi^{(2)})^{-1} \qquad (6.49)\\
&= \omega \cdot \partial_\varphi - \Phi^{(2)} a_3^{(1)} \partial_x^3 (\Phi^{(2)})^{-1} - 2\Phi^{(2)} (a_3^{(1)})_x \partial_x^2 (\Phi^{(2)})^{-1} - \Phi^{(2)} a_1^{(1)} \partial_x (\Phi^{(2)})^{-1}\\
&\quad - \sum_{k=0}^{M} \Phi^{(2)} a_{-k}^{(1)} \partial_x^{-k} (\Phi^{(2)})^{-1} - \Phi^{(2)} Q_{-1}^{kdv}(D;\omega)(\Phi^{(2)})^{-1}\\
&\quad + \Phi^{(2)} \mathcal{R}_M^{(1)} (\Phi^{(2)})^{-1} + \Phi^{(2)} \big(\omega \cdot \partial_\varphi \ (\Phi^{(2)})^{-1}\big).
\end{aligned}$$

By (6.40), (6.20), (6.24) and Lemmata 2.1, 2.3, the estimate (6.41) follows. Using (6.1) with $\mu_0 > 0$ large enough, the estimate (6.41) 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 (2.79), (2.107), (2.110). Now we apply Proposition 2.31 to expand the terms

$$\Phi^{(2)} a_3^{(1)} \partial_x^3 (\Phi^{(2)})^{-1}, \quad 2\Phi^{(2)} (a_3^{(1)})_x \partial_x^2 (\Phi^{(2)})^{-1},$$
$$\Phi^{(2)} a_{1-k}^{(1)} \partial_x^{1-k} (\Phi^{(2)})^{-1}, \quad 0 \le k \le M+1,$$

Lemma 2.35 to expand the term $\Phi^{(2)} Q_{-1}^{kdv}(D;\omega)(\Phi^{(2)})^{-1}$, and Proposition 2.34 to expand $\Phi^{(2)}\big(\omega \cdot \partial_\varphi \ (\Phi^{(2)})^{-1}\big)$. Using also the bounds (6.11), (6.13), (6.41) one deduces (6.43), (6.44). By the choice of $\breve{\beta}^{(2)}$ in (6.40) and Proposition 2.31, the coefficient of the highest order term of $\Phi^{(2)} a_3^{(1)} \partial_x^3 (\Phi^{(2)})^{-1}$ (and hence 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 (6.25). Since $\Phi^{(2)}$ is symplectic, the operator $\mathcal{L}_\omega^{(2)}$ is Hamiltonian and hence by Lemma 2.7, 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 2.31 (estimate of $\mathcal{R}_N(\tau, \varphi)$), using Lemma 2.28 to estimate $\Phi^{(2)}$, $(\Phi^{(2)})^{-1}$, the estimate (6.31) for $\mathcal{R}_M^{(1)}$, the estimate (6.41) of $\beta^{(2)}$, $\breve{\beta}^{(2)}$, and (6.1) with μ_0 large enough. The estimates (6.46) follow from (2.73) and (6.41). The bounds (6.45), (6.48) are derived by similar arguments. □

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 (6.39), (6.42). To this end, we conjugate the operator $\mathcal{L}_\omega^{(2)}$ by the variable transformation induced by the flow $\Phi^{(3)}(\tau, \varphi)$, acting on $L^2_\perp(\mathbb{T}_1)$, defined by

$$\partial_\tau \Phi^{(3)}(\tau, \varphi) = \Pi_\perp\big(b^{(3)}(\varphi, x)\partial_x^{-1}\Phi^{(3)}(\tau, \varphi)\big), \quad \Phi^{(3)}(0) = \mathrm{Id}_\perp, \tag{6.50}$$

where $b^{(3)}(\varphi, x)$ is a small, real valued, periodic function, chosen in (6.52) 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 $\frac{1}{2}\int_{\mathbb{T}_1} b^{(3)}(\partial_x^{-1}h)^2\, \mathrm{d}x$), 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,

$$\mathcal{L}_\omega^{(3)} := \Phi^{(3)}\mathcal{L}_\omega^{(2)}\big(\Phi^{(3)}\big)^{-1}, \quad \Phi^{(3)} := \Phi^{(3)}(1), \tag{6.51}$$

is Hamiltonian.

Lemma 6.6. *Let $b^{(3)}(\varphi, x; \omega)$ be the real valued periodic function*

$$b^{(3)}(\varphi, x; \omega) := \frac{1}{3m_3}\partial_x^{-1}\Big(a_1^{(2)}(\varphi, x; \omega) - \langle a_1^{(2)}\rangle_x(\varphi; \omega)\Big),$$
$$\langle a_1^{(2)}\rangle_x(\varphi; \omega) := \int_{\mathbb{T}_1} a_1^{(2)}(\varphi, x; \omega)\, dx \tag{6.52}$$

and let $M \in \mathbb{N}$. Then there exists $\sigma_M > 0$ with the following properties:

(i) *For any $s \geqq s_0$,*

$$\|b^{(3)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)} \tag{6.53}$$

and the symplectic maps $(\Phi^{(3)})^{\pm 1}$ are $\mathrm{Lip}(\gamma)$-tame and with tame constants satisfying

$$\mathfrak{M}_{(\Phi^{(3)})^{\pm 1}}(s) \lesssim_{s,M} 1 + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}. \tag{6.54}$$

(ii) *The Hamiltonian operator in (6.51) admits an expansion of the form*

$$\mathcal{L}_\omega^{(3)} = \omega \cdot \partial_\varphi - \big(m_3\partial_x^3 + a_1^{(3)}(\varphi)\partial_x + \mathrm{Op}(r_0^{(3)}) + Q_{-1}^{kdv}(D; \omega)\big) + \mathcal{R}_M^{(3)} \tag{6.55}$$

where the real valued, periodic function $a_1^{(3)}(\varphi; \omega) := \langle a_1^{(2)}\rangle_x(\varphi; \omega)$ satisfies

$$\|a_1^{(3)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}, \tag{6.56}$$

and $r_0^{(3)} := r_0^{(3)}(\varphi, x, \xi; \omega)$ *is a pseudo-differential symbol in* S^0 *satisfying for any* $s \geqq s_0$,

$$|\mathrm{Op}(r_0^{(3)})|_{0,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}. \tag{6.57}$$

Let $s_1 \geqq s_0$ *and let* $\breve{\iota}_1, \breve{\iota}_2$ *be torus embeddings satisfying* (6.1) *with* $\mu_0 \geqq s_1+\sigma_M$. *Then*

$$\begin{gathered}\|\Delta_{12} b^{(3)}\|_{s_1}, \|\Delta_{12} a_1^{(3)}\|_{s_1} \lesssim_{s_1,M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}, \\ |\Delta_{12}\mathrm{Op}(r_0^{(3)})|_{0,s_1,0} \lesssim_{s_1,M} \|\iota_1 - \iota_2\|_{s_1+\sigma_M}.\end{gathered} \tag{6.58}$$

(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* $\lambda \leqq \lambda_0$ *and* $n_1 + n_2 + \lambda_0 \leqq M - 1$, *the operator* $\langle D\rangle^{n_1}\partial_{\varphi_m}^{\lambda}\mathcal{R}_M{}^{(3)}\langle D\rangle^{n_2}$, *is* Lip$(\gamma)$*-tame with tame constant satisfying*

$$\mathfrak{M}_{\partial_{\varphi_m}^{\lambda}\langle D\rangle^{n_1}\mathcal{R}_M^{(3)}\langle D\rangle^{n_2}}(s) \lesssim_{S,M,\lambda_0} \varepsilon + \|\iota\|_{s+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)}, \quad \forall \mathfrak{s}_M \leqq s \leqq S. \tag{6.59}$$

Let $s_1 \geqq \mathfrak{s}_M$ *and let* $\breve{\iota}_1, \breve{\iota}_2$ *be torus embeddings satisfying* (6.1) *with* $\mu_0 \geqq s_1 + \sigma_M(\lambda_0)$. *Then*

$$\|\partial_{\varphi_m}^{\lambda}\langle D\rangle^{n_1}\Delta_{12}\mathcal{R}_M^{(3)}\langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,M,\lambda_0} \|\iota_1 - \iota_2\|_{s_1+\sigma_M(\lambda_0)}. \tag{6.60}$$

Proof. The estimate (6.53) follows by the definition (6.52) and (6.43), (6.20). 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 2.10, and (6.53), one infers that for any $s \geqq s_0$, $|\Pi_\perp b^{(3)}\partial_x^{-1}|_{-1,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}$. Therefore, by Lemma 2.13, there exists $\sigma_M > 0$ such that, if (6.1) holds with $\mu_0 \geqq \sigma_M$, then, for any $s \geqq s_0$,

$$\sup_{\tau\in[-1,1]} |\Phi^{(3)}(\tau) - \mathrm{Id}|_{0,s,0}^{\mathrm{Lip}(\gamma)} \lesssim_s \varepsilon + \|\iota\|_{s+\sigma_M}^{\mathrm{Lip}(\gamma)}. \tag{6.61}$$

The latter estimate, together with Lemma 2.18, imply (6.54).

By (6.42) and using Lemma 6.3 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

$$\begin{aligned}\mathcal{R}_0^{(I)} := & -\Phi^{(3)}\mathrm{Op}(r_0^{(2)})(\Phi^{(3)})^{-1} + \Phi^{(3)}\big(\omega\cdot\partial_\varphi(\Phi^{(3)})^{-1}\big) \\ & - (\Phi^{(3)} - \mathrm{Id}_\perp)\Pi_\perp\Big(\sum_{k=1}^{M} c_{-k}^{kdv}(\omega)\partial_x^{-k}\Big)(\Phi^{(3)})^{-1} \\ & - \Pi_\perp\Big(\sum_{k=1}^{M} c_{-k}^{kdv}(\omega)\partial_x^{-k}\Big)\Big((\Phi^{(3)})^{-1} - \mathrm{Id}_\perp\Big), \\ \mathcal{R}_M^{(3)} := & \ \Phi^{(3)}\mathcal{R}_M^{(2)}(\Phi^{(3)})^{-1} - (\Phi^{(3)} - \mathrm{Id}_\perp)\mathcal{R}_M(\omega, Q_{-1}^{kdv})(\Phi^{(3)})^{-1} \\ & - \mathcal{R}_M(\omega, Q_{-1}^{kdv})\big((\Phi^{(3)})^{-1} - \mathrm{Id}_\perp\big).\end{aligned} \tag{6.62}$$

Note that $\mathcal{R}_0^{(I)}$ is a pseudo-differential operator in OPS^0 (cf. Lemma 2.13). Moreover, by a Lie expansion and using (6.50), one has

$$\begin{aligned}
&\Phi^{(3)}\big(m_3\partial_x^3 + a_1^{(2)}\partial_x\big)(\Phi^{(3)})^{-1} = m_3\partial_x^3 + a_1^{(2)}\partial_x + [\Pi_\perp b^{(3)}\partial_x^{-1},\, m_3\partial_x^3 + a_1^{(2)}\partial_x] \\
&\quad + \int_0^1 (1-\tau)\Phi^{(3)}(\tau)\Big[\Pi_\perp b^{(3)}\partial_x^{-1}, \Big[\Pi_\perp b^{(3)}\partial_x^{-1}, m_3\partial_x^3 + a_1^{(2)}\partial_x\Big]\Big]\Phi^{(3)}(\tau)^{-1}\,\mathrm{d}\tau \\
&\quad = m_3\partial_x^3 + \big(a_1^{(2)} - 3m_3 b_x^{(3)}\big)\partial_x + \mathcal{R}_0^{(II)}, \\
&\mathcal{R}_0^{(II)} := -3m_3 b_{xx}^{(3)} - m_3 b_{xxx}^{(3)}\partial_x^{-1} + [\Pi_\perp b^{(3)}\partial_x^{-1}, a_1^{(2)}\partial_x] + [(\Pi_\perp - \mathrm{Id})b^{(3)}\partial_x^{-1}, m_3\partial_x^3] \\
&\quad + \int_0^1 (1-\tau)\Phi^{(3)}(\tau)\Big[\Pi_\perp b^{(3)}\partial_x^{-1},\, \Big[\Pi_\perp b^{(3)}\partial_x^{-1}, m_3\partial_x^3 + a_1^{(2)}\partial_x\Big]\Big]\Phi^{(3)}(\tau)^{-1}\,\mathrm{d}\tau \in OPS^0.
\end{aligned} \tag{6.63}$$

Note that $\mathcal{R}_0^{(II)}$ is also a pseudo-differential operator in OPS^0 (cf. Lemma 2.13). Hence, (6.62)–(6.63) and the choice of $b^{(3)}$ in (6.52) lead to the expansion (6.55) with $\mathcal{R}_M^{(3)}$ given by (6.62) and

$$\mathrm{Op}(r_0^{(3)}) := -\,\mathcal{R}_0^{(I)} + \mathcal{R}_0^{(II)}. \tag{6.64}$$

The estimate (6.56) follows by (6.24).

The estimate (6.57) on the operator $\mathrm{Op}(r_0^{(3)})$ follows by the definitions (6.62), (6.63), (6.64), by applying the estimates (6.20), (6.43), (6.44), (6.53), (6.61), (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 (6.62). We only consider the second term in the definition of $\mathcal{R}_M^{(3)}$, since the estimates of 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 $\lambda \leqq \lambda_0$ and $n_1 + n_2 + \lambda_0 \leqq M - 2$, one has

$$\begin{aligned}
&\langle D\rangle^{n_1}\partial_{\varphi_m}^{\lambda}\Big((\Phi^{(3)} - \mathrm{Id}_\perp)\mathcal{R}_M(Q_{-1}^{kdv};\omega)(\Phi^{(3)})^{-1}\Big)\langle D\rangle^{n_2} \\
&\quad = \sum_{\lambda_1+\lambda_2=\lambda} C_{\lambda_1,\lambda_2}\langle D\rangle^{n_1}\partial_{\varphi_m}^{\lambda_1}(\Phi^{(3)} - \mathrm{Id}_\perp)\mathcal{R}_M(Q_{-1}^{kdv};\omega)\partial_{\varphi_m}^{\lambda_2}(\Phi^{(3)})^{-1}\langle D\rangle^{n_2} \\
&\quad = \sum_{\lambda_1+\lambda_2=\lambda} C_{\lambda_1,\lambda_2}\Big(\langle D\rangle^{n_1}\partial_{\varphi_m}^{\lambda_1}(\Phi^{(3)} - \mathrm{Id}_\perp)\langle D\rangle^{-n_1}\Big) \\
&\qquad \Big(\langle D\rangle^{n_1}\mathcal{R}_M(Q_{-1}^{kdv};\omega)\langle D\rangle^{n_2}\Big)\Big(\langle D\rangle^{-n_2}\partial_{\varphi_m}^{\lambda_2}(\Phi^{(3)})^{-1}\langle D\rangle^{n_2}\Big).
\end{aligned} \tag{6.65}$$

By the estimates (2.21), (2.24), (6.61) and Lemma 2.18, 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 2.16, 6.16 and using (6.1), the operator (6.65) satisfies (6.59). The estimates (6.58), (6.60) 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 (6.51), (6.55). 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 (6.67) below. Note 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 $H^s_\perp(\mathbb{T}_1)$, and hence the conjugated operator

$$\mathcal{L}_\omega^{(4)} := \Phi^{(4)}\mathcal{L}_\omega^{(3)}\big(\Phi^{(4)}\big)^{-1} \tag{6.66}$$

is Hamiltonian.

Proposition 6.7. (Reduction of $\mathcal{L}_\omega$ to constant coefficients up to order zero) *Assume that $\omega \in \mathtt{DC}(\gamma,\tau)$ (cf. (4.4)). Let $b^{(4)}(\varphi)$ be the real valued, periodic function*

$$b^{(4)}(\varphi;\omega) := -(\omega\cdot\partial_\varphi)^{-1}\big(a_1^{(3)}(\varphi;\omega) - m_1\big), \quad m_1 := \frac{1}{(2\pi)^{|\mathbb{S}_+|}}\int_{\mathbb{T}^{\mathbb{S}_+}} a_1^{(3)}(\varphi;\omega)\,d\varphi \tag{6.67}$$

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*

$$|m_1|^{\mathrm{Lip}(\gamma)} \lesssim_M \varepsilon\gamma^{-2}, \quad \|b^{(4)}\|_s^{\mathrm{Lip}(\gamma)} \lesssim_{s,M} \gamma^{-1}\big(\varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M}\big), \quad \forall s \geqq s_0. \tag{6.68}$$

(ii) *The Hamiltonian operator in (6.66) admits an expansion of the form*

$$\mathcal{L}_\omega^{(4)} = \omega\cdot\partial_\varphi - \big(m_3\partial_x^3 + m_1\partial_x + \mathrm{Op}(r_0^{(4)}) + Q_{-1}^{kdv}(D;\omega)\big) + \mathcal{R}_M^{(4)} \tag{6.69}$$

where m_3, given by (6.18), satisfies $|m_3+1|^{\mathrm{Lip}(\gamma)} \lesssim_M \varepsilon$ (cf. (6.20)), and $r_0^{(4)} := r_0^{(4)}(\varphi,x,\xi;\omega)$ is a pseudo-differential symbol in S^0 satisfying

$$|\mathrm{Op}(r_0^{(4)})|^{\mathrm{Lip}(\gamma)}_{0,s,0} \lesssim_{s,M} \varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M}, \quad \forall s\geqq s_0. \tag{6.70}$$

Let $s_1 \geqq s_0$ and let $\breve{\iota}_1, \breve{\iota}_2$ be torus embeddings satisfying (6.1) with $\mu_0 \geqq s_1+\sigma_M$. Then

$$\begin{aligned} &|\Delta_{12}m_1|,\ \|\Delta_{12}b^{(4)}\|_{s_1} \lesssim_{s_1,M} \|\iota_1-\iota_2\|_{s_1+\sigma_M}, \\ &|\Delta_{12}\mathrm{Op}(r_0^{(4)})|_{0,s_1,0} \lesssim_{s_1,M} \|\iota_1-\iota_2\|_{s_1+\sigma_M}. \end{aligned} \tag{6.71}$$

(iii) *Let $S > \mathfrak{s}_M$. Then the maps $(\Phi^{(4)})^{\pm1}$ are* Lip(γ)*-tame operators with tame constants satisfying*

$$\mathfrak{M}_{(\Phi^{(4)})^{\pm1}}(s) \lesssim_{S,M} 1 + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_M}, \quad \forall s_0 \leqq s \leqq S. \tag{6.72}$$

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 $\lambda \leqq \lambda_0$ and $n_1 + n_2 + 2\lambda_0 \leqq M - 3$, 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*

$$\mathfrak{M}_{\partial_{\varphi_m}^{\lambda} \langle D\rangle^{n_1} \mathcal{R}_M^{(4)} \langle D\rangle^{n_2}}(s) \lesssim_{S,M,\lambda_0} \varepsilon + \|\iota\|_{s+\sigma_M(\lambda_0)}^{\mathrm{Lip}(\gamma)}, \quad \forall \mathfrak{s}_M \leqq s \leqq S. \tag{6.73}$$

Let $s_1 \geqq \mathfrak{s}_M$ and let $\breve{\iota}_1, \breve{\iota}_2$ be two tori satisfying (6.1) with $\mu_0 \geqq s_1 + \sigma_M(\lambda_0)$. Then

$$\|\partial_{\varphi_m}^{\lambda} \langle D\rangle^{n_1} \Delta_{12} \mathcal{R}_M^{(4)} \langle D\rangle^{n_2}\|_{\mathcal{B}(H^{s_1})} \lesssim_{s_1,M,\lambda_0} \|\iota_1 - \iota_2\|_{s_1+\sigma_M(\lambda_0)}. \tag{6.74}$$

Proof. The estimates (6.68) are direct consequences of (6.56) and of (6.1). Note that

$$\Phi^{(4)} \circ \omega \cdot \partial_\varphi \circ (\Phi^{(4)})^{-1} = \omega \cdot \partial_\varphi - \big(\omega \cdot \partial_\varphi b^{(4)}\big)\partial_x.$$

A straightforward calculation then shows that for any pseudo-differential operator $\mathrm{Op}(a(\varphi, x, \xi))$,

$$\Phi^{(4)} \mathrm{Op}(a(\varphi, x, \xi))\big(\Phi^{(4)}\big)^{-1} = \mathrm{Op}(a(\varphi, x + b^{(4)}(\varphi), \xi)).$$

Hence, by (6.55) and the definition (6.67), one obtains (6.69) with

$$\mathrm{Op}\big(r_0^{(4)}(\varphi, x, \xi)\big) = \mathrm{Op}\big(r_0^{(3)}(\varphi, x + b^{(4)}(\varphi), \xi)\big), \quad \mathcal{R}_M^{(4)} := \Phi^{(4)} \mathcal{R}_M^{(3)} (\Phi^{(4)})^{-1}. \tag{6.75}$$

The estimates (6.70) for $\mathrm{Op}(r_0^{(4)})$ follow from Lemma 2.1, using (6.68), (6.57), and (6.1). The estimates (6.73) for the operator $\mathcal{R}_M^{(4)}$ ared obtained from (6.59), (6.68) 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 2.31. The estimates (6.72) follow by Lemma 2.1 and (6.68). Finally, the estimates (6.71), (6.74) are obtained by similar arguments. □

7. KAM Reduction of the Linearized Operator

The main result of this section is Theorem 7.11, stating that the assumptions **A-I** concerning the almost-invertibility of $\mathcal{L}_\omega$ in Section 5 are satisfied. The key ingredient for its proof is Theorem 7.3, which affirms that the Hamiltonian operator $\mathcal{L}_\omega^{(4)}$ in (6.69), renamed $\mathtt{L}_0$ in (7.1), can be brought in almost diagonal form. This completes the diagonalization of the Hamiltonian operator $\mathcal{L}_\omega$, defined in (6.7), which was started in Section 6. The Hamiltonian operator $\mathtt{L}_0$ is diagonalized by applying a KAM-reducibility iterative scheme, developed in [10], for perturbations of diagonal operators which are modulo-tame (cf. Definition 2.19). In Lemma 7.1 we prove that the initial remainder $\mathtt{R}_0$ is indeed modulo-tame. We recall that the class of modulo-tame operators is closed under the operations coming up in the KAM reduction procedure, namely: sum and composition (Lemma 2.21); projections (Lemma 2.23); solution of the homological equation (Lemma 7.5).

As in Section 6, we again consider in the sequel torus embeddings $\breve{\imath}(\varphi) = (\varphi, 0, 0) + \imath(\varphi)$ with $\imath(\cdot) \equiv \imath(\cdot\,;\omega)$, $\omega \in \mathtt{DC}(\gamma, \tau)$ (cf. (4.4)), satisfying (6.1), $\|\imath\|_{\mu_0}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}$, $\varepsilon\gamma^{-2} \leqq \delta(S)$, where $\mu_0 := \mu_0(\tau, \mathbb{S}_+) > s_0$ and $S > \mu_0$ are sufficiently large, $0 < \delta(S) < 1$ is sufficiently small, and $0 < \gamma < 1$.

Recall that by (6.69), the operator $\mathcal{L}_\omega^{(4)}$ is given by $\omega \cdot \partial_\varphi - \big(m_3\partial_x^3 + m_1\partial_x + \mathrm{Op}(r_0^{(4)}) + Q_{-1}^{kdv}(D;\omega)\big) + \mathcal{R}_M^{(4)}$ and acts on $H_\perp^s$. In view of the reduction scheme, implemented in this section, it is useful to rename it to

$$\mathtt{L}_0 := \omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D}_0 + \mathtt{R}_0, \tag{7.1}$$

where $\omega \in \mathtt{DC}(\gamma, \tau)$ (cf. (4.4)) and, in view of (6.9), (3.13), (4.3),

$$\mathtt{D}_0 := \mathrm{diag}_{j\in\mathbb{S}^\perp}(\mu_j^0), \quad \mu_j^0 := m_3(2\pi j)^3 - m_1 2\pi j - q_j(\omega),$$

$$q_j(\omega) := \omega_j^{kdv}\big(\nu(\omega), 0\big) - (2\pi j)^3, \tag{7.2}$$

$$\mathtt{R}_0 := -\,\mathrm{Op}(r_0^{(4)}) + \mathcal{R}_M^{(4)}. \tag{7.3}$$

We recall that $m_3 : \Omega \to \mathbb{R}$ is given by (6.18) and $m_1 : \mathtt{DC}(\gamma, \tau) \to \mathbb{R}$ by (6.67). Furthermore, note that $\mu_{-j}^0 = -\mu_j^0$ for any $j \in \mathbb{S}^\perp$. By (3.67) we have

$$\sup_{j\in\mathbb{S}^\perp} |j||q_j|^{\sup}, \quad \sup_{j\in\mathbb{S}^\perp} |j||q_j|^{\mathrm{lip}} \lesssim 1, \tag{7.4}$$

and, by (6.20), (6.68) and $\varepsilon\gamma^{-3} \leqq 1$,

$$|\mu_j^0 - \mu_{j'}^0|^{\mathrm{lip}} \lesssim_M |j^3 - j'^3|, \quad \forall j, j' \in \mathbb{S}^\perp. \tag{7.5}$$

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 $[\mathtt{a}]$ denotes the integer part of $\mathtt{a}$ and the constants $\sigma_M, \sigma_M(\mathtt{b})$ are the ones introduced in Lemma 6.7. The integer M is related to the order of smoothing of the remainder $\mathcal{R}_M^{(4)}$ in (6.69) (cf. (6.73)). 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 *be given as in* (7.6) *and* $S > \mathfrak{s}_M$ *with* $\mathfrak{s}_M$ *given by* (2.55).

(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* $\mathrm{Lip}(\gamma)$*-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 $\mathfrak{s}_M \leqq s \leqq S$,

$$\mathfrak{M}_0(s, \mathtt{b}) := \max\{\mathbb{M}_0(s), \mathbb{M}_0(s, \mathtt{b})\} \lesssim_S \varepsilon + \|\imath\|_{s+\mu(\mathtt{b})}^{\mathrm{Lip}(\gamma)}. \tag{7.9}$$

Assuming that (6.1) *(ansatz for* $\breve{\imath}$*) holds with* $\mu_0 \geqq \mathfrak{s}_M + \mu(\mathtt{b})$, *the latter estimate yields* $\mathfrak{M}_0(\mathfrak{s}_M, \mathtt{b}) \lesssim_S \varepsilon\gamma^{-2}$.

(ii) *For any torus embeddings* $\breve{\imath}_1, \breve{\imath}_2$ *satisfying* (6.1), *one has for any* $m \in \mathbb{S}_+$ *and any* $\lambda \in \mathbb{N}$ *with* $\lambda \leqq s_0 + \mathtt{b}$

$$\|\partial_{\varphi_m}^{\lambda} \Delta_{12}\mathrm{R}_0\|_{\mathcal{B}(H^{\mathfrak{s}_M})}, \ \|\partial_{\varphi_m}^{\lambda} [\Delta_{12}\mathrm{R}_0, \partial_x]\|_{\mathcal{B}(H^{\mathfrak{s}_M})} \lesssim \|\iota_1 - \iota_2\|_{\mathfrak{s}_M+\mu(\mathtt{b})}. \tag{7.10}$$

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 $\mathfrak{M}_{\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathrm{R}_0, \partial_x]}(s)$, $m \in \mathbb{S}_+$, 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 R_0 are treated separately. By Lemma 2.18, for any $m \in \mathbb{S}_+$, the operator $\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathrm{Op}(r_0^{(4)}), \partial_x] = -\mathrm{Op}\big(\partial_{\varphi_m}^{s_0+\mathtt{b}}\partial_x r_0^{(4)}\big)$ is Lip(γ)-tame with a tame constant satisfying for any $s_0 \leqq s \leqq S$,

$$\begin{aligned}\mathfrak{M}_{\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathrm{Op}(r_0^{(4)}), \partial_x]}(s) &\overset{(2.31)}{\lesssim_s} \left|\mathrm{Op}\Big(\partial_{\varphi_m}^{s_0+\mathtt{b}}\partial_x r_0^{(4)}\Big)\right|_{0,s,0}^{\mathrm{Lip}(\gamma)} \\ &\lesssim_s \left|\mathrm{Op}(r_0^{(4)})\right|_{0,s+s_0+\mathtt{b}+1,0}^{\mathrm{Lip}(\gamma)} \\ &\overset{(6.70)}{\lesssim_s} \varepsilon + \|\iota\|_{s+s_0+\mathtt{b}+1+\sigma_M}^{\mathrm{Lip}(\gamma)}.\end{aligned} \tag{7.11}$$

Next consider, for any given $m \in \mathbb{S}_+$, the operator $\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathcal{R}_M^{(4)}, \partial_x]$. Recalling that $\langle D\rangle$ denotes the Fourier multiplier with symbol $\langle \xi \rangle$, one has

$$\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathcal{R}_M^{(4)}, \partial_x] = \partial_{\varphi_m}^{s_0+\mathtt{b}}\mathcal{R}_M^{(4)}\langle D\rangle\langle D\rangle^{-1}\partial_x - \langle D\rangle^{-1}\partial_x\langle D\rangle\partial_{\varphi_m}^{s_0+\mathtt{b}}\mathcal{R}_M^{(4)}.$$

Since $\langle D\rangle^{-1}\partial_x$ admits a tame constant $\mathfrak{M}_{\langle D\rangle^{-1}\partial_x}(s)$ bounded by 1, it follows by (6.73) that, for any $\mathfrak{s}_M \leqq s \leqq S$,

$$\mathfrak{M}_{\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathcal{R}_M^{(4)}, \partial_x]}(s) \lesssim_S \varepsilon + \|\iota\|_{s+\sigma_M(\mathtt{b})}^{\mathrm{Lip}(\gamma)}. \tag{7.12}$$

Combining (7.11), (7.12) and recalling the definition of $\mu(\mathtt{b})$ in (7.6) one infers that $\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathrm{R}_0, \partial_x]$ admits a tame constant $\mathfrak{M}_{\partial_{\varphi_m}^{s_0+\mathtt{b}}[\mathrm{R}_0, \partial_x]}(s)$, satisfying the claimed bound.

(ii) The estimate (7.10) follows by similar arguments using (6.71) and (6.74) with $s_1 = \mathfrak{s}_M$. □

We perform the almost reducibility scheme for L_0 along the scale

$$N_{-1} := 1, \quad N_\nu := N_0^{\chi^\nu}, \ \nu \geqq 0, \quad \chi := 3/2 \tag{7.13}$$

(with N_0 specified in Theorem 7.2 below), assuming that at each induction step the second order Melnikov non-resonance conditions (7.18) hold.

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* $N_0 := N_0(S, \mathtt{b}) \in \mathbb{N}$, *satisfying*

$$N_0^{\overline{\tau}}\mathfrak{M}_0(\mathfrak{s}_M, \mathtt{b})\gamma^{-1} \leqq 1, \tag{7.14}$$

so that 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*

$$\mathtt{L}_\nu := \omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D}_\nu + \mathtt{R}_\nu, \quad \mathtt{D}_\nu := \mathrm{diag}_{j \in \mathbb{S}^\perp} \mu^\nu_j, \quad \mu^\nu_j \in \mathbb{R}, \tag{7.15}$$

where for any $j \in \mathbb{S}^\perp$, μ^ν_j, *also denoted by* $\mu^\nu_j(\omega)$ *or* $\mu^\nu_j(\omega; \iota)$, *is a* Lip(γ)-*function of the form*

$$\mu^\nu_j(\omega) := \mu^0_j(\omega) + r^\nu_j(\omega), \tag{7.16}$$

with $\mu^{(0)}_j$ *defined by* (7.2) *and*

$$\mu^\nu_{-j} = -\mu^\nu_j, \quad |r^\nu_j|^{\mathrm{Lip}(\gamma)} \leqq C(S)\varepsilon\gamma^{-2}. \tag{7.17}$$

If $\nu = 0$, *the set of frequency vectors* Ω^γ_ν *is defined to be the set* $\Omega^\gamma_0 := \mathtt{DC}(\gamma, \tau)$, *and if* $\nu \geqq 1$,

$$\Omega^\gamma_\nu = \Omega^\gamma_\nu(\iota) := \Big\{ \omega \in \Omega^\gamma_{\nu-1} \, : |\omega \cdot \ell + \mu^{\nu-1}_j - \mu^{\nu-1}_{j'}| \geqq \gamma \frac{|j^3 - j'^3|}{\langle \ell \rangle^\tau}, \ \forall |\ell| \leqq N_{\nu-1}, j, j' \in \mathbb{S}^\perp \Big\}. \tag{7.18}$$

Note that $\Omega^\gamma_{\nu+1} \subseteq \Omega^\gamma_\nu$ *for any* $\nu \geq 0$. *The operators* $\mathtt{R}_\nu$ *and* $\langle \partial_\varphi \rangle^{\mathtt{b}} \mathtt{R}_\nu$ *are* Lip(γ)-*modulo-tame with modulo-tame constants*

$$\mathfrak{M}^\sharp_\nu(s) := \mathfrak{M}^\sharp_{\mathtt{R}_\nu}(s), \quad \mathfrak{M}^\sharp_\nu(s, \mathtt{b}) := \mathfrak{M}^\sharp_{\langle \partial_\varphi \rangle^{\mathtt{b}} \mathtt{R}_\nu}(s)\,, \tag{7.19}$$

satisfying, for some $C_*(\mathfrak{s}_M, \mathtt{b}) > 0$ *and any* $s \in [\mathfrak{s}_M, S]$,

$$\mathfrak{M}^\sharp_\nu(s) \leqq C_*(\mathfrak{s}_M, \mathtt{b})\mathfrak{M}_0(s, \mathtt{b}) N^{-\mathtt{a}}_{\nu-1}, \quad \mathfrak{M}^\sharp_\nu(s, \mathtt{b}) \leqq C_*(\mathfrak{s}_M, \mathtt{b})\mathfrak{M}_0(s, \mathtt{b}) N_{\nu-1} \tag{7.20}$$

with $N_{\nu-1}$ *given by* (7.13). *Moreover, if* $\nu \geqq 1$ *and* $\omega \in \Omega^\gamma_\nu$, *there exists a Hamiltonian operator* $\Psi_{\nu-1}$ *acting on* $H^s_\perp$, *so that the corresponding symplectic time one flow*

$$\Phi_{\nu-1} := \exp(\Psi_{\nu-1}) \tag{7.21}$$

conjugates $\mathtt{L}_{\nu-1}$ *to*

$$\mathtt{L}_\nu = \Phi_{\nu-1} \mathtt{L}_{\nu-1} \Phi^{-1}_{\nu-1}. \tag{7.22}$$

The operators $\Psi_{\nu-1}$ *and* $\langle \partial_\varphi \rangle^{\mathtt{b}} \Psi_{\nu-1}$ *are* Lip(γ)-*modulo-tame with modulo-tame constants satisfying, for any* $s \in [\mathfrak{s}_M, S]$, *(with* τ_1, $\mathtt{a}$ *defined in* (7.6)*)*

$$\begin{aligned} \mathfrak{M}^\sharp_{\Psi_{\nu-1}}(s) &\leqq \frac{C(\mathfrak{s}_M, \mathtt{b})}{\gamma} N^{\tau_1}_{\nu-1} N^{-\mathtt{a}}_{\nu-2} \mathfrak{M}_0(s, \mathtt{b}), \\ \mathfrak{M}^\sharp_{\langle \partial_\varphi \rangle^{\mathtt{b}} \Psi_{\nu-1}}(s) &\leqq \frac{C(\mathfrak{s}_M, \mathtt{b})}{\gamma} N^{\tau_1}_{\nu-1} N_{\nu-2} \mathfrak{M}_0(s, \mathtt{b}). \end{aligned} \tag{7.23}$$

(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 a Lipschitz extension of* m_1, *satisfying* $|\widetilde{m}_1|^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}$ *(recall that by (6.18)* m_3 *is defined on* Ω*); for any* $\nu \geqq 1$,

$$|\widetilde{\mu}_j^\nu - \widetilde{\mu}_j^{\nu-1}|^{\mathrm{Lip}(\gamma)} \lesssim \mathfrak{M}_{\nu-1}^\sharp(\mathfrak{s}_M) \lesssim \mathfrak{M}_0(\mathfrak{s}_M, \mathrm{b}) N_{\nu-1}^{-\mathrm{a}}.$$

If needed, we indicate the dependence of $\widetilde{\mu}_j^\nu$ *on the torus embedding by writing* $\widetilde{\mu}_j^\nu(\omega; \iota)$ *or* $\widetilde{\mu}_j^\nu(\iota)$.

(S3)$_\nu$ *Let* $\breve{\iota}_1, \breve{\iota}_2$ *be torus embeddings satisfying (6.1) with* $\mu_0 \geqq \mathfrak{s}_M + \mu(\mathrm{b})$. *Then, for all* $\omega \in \Omega_\nu^{\gamma_1}(\iota_1) \cap \Omega_\nu^{\gamma_2}(\iota_2)$ *with* $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$ *and* $0 < \gamma < 1/2$, *we have*

$$\| |\mathrm{R}_\nu(\iota_1) - \mathrm{R}_\nu(\iota_2)| \|_{\mathcal{B}(H^{\mathfrak{s}_M})} \lesssim_S N_{\nu-1}^{-\mathrm{a}} \|\iota_1 - \iota_2\|_{\mathfrak{s}_M + \mu(\mathrm{b})}, \tag{7.24}$$

$$\| |\langle\partial_\varphi\rangle^{\mathrm{b}}(\mathrm{R}_\nu(\iota_1) - \mathrm{R}_\nu(\iota_2))| \|_{\mathcal{B}(H^{\mathfrak{s}_M})} \lesssim_S N_{\nu-1} \|\iota_1 - \iota_2\|_{\mathfrak{s}_M + \mu(\mathrm{b})}. \tag{7.25}$$

Moreover, if $\nu \geqq 1$, *then for any* $j \in \mathbb{S}^\perp$, r_j^ν, *given by (7.16), satisfies*

$$\left|(r_j^\nu(\iota_1) - r_j^\nu(\iota_2)) - (r_j^{\nu-1}(\iota_1) - r_j^{\nu-1}(\iota_2))\right| \lesssim \| |\mathrm{R}_\nu(\iota_1) - \mathrm{R}_\nu(\iota_2)| \|_{\mathcal{B}(H^{\mathfrak{s}_M})}, \tag{7.26}$$

$$|r_j^\nu(\iota_1) - r_j^\nu(\iota_2)| \lesssim_S \|\iota_1 - \iota_2\|_{\mathfrak{s}_M + \mu(\mathrm{b})}. \tag{7.27}$$

(S4)$_\nu$ *Let* $\breve{\iota}_1, \breve{\iota}_2$ *be torus embeddings as in* **(S3)**$_\nu$ *and* $0 < \rho < \gamma/2$. *Then*

$$C(S) N_{\nu-1}^\tau \|\iota_1 - \iota_2\|_{\mathfrak{s}_M + \mu(\mathrm{b})} \leqq \rho \quad \Longrightarrow \quad \Omega_\nu^\gamma(\iota_1) \subseteq \Omega_\nu^{\gamma-\rho}(\iota_2). \tag{7.28}$$

Before proving Theorem 7.2 in the subsequent section, we discuss the following application. Theorem 7.2 implies that for any $n \geq 1$, the symplectic invertible operator

$$U_n := \Phi_{n-1} \circ \ldots \circ \Phi_0, \quad \omega \in \bigcap_{\nu=0}^{n} \Omega_\nu^\gamma, \tag{7.29}$$

almost diagonalizes L_0, meaning that (7.32) below holds. Note that since $\Omega_\nu^\gamma \subset \Omega_{\nu-1}^\gamma$ (cf. (7.18)), one has

$$\bigcap_{\nu=0}^{n} \Omega_\nu^\gamma = \Omega_n^\gamma. \tag{7.30}$$

By the same arguments as in [10], one infers the following theorem from Theorem 7.2.

Theorem 7.3. (KAM almost-reducibility) *Assume the ansatz (6.1) with* $\mu_0 \geqq \mathfrak{s}_M + \mu(\mathrm{b})$. *Then for any* $S > \mathfrak{s}_M$ *there exist* $N_0 := N_0(S, \mathrm{b}) > 0$, $0 < \delta_0 := \delta_0(S) < 1$, *so that if*

$$N_0^{\overline{\tau}} \varepsilon \gamma^{-3} \leqq \delta_0, \tag{7.31}$$

with $\overline{\tau} = \overline{\tau}(\tau, \mathbb{S}_+)$ *given by Theorem 7.2, the following holds: For any* $n \in \mathbb{N}$ *and* ω *in* Ω_n^γ*, the operator* U_n*, introduced in (7.29), is well defined and* $\mathtt{L}_n := U_n \mathtt{L}_0 U_n^{-1}$ *satisfies*

$$\mathtt{L}_n = \omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D}_n + \mathtt{R}_n \tag{7.32}$$

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*

$$\mathfrak{M}^\sharp_{\mathtt{R}_n}(s) \lesssim_S N_{n-1}^{-\mathtt{a}}(\varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\mu(\mathtt{b})}), \quad \forall \mathfrak{s}_M \leqq s \leqq S. \tag{7.33}$$

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*

$$\mathfrak{M}^\sharp_{U_n^{\pm1} - \mathrm{Id}_\perp}(s) \lesssim_S \gamma^{-1} N_0^{\tau_1}(\varepsilon + \|\iota\|^{\mathrm{Lip}(\gamma)}_{s+\mu(\mathtt{b})}), \quad \forall \mathfrak{s}_M \leqq s \leqq S. \tag{7.34}$$

Here $\mathrm{Id}_\perp$ *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*

Theorem 7.2 is proved by induction. We first prove the base case $\nu = 0$ and then the induction step.

Base case $\nu = 0$. The items $(\mathbf{S1})_0$, …, $(\mathbf{S4})_0$ are proved separately.

Proof of $(\mathbf{S1})_0$. Properties (7.15)–(7.17) for $\nu = 0$ follow by (7.1)–(7.2) with $r_j^0 = 0$. Furthermore (7.20) holds for $\nu = 0$ in view of the following lemma, which can be proved by the same arguments used in the proof of Lemma 7.6 in [10].

Lemma 7.4. $\mathfrak{M}^\sharp_0(s)$*,* $\mathfrak{M}^\sharp_0(s, \mathtt{b}) \lesssim_{\mathtt{b}} \mathfrak{M}_0(s, \mathtt{b})$ *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 extension Theorem for Lipschitz functions (which is a particular case of the Whitney extension Theorem as recorded in [1, Appendix B], see also Theorem 3, page 174 in [31]) and (6.68) 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 $\nu = 0$ follow by arguing as in the proof of $(\mathbf{S3})_0$ in [10].

Proof of $(\mathbf{S4})_0$. By the definition of Ω_0^γ one has $\Omega_0^\gamma(\iota_1) = \mathtt{DC}(\gamma, \tau) \subseteq \mathtt{DC}(\gamma - \rho, \tau) = \Omega_0^{\gamma-\rho}(\iota_2)$.

Induction step. In this paragraph, 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, for example we write $\mathtt{L}$ for $\mathtt{L}_\nu$, $\mathtt{L}_+$ for $\mathtt{L}_{\nu+1}$, Ψ for Ψ_ν, N for N_ν, $\mathfrak{M}^\sharp(s)$ for $\mathfrak{M}^\sharp_\nu(s)$, etc. We conjugate $\mathtt{L}$ by the symplectic time one flow map

$$\Phi := \exp(\Psi) \tag{7.35}$$

generated by a Hamiltonian vector field Ψ acting in $H^s_\perp$. By a Lie expansion we get

$$
\begin{aligned}
\Phi \mathrm{L} \Phi^{-1} &= \Phi(\omega\cdot\partial_\varphi + \mathrm{i}\mathrm{D})\Phi^{-1} + \Phi \mathrm{R} \Phi^{-1} \\
&= \omega\cdot\partial_\varphi + \mathrm{i}\mathrm{D} - \omega\cdot\partial_\varphi \Psi - \mathrm{i}[\mathrm{D},\Psi] + \Pi_N \mathrm{R} + \Pi_N^\perp \mathrm{R} - \int_0^1 \exp(\tau\Psi)[\mathrm{R},\Psi]\exp(-\tau\Psi)\,\mathrm{d}\tau \\
&\quad + \int_0^1 (1-\tau)\exp(\tau\Psi)\Big[\omega\cdot\partial_\varphi\Psi + \mathrm{i}[\mathrm{D},\Psi],\Psi\Big]\exp(-\tau\Psi)\,\mathrm{d}\tau
\end{aligned}
\tag{7.36}
$$

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

$$
-\omega\cdot\partial_\varphi\Psi - \mathrm{i}[\mathrm{D},\Psi] + \Pi_N\mathrm{R} = [\mathrm{R}] \quad \text{where} \quad [\mathrm{R}] := \mathrm{diag}_{j\in\mathbb{S}^\perp}\mathrm{R}_j^j(0). \tag{7.37}
$$

The solution of (7.37) is

$$
\Psi_j^{j'}(\ell) := \begin{cases} \dfrac{\mathrm{R}_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + \mu_j - \mu_{j'})} & \forall(\ell,j,j')\neq(0,j,j),\ |\ell|\leqq N,\ j,j'\in\mathbb{S}^\perp \\ 0 \quad \text{otherwise.} \end{cases} \tag{7.38}
$$

Note that the denominators in (7.38) do not vanish for $\omega\in\Omega_{\nu+1}^\gamma$ (cf. (7.18)).

Lemma 7.5. (Homological equations) (i) *The solution Ψ of the homological equation (7.37), given by (7.38) for $\omega\in\Omega_{\nu+1}^\gamma$, is a* Lip$(\gamma)$*-modulo-tame operator with a modulo-tame constant satisfying*

$$
\mathfrak{M}^\sharp_\Psi(s) \lesssim N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp(s), \quad \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathtt{b}}\Psi}(s) \lesssim N^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp(s,\mathtt{b}), \tag{7.39}
$$

where $\tau_1 := 2\tau + 1$. Moreover Ψ is Hamiltonian.
(ii) *Let $\breve\iota_1, \breve\iota_2$ be torus embeddings and define $\Delta_{12}\Psi := \Psi(\iota_2) - \Psi(\iota_1)$. If $\gamma/2 \leqq \gamma_1,\gamma_2 \leqq 2\gamma < 1$ then, for any $\omega\in\Omega_{\nu+1}^{\gamma_1}(\iota_1)\cap\Omega_{\nu+1}^{\gamma_2}(\iota_2)$,*

$$
|||\Delta_{12}\Psi|||_{\mathcal{B}(H^{\mathfrak{s}_M})} \leqq C N^{2\tau}\gamma^{-2}\big(|||\mathrm{R}(\iota_2)|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\|\iota_1-\iota_2\|_{\mathfrak{s}_M+\mu(\mathtt{b})} + |||\Delta_{12}\mathrm{R}|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\big), \tag{7.40}
$$

$$
|||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\Psi|||_{\mathcal{B}(H^{\mathfrak{s}_M})} \lesssim_{\mathtt{b}} N^{2\tau}\gamma^{-2}\big(|||\langle\partial_\varphi\rangle^{\mathtt{b}}\mathrm{R}(\iota_2)|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\|\iota_1-\iota_2\|_{\mathfrak{s}_M+\mu(\mathtt{b})} + |||\langle\partial_\varphi\rangle^{\mathtt{b}}\Delta_{12}\mathrm{R}|||_{\mathcal{B}(H^{\mathfrak{s}_M})}\big). \tag{7.41}
$$

Proof. Since R is Hamiltonian, one infers from Definition 2.5 and Lemma 2.6-(iii) that the operator Ψ defined in (7.38) is Hamiltonian as well. We now prove (7.39). Let $\omega\in\Omega_{\nu+1}^\gamma$. By (7.18), and the definition of Ψ in (7.38), it follows that for any $(\ell,j,j')\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp\times\mathbb{S}^\perp$, with $|\ell|\leqq N$, $(\ell,j,j')\neq(0,j,j)$,

$$
|\Psi_j^{j'}(\ell)| \lesssim \langle\ell\rangle^\tau\gamma^{-1}|\mathrm{R}_j^{j'}(\ell)| \tag{7.42}
$$

and

$$
\Delta_\omega\Psi_j^{j'}(\ell) = \frac{\Delta_\omega\mathrm{R}_j^{j'}(\ell)}{\delta_{\ell jj'}(\omega_1)} - \mathrm{R}_j^{j'}(\ell;\omega_2)\frac{\Delta_\omega\delta_{\ell jj'}}{\delta_{\ell jj'}(\omega_1)\delta_{\ell jj'}(\omega_2)},
$$

$$\delta_{\ell jj'}(\omega) := \mathrm{i}(\omega\cdot\ell + \mu_j - \mu_{j'}).$$

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

$$|\Delta_\omega \Psi_j^{j'}(\ell)| \lesssim \langle\ell\rangle^\tau \gamma^{-1} |\Delta_\omega \mathrm{R}_j^{j'}(\ell)| + \langle\ell\rangle^{2\tau+1}\gamma^{-2}|\mathrm{R}_j^{j'}(\ell;\omega_2)||\omega_1-\omega_2|. \tag{7.43}$$

Recalling the definitions (2.33), (7.19), using (7.42), (7.43), and arguing as in the proof of the estimates (7.61) in [10, Lemma 7.7], one then deduces (7.39). The estimates (7.40)–(7.41) can be obtained similarly. □

By (7.36)–(7.37) one has

$$\mathrm{L}_+ = \Phi \mathrm{L} \Phi^{-1} = \omega\cdot\partial_\varphi + \mathrm{i}\mathrm{D}_+ + \mathrm{R}_+$$

which proves (7.22) and (7.15) at the step $\nu+1$, with

$$\begin{aligned} \mathrm{i}\mathrm{D}_+ &:= \mathrm{i}\mathrm{D} + [\mathrm{R}], \\ \mathrm{R}_+ &= \Pi_N^\perp \mathrm{R} - \int_0^1 \exp(\tau\Psi)[\mathrm{R},\Psi]\exp(-\tau\Psi)\,\mathrm{d}\tau \\ &\quad + \int_0^1 (1-\tau)\exp(\tau\Psi)\big[\Pi_N\mathrm{R} - [\mathrm{R}],\Psi\big]\exp(-\tau\Psi)\,\mathrm{d}\tau. \end{aligned} \tag{7.44}$$

The operator L_+ has the same form as L. More precisely, D_+ is diagonal and R_+ is the sum of an operator supported on high frequencies and one which is quadratic in Ψ and R. The new normal form D_+ has the following properties:

Lemma 7.6. (New diagonal part) (i) *The new normal form is*

$$\begin{aligned} &\mathrm{D}_+ = \mathrm{D} - \mathrm{i}[\mathrm{R}], \quad \mathrm{D}_+ := \mathrm{diag}_{j\in\mathbb{S}^\perp}\mu_j^+, \\ &\mu_j^+ := \mu_j + \mathrm{r}_j \in \mathbb{R}, \quad \mathrm{r}_j := -\mathrm{i}\mathrm{R}_j^j(0), \quad \forall j \in \mathbb{S}^\perp, \end{aligned} \tag{7.45}$$

with

$$\mu_{-j}^+ = -\mu_j^+, \quad |\mu_j^+ - \mu_j|^{\mathrm{Lip}(\gamma)} = |\mathrm{r}_j|^{\mathrm{Lip}(\gamma)} \lesssim \mathfrak{M}^\sharp(\mathfrak{s}_M).$$

(ii) *For any tori* $\breve{\iota}_1(\omega)$, $\breve{\iota}_2(\omega)$ *and any* $\omega \in \Omega_\nu^{\gamma_1}(\iota_1)\cap\Omega_\nu^{\gamma_2}(\iota_2)$, *one has*

$$|\mathrm{r}_j(\iota_1) - \mathrm{r}_j(\iota_2)| \lesssim |||\Delta_{12}\mathrm{R}|||_{\mathcal{B}(H^{\mathfrak{s}_M})}. \tag{7.46}$$

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 $|\mu_j^+ - \mu_j|^{\mathrm{Lip}(\gamma)} \leqq |\mathrm{R}_j^j(0)|^{\mathrm{Lip}(\gamma)} \lesssim \mathfrak{M}^\sharp(\mathfrak{s}_M)$. Since $\mathrm{R}(\varphi)$ is Hamiltonian, Lemma 2.6 implies that $\mathrm{r}_j = -\mathrm{i}\mathrm{R}_j^j(0)$, $j \in \mathbb{S}^\perp$, are odd in j and real. The estimate (7.46) is proved in the same way by using $|\Delta_{12}\mathrm{R}_j^j(0)| \leqq C|||\Delta_{12}\mathrm{R}|||_{\mathcal{B}(H^{\mathfrak{s}_M})}$. □

Induction step. Assuming that $(\mathbf{S1})_\nu$–$(\mathbf{S4})_\nu$ are true for a given $\nu \geqq 0$, we show in this paragraph that each of the statements $(\mathbf{S1})_{\nu+1}$, $(\mathbf{S2})_{\nu+1}$, $(\mathbf{S3})_{\nu+1}$, and $(\mathbf{S4})_{\nu+1}$ hold.

Proof of $(\mathbf{S1})_{\nu+1}$. By Lemma 7.5, for any $\omega \in \Omega^\gamma_{\nu+1}$, the solution Ψ_ν of the homological equation (7.37), defined in (7.38), is well defined and that by (7.39), (7.20), it satisfies the estimates (7.23) at $\nu+1$. In particular, the estimate (7.23) for $\nu+1$, $s = \mathfrak{s}_M$ together with (7.6), (7.14) imply that

$$\mathfrak{M}^\sharp_{\Psi_\nu}(\mathfrak{s}_M) \lesssim_{\mathrm{b}} N_\nu^{\tau_1} N_{\nu-1}^{-\mathrm{a}} \gamma^{-1} \mathfrak{M}_0(\mathfrak{s}_M, \mathrm{b}) \leqq 1. \tag{7.47}$$

By Lemma 2.22 and using again Lemma 7.5 one infers that

$$\begin{aligned}
\mathfrak{M}^\sharp_{\Phi_\nu^{\pm 1}}(\mathfrak{s}_M) &\lesssim 1, \\
\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}\Phi_\nu^{\pm 1}}(\mathfrak{s}_M) &\lesssim 1 + \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}\Psi_\nu}(\mathfrak{s}_M) \lesssim 1 + N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(\mathfrak{s}_M, \mathrm{b}), \\
\mathfrak{M}^\sharp_{\Phi_\nu^{\pm 1}}(s) &\lesssim 1 + \mathfrak{M}^\sharp_{\Psi_\nu}(s) \lesssim_s 1 + N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(s), \\
\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}\Phi_\nu^{\pm 1}}(s) &\lesssim 1 + \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}\Psi_\nu}(s) + \mathfrak{M}^\sharp_{\Psi_\nu}(s)\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}\Psi_\nu}(\mathfrak{s}_M), \\
&\overset{(7.14),(7.20),(7.39)}{\lesssim} 1 + N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(s, \mathrm{b}) + N_\nu^{2\tau_1}N_{\nu-1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(s).
\end{aligned} \tag{7.48}$$

By Lemma 7.6, by the estimate (7.20) and Lemma 7.1, the operator $\mathrm{D}_{\nu+1}$ is diagonal and its eigenvalues $\mu_j^{\nu+1} : \Omega^\gamma_{\nu+1} \to \mathbb{R}$ satisfy (7.17) at $\nu+1$.

Next we estimate the remainder $\mathrm{R}_{\nu+1}$ defined in (7.44).

Lemma 7.7. (Nash–Moser iterative scheme) *The operator* $\mathrm{R}_{\nu+1}$ *is* $\mathrm{Lip}(\gamma)$*-modulo-tame with a modulo-tame constant satisfying*

$$\mathfrak{M}^\sharp_{\nu+1}(s) \lesssim N_\nu^{-\mathrm{b}}\mathfrak{M}^\sharp_\nu(s, \mathrm{b}) + N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(s)\mathfrak{M}^\sharp_\nu(\mathfrak{s}_M). \tag{7.49}$$

The operator $\langle\partial_\varphi\rangle^{\mathrm{b}}\mathrm{R}_{\nu+1}$ *is* $\mathrm{Lip}(\gamma)$*-modulo-tame with a modulo-tame constant satisfying*

$$\begin{aligned}
\mathfrak{M}^\sharp_{\nu+1}(s, \mathrm{b}) \lesssim_{\mathrm{b}} &\ \mathfrak{M}^\sharp_\nu(s, \mathrm{b}) + N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(s, \mathrm{b})\mathfrak{M}^\sharp_\nu(\mathfrak{s}_M) \\
&+ N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(\mathfrak{s}_M, \mathrm{b})\mathfrak{M}^\sharp_\nu(s).
\end{aligned} \tag{7.50}$$

Proof. We treat each of the three terms in the definition (7.44) of $\mathrm{R}_{\nu+1}$ separately. By Lemma 2.23 and the definition (7.19) of $\mathfrak{M}^\sharp_\nu(s, \mathrm{b})$, we have

$$\begin{aligned}
\mathfrak{M}^\sharp_{\Pi^\perp_{N_\nu}\mathrm{R}_\nu}(s) &\leqq N_\nu^{-\mathrm{b}}\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}\mathrm{R}_\nu}(s) = N_\nu^{-\mathrm{b}}\mathfrak{M}^\sharp_\nu(s, \mathrm{b}), \\
\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}\Pi^\perp_{N_\nu}\mathrm{R}_\nu}(s) &\leqq \mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}\mathrm{R}_\nu}(s) = \mathfrak{M}^\sharp_\nu(s, \mathrm{b}).
\end{aligned} \tag{7.51}$$

We now estimate the second term $\mathrm{G}_\nu := -\int_0^1 \exp(\tau\Psi_\nu)[\mathrm{R}_\nu, \Psi_\nu]\exp(-\tau\Psi_\nu)\, d\tau$ in (7.44). By applying (7.39), together with the composition Lemma 2.21, one obtains that

$$\begin{aligned}
\mathfrak{M}^\sharp_{[\mathrm{R}_\nu, \Psi_\nu]}(s) &\lesssim N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}^\sharp_\nu(s)\mathfrak{M}^\sharp_\nu(\mathfrak{s}_M), \\
\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}[\mathrm{R}_\nu, \Psi_\nu]}(s) &\lesssim N_\nu^{\tau_1}\gamma^{-1}\big(\mathfrak{M}^\sharp_\nu(s, \mathrm{b})\mathfrak{M}^\sharp_\nu(\mathfrak{s}_M) + \mathfrak{M}^\sharp_\nu(\mathfrak{s}_M, \mathrm{b})\mathfrak{M}^\sharp_\nu(s)\big).
\end{aligned} \tag{7.52}$$

Note that the estimates (7.48) also hold for the maps $\exp(\pm\tau\Psi_\nu)$, uniformly in $\tau \in [-1, 1]$. Therefore, taking into account (7.52) and (7.48), Lemma 2.21, the induction hypothesis for the estimate (7.20), the smallness condition (7.14) (with $\bar\tau$ large enough) and (7.6), $N_\nu^{\tau_1}\mathfrak{M}_\nu^\sharp(\mathfrak{s}_M)\gamma^{-1} \leqq 1$, one concludes that

$$\begin{aligned} &\mathfrak{M}^\sharp_{\mathrm{G}_n}(s) \lesssim N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}_\nu^\sharp(s)\mathfrak{M}_\nu^\sharp(\mathfrak{s}_M), \\ &\mathfrak{M}^\sharp_{\langle\partial_\varphi\rangle^{\mathrm{b}}\mathrm{G}_n}(s) \lesssim N_\nu^{\tau_1}\gamma^{-1}\big(\mathfrak{M}_\nu^\sharp(s,\mathrm{b})\mathfrak{M}_\nu^\sharp(\mathfrak{s}_M) + \mathfrak{M}_\nu^\sharp(\mathfrak{s}_M,\mathrm{b})\mathfrak{M}_\nu^\sharp(s)\big). \end{aligned} \tag{7.53}$$

The estimate of the third term in (7.44) is obtained in a similar way. Together with (7.51) and (7.53) it yields (7.49)–(7.50). □

The estimates (7.49), (7.50), and (7.6), now allow us to prove that (7.20) holds at the step $\nu + 1$.

Lemma 7.8. $\mathfrak{M}^\sharp_{\nu+1}(s) \leqq C_*(\mathfrak{s}_M,\mathrm{b})\mathfrak{M}_0(s,\mathrm{b})N_\nu^{-\mathrm{a}}$ *and* $\mathfrak{M}^\sharp_{\nu+1}(s,\mathrm{b}) \leqq C_*(\mathfrak{s}_M,\mathrm{b})$ $\mathfrak{M}_0(s,\mathrm{b})N_\nu$.

Proof. By (7.49), the induction hypothesis for the estimate (7.20) we get

$$\begin{aligned} \mathfrak{M}^\sharp_{\nu+1}(s) &\leqq C_*(\mathfrak{s}_M,\mathrm{b})N_\nu^{-\mathrm{b}}N_{\nu-1}\mathfrak{M}_0(s,\mathrm{b}) + CC_*(\mathfrak{s}_M,\mathrm{b})^2 N_\nu^{\tau_1}\gamma^{-1}\mathfrak{M}_0(s,\mathrm{b})\mathfrak{M}_0(\mathfrak{s}_M,\mathrm{b})N_{\nu-1}^{-2\mathrm{a}} \\ &\leqq C_*(s_0,\mathrm{b})N_\nu^{-\mathrm{a}}\mathfrak{M}_0(s,\mathrm{b}) \end{aligned}$$

where for the latter inequality we used the definition (7.6) of the constants and the bound (7.14) with $\bar\tau$ and $N_0 := N_0(S,\mathrm{b}) > 0$ large enough. Similarly, by (7.50), (7.20),

$$\begin{aligned} \mathfrak{M}^\sharp_{\nu+1}(s,\mathrm{b}) &\leqq C_*(s_0,\mathrm{b})N_{\nu-1}\mathfrak{M}_0(s,\mathrm{b}) + CC_*(s_0,\mathrm{b})^2 N_\nu^{\tau_1}N_{\nu-1}^{1-\mathrm{a}}\gamma^{-1}\mathfrak{M}_0(s,\mathrm{b})\mathfrak{M}_0(\mathfrak{s}_M,\mathrm{b}) \\ &\leqq C_*(s_0,\mathrm{b})N_\nu\mathfrak{M}_0(s,\mathrm{b}) \end{aligned}$$

where for the latter inequality, we again used (7.6) and (7.14) with $N_0 := N_0(S,\mathrm{b}) > 0$ large enough. □

PROOF OF $(\mathbf{S2})_{\nu+1}$. By Lemma 7.6, for any $j \in \mathbb{S}^\perp$, $\mu_j^{\nu+1} = \mu_j^\nu + \mathrm{r}_j^\nu$ where $|\mathrm{r}_j^\nu|^{\mathrm{Lip}(\gamma)} \lesssim \mathfrak{M}_0(\mathfrak{s}_M,\mathrm{b})N_\nu^{-\mathrm{a}}$. Then $(\mathbf{S2})_{\nu+1}$ follows by defining $\widetilde\mu_j^{\nu+1} := \widetilde\mu_j^\nu + \widetilde{\mathrm{r}}_j^\nu$ where $\widetilde{\mathrm{r}}_j^\nu : \Omega \to \mathbb{R}$ is a Lipschitz extension of r_j^ν (using again the 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]. □

7.2. Almost-Invertibility of $\mathcal{L}_\omega$

By (7.32), for any $\omega \in \Omega_n^\gamma$, we have that $\mathrm{L}_0 = U_n^{-1}\mathrm{L}_n U_n$ where U_n is defined in (7.29) and thus

$$\mathcal{L}_\omega = \mathcal{V}_n^{-1}\mathrm{L}_n\mathcal{V}_n, \quad \mathcal{V}_n := U_n\Phi^{(4)}\cdots\Phi^{(1)}, \tag{7.54}$$

where $\mathcal{L}_\omega$ is the operator introduced in (5.26) and $\Phi^{(1)}, \ldots, \Phi^{(4)}$ are the transformations constructed in Lemmas 6.4, 6.5, 6.6, and respectively, Proposition 6.7.

Lemma 7.9. *There exists $\sigma = \sigma(\tau, \mathbb{S}_+) > 0$ such that, if (7.31) and (6.1) with $\mu_0 \geqq \mathfrak{s}_M + \mu(\mathtt{b}) + \sigma$ hold, then the operators $\mathcal{V}_n^{\pm 1}$ satisfy for any $\mathfrak{s}_M \leqq s \leqq S$ the estimate*

$$\|\mathcal{V}_n^{\pm 1} h\|_s^{\mathrm{Lip}(\gamma)} \lesssim_S \|h\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + N_0^{\tau_1} \gamma^{-1} \|\iota\|_{s+\mu(\mathtt{b})+\sigma}^{\mathrm{Lip}(\gamma)} \|h\|_{\mathfrak{s}_M+\sigma}^{\mathrm{Lip}(\gamma)}. \tag{7.55}$$

Proof. The claimed estimates follow from the estimates (6.30), (6.46), (6.54), (6.72), and (7.34) together with the Lemmata 2.16, 2.17, 2.20. □

We now decompose the operator $\mathtt{L}_n = \omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D}_n + \mathtt{R}_n$ in (7.32) as

$$\mathtt{L}_n = \mathfrak{L}_n^< + \mathtt{R}_n + \mathtt{R}_n^\perp \tag{7.56}$$

with

$$\mathfrak{L}_n^< := \Pi_{K_n}\big(\omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D}_n\big)\Pi_{K_n} + \Pi_{K_n}^\perp, \quad \mathtt{R}_n^\perp := \Pi_{K_n}^\perp\big(\omega \cdot \partial_\varphi + \mathrm{i}\mathtt{D}_n\big)\Pi_{K_n}^\perp - \Pi_{K_n}^\perp, \tag{7.57}$$

where the diagonal operator $\mathtt{D}_n$ is defined in (7.15) (with $\nu = n$), $K_n = K_0^{\chi^n}$ is the scale of the nonlinear Nash–Moser iterative scheme introduced in (5.28), and $\Pi_{K_n}^\perp = \mathrm{Id}_\perp - \Pi_{K_n}$ with Π_{K_n} denoting the projector in (2.2). The diagonal constant coefficient operator $\mathfrak{L}_n^<$ can be inverted assuming the following standard non-resonance conditions:

Lemma 7.10. (First order Melnikov non-resonance conditions) *Let $n \geq 0$. Then for any ω in*

$$\Lambda_{n+1}^\gamma := \Lambda_{n+1}^\gamma(\iota) := \big\{\omega \in \Omega : |\omega \cdot \ell + \widetilde{\mu}_j^n| \geqq 2\gamma |j|^3 \langle \ell \rangle^{-\tau}, \quad \forall |\ell| \leqq K_n,\ j \in \mathbb{S}^\perp\big\}, \tag{7.58}$$

the operator $\mathfrak{L}_n^<$ in (7.57) is invertible and

$$\|(\mathfrak{L}_n^<)^{-1} g\|_s^{\mathrm{Lip}(\gamma)} \lesssim \gamma^{-1} \|g\|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)}. \tag{7.59}$$

By (7.54), (7.56), Theorem 7.3, estimates (7.59), (7.60), (7.55), and using that, for all $b > 0$,

$$\|\mathtt{R}_n^\perp h\|_{\mathfrak{s}_M}^{\mathrm{Lip}(\gamma)} \lesssim K_n^{-b} \|h\|_{\mathfrak{s}_M+b+3}^{\mathrm{Lip}(\gamma)}, \quad \|\mathtt{R}_n^\perp h\|_s^{\mathrm{Lip}(\gamma)} \lesssim \|h\|_{s+3}^{\mathrm{Lip}(\gamma)}, \tag{7.60}$$

we deduce the following theorem, stating the assumption **A-I** on the almost-invertibility of $\mathcal{L}_\omega$ in Section 5:

Theorem 7.11. (Almost-invertibility of $\mathcal{L}_\omega$) *Assume the ansatz (6.1) with $\mu_0 \geqq \mathfrak{s}_M + \mu(\mathtt{b})$. Let $\mathtt{a}, \mathtt{b}, M$ as in (7.6), and $S > \mathfrak{s}_M$. There exists $\sigma = \sigma(\tau, \mathbb{S}_+) > 0$ so that, if (7.31) and (6.1) hold with $\mu_0 \geqq \mathfrak{s}_M + \mu(\mathtt{b}) + \sigma$, then, for any $n \geq 0$ and any*

$$\omega \in \mathbf{\Omega}_{n+1}^\gamma = \mathbf{\Omega}_{n+1}^\gamma(\iota) := \Omega_{n+1}^\gamma(\iota) \cap \Lambda_{n+1}^\gamma(\iota) \tag{7.61}$$

(see (7.18), (7.58)), the operator $\mathcal{L}_\omega$, defined in (5.26), can be decomposed as

$$\mathcal{L}_\omega = \mathcal{L}_\omega^< + \mathcal{R}_\omega + \mathcal{R}_\omega^\perp, \quad \mathcal{L}_\omega^< := \mathcal{V}_n^{-1} \mathfrak{L}_n^< \mathcal{V}_n, \quad \mathcal{R}_\omega := \mathcal{V}_n^{-1} \mathtt{R}_n \mathcal{V}_n, \quad \mathcal{R}_\omega^\perp := \mathcal{V}_n^{-1} \mathtt{R}_n^\perp \mathcal{V}_n, \tag{7.62}$$

where $\mathcal{L}_\omega^<$ is invertible and satisfies (5.32) and the operators $\mathcal{R}_\omega$ and $\mathcal{R}_\omega^\perp$ satisfy (5.30)–(5.31).

8. Proof of Theorem 1.1

The goal of this section is to prove Theorem 1.1. In Section 8.3 we deduce Theorem 1.1 from Theorem 4.1. The proof of the latter theorem is also given in Section 8.3, relying on the Nash–Moser Theorem 8.1 and its Corollary 8.4 in Section 8.1, and the measure estimate of Proposition 8.7 in Section 8.2.

8.1. The Nash–Moser Iteration

The main result of this section is Theorem 8.1 which implements a Nash–Moser iteration scheme by providing a sequence of better and better approximate solutions of the equation $\mathcal{F}_\omega(\iota,\zeta) = 0$ (cf. (4.6)) under the assumption that ω satisfies nonresonance conditions, that is ω belongs to the sets $\mathcal{G}_n$ defined in (8.10). A key ingredient into its proof is Theorem 5.7, concerning the existence of an approximate right inverse of $d_{\iota,\zeta}\mathcal{F}_\omega(\iota,\zeta)$.

To describe the iteration scheme, first recall that $L^2_\varphi = L^2_\varphi(\mathbb{T}^{\mathbb{S}_+}, \mathbb{R}^{\mathbb{S}_+})$ (cf. (4.9)) and $L^2_\perp = L^2(\mathbb{T}^{\mathbb{S}_+}, L^2_\perp(\mathbb{T}_1))$ (cf. (1.24)). We then introduce finite-dimensional subspaces of $L^2_\varphi \times L^2_\varphi \times L^2_\perp$, defined for any $n \in \mathbb{N}$ as

$$\mathrm{E}_n := \big\{\iota(\varphi) = (\Theta, y, w)(\varphi), \ \ \Theta = \Pi_n\Theta, \ y = \Pi_n y, \ w = \Pi_n w\big\}$$

where, by a slight abuse of notation, $\Pi_n : L^2_\perp \to \cap_{s\geq 0} H^s_\perp$ denotes the projector Π_{K_n}, introduced in (2.2),

$$\Pi_n : w = \sum_{\ell\in\mathbb{Z}^{\mathbb{S}_+}, j\in\mathbb{S}^\perp} w_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi+2\pi j x)} \quad\mapsto\quad \Pi_n w := \sum_{|(\ell,j)|\leqq K_n} w_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi+2\pi j x)}, \tag{8.1}$$

with $K_n = K_0^{\chi^n}$, $n \geq 1$, (cf. (5.28)) and $K_0 > 1$ a constant chosen in Theorem 8.1 below. We also denote by Π_n the L^2- orthogonal projector on L^2_φ, $L^2_\varphi \to \cap_{s\geq 0} H^s_\varphi$, $v = \sum_{\ell\in\mathbb{Z}^{\mathbb{S}_+}} v_\ell e^{\mathrm{i}\ell\cdot\varphi} \mapsto \Pi_n(v) = \sum_{|\ell|\leqq K_n} v_\ell e^{\mathrm{i}\ell\cdot\varphi}$. The projectors Π_n, $n \geq 0$, are smoothing operators on the Sobolev spaces $H^s_\perp$ (and H^s_φ), meaning that 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

$$\overline{\sigma} := \max\{\sigma_1, \sigma_2\}, \quad \mathtt{b} := [\mathtt{a}] + 2, \quad \mathtt{a} = 3\tau_1 + 1, \quad \tau_1 = 2\tau + 1, \quad \chi = 3/2, \tag{8.2}$$

$$\mathtt{a}_1 := \max\{12\overline{\sigma} + 13, \ p\tau + 3 + \chi(\mu(\mathtt{b}) + 2\overline{\sigma})\}, \quad \mathtt{a}_2 := \chi^{-1}\mathtt{a}_1 - \mu(\mathtt{b}) - 2\overline{\sigma}, \tag{8.3}$$

$$\mathtt{b}_1 := \mathtt{a}_1 + \mu(\mathtt{b}) + 3\overline{\sigma} + 4 + \frac{2}{3}\mu_1, \quad \mu_1 := 3(\mu(\mathtt{b}) + 2\overline{\sigma} + 2) + 1, \quad S := \mathfrak{s}_M + \mathtt{b}_1, \tag{8.4}$$

where σ_1 is defined in Lemma 4.3, σ_2 in Theorem 5.7, $\mathtt{a}$, $\mu(\mathtt{b})$ in (7.6) , and $\mathfrak{s}_M$ in (2.55). The number p is the exponent in (5.27) and is requested to satisfy

$$p\mathtt{a} > (\chi - 1) \cdot \mathtt{a}_1 + \chi \cdot (\overline{\sigma} + 4) \overset{\chi=3/2}{=} \frac{1}{2}\mathtt{a}_1 + \frac{3}{2}(\overline{\sigma} + 4). \tag{8.5}$$

In view of the definition (8.3) of $\mathtt{a}_1$, we can define $p := p(\tau, \mathbb{S}_+)$ as

$$p := \frac{12\overline{\sigma} + 17 + \chi \cdot (\mu(\mathtt{b}) + 2\overline{\sigma})}{\mathtt{a}}. \tag{8.6}$$

We denote by $\|W\|_s^{\mathrm{Lip}(\gamma)} := \max\{\|\iota\|_s^{\mathrm{Lip}(\gamma)}, |\zeta|^{\mathrm{Lip}(\gamma)}\}$ the norm of a map

$$W := (\iota, \zeta) : \Omega \to \left(H^s_\varphi \times H^s_\varphi \times H^s_\perp\right) \times \mathbb{R}^{\mathbb{S}_+}, \quad \omega \mapsto W(\omega) = (\iota(\omega), \zeta(\omega)).$$

Theorem 8.1. (Nash–Moser) *There exist* $0 < \delta_0 < 1$ *(small) and* $C_* > 0$ *(large) so that if*

$$\tau_2 := \max\{p\overline{\tau} + 3,\ 4\overline{\sigma} + 4 + \mathtt{a}_1\}, \quad \gamma := \varepsilon^{\mathfrak{a}}, \quad 0 < \mathfrak{a} < \frac{1}{\tau_2},$$
$$K_0 := \gamma^{-1}, \quad \varepsilon K_0^{\tau_2} = \varepsilon^{1-\mathfrak{a}\tau_2} < \delta_0, \tag{8.7}$$

where $\overline{\tau} := \overline{\tau}(\tau, \mathbb{S}_+)$ *is defined as in Theorem 7.2, then the following holds: for any* $n \in \mathbb{N}$

$(\mathcal{P}1)_n$ (Estimates in low norms) *Let* $\tilde{W}_0 := (0, 0)$. *If* $n \geq 1$, *then there exists a* $\mathrm{Lip}(\gamma)$*-function*

$$\tilde{W}_n : \mathbb{R}^{\mathbb{S}_+} \to \mathrm{E}_{n-1} \times \mathbb{R}^{\mathbb{S}_+},\ \omega \mapsto \tilde{W}_n(\omega) := (\tilde{\iota}_n, \tilde{\zeta}_n),$$

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* $U_0 := (\varphi, 0, 0, 0)$. *The difference* $\tilde{H}_n := \tilde{U}_n - \tilde{U}_{n-1}$ *satisfies*

$$\|\tilde{H}_1\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M + \mu(\mathtt{b}) + \overline{\sigma}} \lesssim \varepsilon\gamma^{-2} \quad (n = 1), \quad \|\tilde{H}_n\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M + \mu(\mathtt{b}) + \overline{\sigma}} \lesssim \varepsilon\gamma^{-2} K_{n-1}^{-\mathtt{a}_2} \quad (n \geq 2)\,, \tag{8.9}$$

where $K_n = K_0^{\chi^n}$ *(cf. (5.28)).*
$(\mathcal{P}2)_n$ *Define*

$$\mathcal{G}_0 := \Omega, \quad \mathcal{G}_n := \mathcal{G}_{n-1} \cap\ \mathbf{\Omega}_n^\gamma(\tilde{\iota}_{n-1}) \quad \forall n \geq 1, \tag{8.10}$$

where $\mathbf{\Omega}_n^\gamma(\tilde{\iota}_{n-1})$ *is defined in (7.61). Then for any* $\omega \in \mathcal{G}_n$

$$\|\mathcal{F}_\omega(\tilde{U}_n)\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M} \leqq C_* \varepsilon K_{n-1}^{-\mathtt{a}_1}, \quad K_{-1} := 1. \tag{8.11}$$

$(\mathcal{P}3)_n$ (Estimates in high norms) $\|\tilde{W}_n\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M + \mathtt{b}_1} \leqq C_* \varepsilon K_{n-1}^{\mu_1}, \quad \forall \omega \in \mathcal{G}_n.$

Proof. The theorem can be proved in a by now standard way (cf. [1,10]). We argue by induction. To simplify notation, we write within this proof $\|\cdot\|$ instead of $\|\cdot\|^{\mathrm{Lip}(\gamma)}$.

Step 1: Proof of $(\mathcal{P}1)_0$, $(\mathcal{P}2)_0$, $(\mathcal{P}3)_0$. Note that $(\mathcal{P}1)_0$ and $(\mathcal{P}3)_0$ are trivially satisfied and hence it remains to verify (8.11) for $n = 0$. By (4.6), (4.13), (4.3), and Lemma 4.3, there exists $C_* > 0$ large enough so that $\|\mathcal{F}_\omega(U_0)\|^{\mathrm{Lip}(\gamma)}_{\mathfrak{s}_M} \le \varepsilon C_*$.

Step 2: Proof of induction step. Assuming that $(\mathcal{P}1)_n$, $(\mathcal{P}2)_n$, $(\mathcal{P}3)_n$ hold for some $n \geqq 0$, it is to prove that $(\mathcal{P}1)_{n+1}$, $(\mathcal{P}2)_{n+1}$, $(\mathcal{P}3)_{n+1}$ hold. We are going to define the approximation $\tilde{U}_{n+1}$ by a modified Nash–Moser scheme. To this end, we prove the almost-approximate invertibility of the linearized operator

$$L_n = L_n(\omega) := \mathrm{d}_{\iota,\zeta}\mathcal{F}_\omega(\tilde{\iota}_n(\omega)) \tag{8.12}$$

by applying Theorem 5.7 to $L_n(\omega)$. To prove that the assumptions (5.29)–(5.32) in Theorem 5.7 hold, we apply Theorem 7.11 with $\iota = \tilde{\iota}_n$. By choosing ε small enough it follows from (8.7) that $N_0 = K_0^p = \gamma^{-p} = \varepsilon^{-p\mathfrak{a}}$ and the smallness condition (7.31) required in Theorem 7.11 holds. In addition, (6.1) holds by (8.9). Therefore Theorem 7.11 applies, and we deduce that (5.29)–(5.32) hold for all ω in the set $\mathbf{\Omega}^\gamma_{n+1}(\tilde{\iota}_n)$, defined in (7.61). Now we apply Theorem 5.7 to the linearized operator $L_n(\omega)$ with $\Omega_o = \mathbf{\Omega}^\gamma_{n+1}(\tilde{\iota}_n)$ and $S = \mathfrak{s}_M + \mathrm{b}_1$ (cf. (8.4)). It implies that there exists an almost-approximate inverse $\mathbf{T}_n := \mathbf{T}_n(\omega, \tilde{\iota}_n(\omega))$ satisfying

$$\|\mathbf{T}_n g\|_s \lesssim_{\mathfrak{s}_M+\mathrm{b}_1} \gamma^{-2}\big(\|g\|_{s+\overline{\sigma}} + K_0^{\tau_1 p}\gamma^{-1}\|\tilde{\iota}_n\|_{s+\mu(\mathrm{b})+\overline{\sigma}}\|g\|_{\mathfrak{s}_M+\overline{\sigma}}\big), \quad \forall \mathfrak{s}_M \leqq s \leqq \mathfrak{s}_M + \mathrm{b}_1, \tag{8.13}$$

where we used that $\overline{\sigma} \geqq \sigma_2$ (cf. (8.2)), σ_2 is the loss of regularity constant appearing in the estimate (5.45), and $N_0 = K_0^p$. Furthermore, by (8.7)–(8.8), one obtains

$$K_0^{\tau_1 p}\gamma^{-1}\|\tilde{W}_n\|_{\mathfrak{s}_M+\mu(\mathrm{b})+\overline{\sigma}} \leqq 1, \tag{8.14}$$

and hence, for the special value $s = \mathfrak{s}_M$, (8.13) becomes

$$\|\mathbf{T}_n g\|_{\mathfrak{s}_M} \lesssim_{\mathrm{b}_1} \gamma^{-2}\|g\|_{\mathfrak{s}_M+\overline{\sigma}}. \tag{8.15}$$

For all $\omega \in \mathcal{G}_{n+1} = \mathcal{G}_n \cap \mathbf{\Lambda}^\gamma_{n+1}(\tilde{\iota}_n)$ (cf. (8.10)), we define

$$U_{n+1} := \tilde{U}_n + H_{n+1}, \quad H_{n+1} := (\widehat{\iota}_{n+1}, \widehat{\zeta}_{n+1}) := -\mathbf{\Pi}_n\mathbf{T}_n\Pi_n\mathcal{F}_\omega(\tilde{U}_n) \in \mathrm{E}_n \times \mathbb{R}^{\mathbb{S}_+}, \tag{8.16}$$

where $\mathbf{\Pi}_n$ is defined by (cf. (8.1))

$$\mathbf{\Pi}_n(\iota,\zeta) := (\Pi_n\iota, \zeta), \quad \mathbf{\Pi}_n^\perp(\iota,\zeta) := (\Pi_n^\perp\iota, 0), \quad \forall(\iota,\zeta). \tag{8.17}$$

To 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 $L_n := \mathrm{d}_{\iota,\zeta}\mathcal{F}_\omega(\tilde{U}_n)$ and Q_n is defined by (8.18). Then, by the definition of H_{n+1} in (8.16), and by taking into account (8.17), one has

$$\begin{aligned}\mathcal{F}_\omega(U_{n+1}) &= \mathcal{F}_\omega(\tilde{U}_n) - L_n \mathbf{\Pi}_n \mathbf{T}_n \Pi_n \mathcal{F}_\omega(\tilde{U}_n) + Q_n \\ &= \mathcal{F}_\omega(\tilde{U}_n) - L_n \mathbf{T}_n \Pi_n \mathcal{F}_\omega(\tilde{U}_n) + L_n \mathbf{\Pi}_n^{\perp} \mathbf{T}_n \Pi_n \mathcal{F}_\omega(\tilde{U}_n) + Q_n \\ &= \Pi_n^{\perp} \mathcal{F}_\omega(\tilde{U}_n) + R_n + Q_n + P_n \end{aligned} \tag{8.19}$$

where

$$R_n := L_n \mathbf{\Pi}_n^{\perp} \mathbf{T}_n \Pi_n \mathcal{F}_\omega(\tilde{U}_n), \quad P_n := -(L_n \mathbf{T}_n - \mathrm{Id}) \Pi_n \mathcal{F}_\omega(\tilde{U}_n). \tag{8.20}$$

We first note that for any $\omega \in \Omega$ and any $s \geqq \mathfrak{s}_M$ one has, by (4.6), Lemma 4.3, and (8.2) and (8.8) and using the triangle inequality,

$$\|\mathcal{F}_\omega(\tilde{U}_n)\|_s \lesssim_s \|\mathcal{F}_\omega(U_0)\|_s + \|\mathcal{F}_\omega(\tilde{U}_n) - \mathcal{F}_\omega(U_0)\|_s \lesssim_s \varepsilon + \|\tilde{W}_n\|_{s+\overline{\sigma}}, \tag{8.21}$$

and, by (8.8), (8.7), (8.11),

$$K_0^{\tau_1 p} \gamma^{-1} \|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M} \leqq 1. \tag{8.22}$$

To conclude, we first need to prove the following inductive estimates of Nash–Moser type:

Lemma 8.2. *Let* $\mu_2 := \mu(\mathtt{b}) + 3\overline{\sigma} + 3$. *Then for any* $\omega \in \mathcal{G}_{n+1}$,

$$\begin{aligned} \|\mathcal{F}_\omega(U_{n+1})\|_{\mathfrak{s}_M} &\lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{\mu_2-\mathtt{b}_1}(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}) + K_n^{4\overline{\sigma}+4} \|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}^2 \\ &\quad + \varepsilon K_{n-1}^{-p\mathtt{a}} K_n^{\overline{\sigma}+4} \|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}, \end{aligned} \tag{8.23}$$

$$\begin{aligned} \|W_1\|_{\mathfrak{s}_M+\mathtt{b}_1} &\lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_0^2 \varepsilon, \\ |W_{n+1}\|_{\mathfrak{s}_M+\mathtt{b}_1} &\lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{\mu(\mathtt{b})+2\overline{\sigma}+2}(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}), \quad n \geqq 1. \end{aligned} \tag{8.24}$$

Proof of Lemma 8.2. 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

$$\begin{aligned} \|H_{n+1}\|_{\mathfrak{s}_M+\mathtt{b}_1} &\lesssim_{\mathfrak{s}_M+\mathtt{b}_1} \gamma^{-2}\big(K_n^{\overline{\sigma}} \|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M+\mathtt{b}_1} + K_n^{\mu(\mathtt{b})+2\overline{\sigma}} K_0^{\tau_1 p} \gamma^{-1} \|\tilde{\imath}_n\|_{\mathfrak{s}_M+\mathtt{b}_1} \|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}\big) \\ &\overset{(8.21),(8.22)}{\lesssim_{\mathfrak{s}_M+\mathtt{b}_1}} K_n^{\mu(\mathtt{b})+2\overline{\sigma}} \gamma^{-2}\big(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}\big) \\ &\overset{\gamma^{-1}=K_0 \leqq K_n}{\lesssim_{\mathfrak{s}_M+\mathtt{b}_1}} K_n^{\mu(\mathtt{b})+2\overline{\sigma}+2}\big(\varepsilon + \|\tilde{W}_n\|_{\mathfrak{s}_M+\mathtt{b}_1}\big), \end{aligned} \tag{8.25}$$

$$|H_{n+1}\|_{\mathfrak{s}_M} \overset{(8.15)}{\lesssim_{\mathfrak{s}_M+\mathtt{b}_1}} \gamma^{-2} K_n^{\overline{\sigma}} \|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}. \tag{8.26}$$

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), (8.26), (8.11), and since $\chi 2\overline{\sigma} - \mathtt{a}_1 \leqq 0$ (see (8.3)), we infer that $\|\tilde{W}_n + t H_{n+1}\|_{\mathfrak{s}_M+\overline{\sigma}} \lesssim \varepsilon \gamma^{-2} K_0^{2\overline{\sigma}}$ for all $t \in [0, 1]$. Since $\gamma^{-1} = K_0$, by (8.7) we can apply Lemma 4.3 and by Taylor's formula, using (8.18), (4.6), (8.26), (2.3), and $\gamma^{-1} = K_0 \leqq K_n$, we get

$$\|Q_n\|_{\mathfrak{s}_M} \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} \|H_{n+1}\|_{\mathfrak{s}_M+\overline{\sigma}}^2 \lesssim_{\mathfrak{s}_M+\mathtt{b}_1} K_n^{4\overline{\sigma}+4} \|\mathcal{F}_\omega(\tilde{U}_n)\|_{\mathfrak{s}_M}^2. \tag{8.27}$$

Estimate of P_n. By (5.46), $L_n \mathbf{T}_n - \mathrm{Id} = \mathcal{P}(\tilde{\imath}_n) + \mathcal{P}_\omega(\tilde{\imath}_n) + \mathcal{P}_\omega^{\perp}(\tilde{\imath}_n)$. Accordingly, we decompose P_n in (8.20) as $P_n = -P_n^{(1)} - P_{n,\omega} - P_{n,\omega}^{\perp}$, where

$$P_n^{(1)} := \Pi_n \mathcal{P}(\tilde\iota_n)\Pi_n \mathcal{F}_\omega(\tilde U_n), \quad P_{n,\omega} := \Pi_n \mathcal{P}_\omega(\tilde\iota_n)\Pi_n \mathcal{F}_\omega(\tilde U_n),$$
$$P_{n,\omega}^\perp := \Pi_n \mathcal{P}_\omega^\perp(\tilde\iota_n)\Pi_n \mathcal{F}_\omega(\tilde U_n).$$

By (2.3),

$$\begin{aligned}\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M+\overline{\sigma}} &\leqq \|\Pi_n\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M+\overline{\sigma}} + \|\Pi_n^\perp\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M+\overline{\sigma}} \\ &\leqq K_n^{\overline{\sigma}}(\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M} + K_n^{-\mathrm{b}_1}\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M+\mathrm{b}_1}).\end{aligned} \tag{8.28}$$

By (5.47), (8.14), (8.28), and using (8.21), (8.22), and $\gamma^{-1} = K_0 \leqq K_n$ one obtains

$$\begin{aligned}\|P_n^{(1)}\|_{\mathfrak{s}_M} &\lesssim_{\mathfrak{s}_M+\mathrm{b}_1} \gamma^{-3}K_n^{2\overline{\sigma}}\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M}(\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M} + K_n^{-\mathrm{b}_1}\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M+\mathrm{b}_1}) \\ &\lesssim_{\mathfrak{s}_M+\mathrm{b}_1} K_n^{2\overline{\sigma}+3}\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M}(\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M} + K_n^{\overline{\sigma}-\mathrm{b}_1}(\varepsilon + \|\tilde W_n\|_{\mathfrak{s}_M+\mathrm{b}_1})) \\ &\lesssim_{s_0+\mathrm{b}_1} K_n^{2\overline{\sigma}+3}\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M}^2 + K_n^{3\overline{\sigma}+3-\mathrm{b}_1}(\varepsilon + \|\tilde W_n\|_{\mathfrak{s}_M+\mathrm{b}_1}).\end{aligned} \tag{8.29}$$

By (5.48), (8.14), (8.8), (2.3), we have

$$\|P_{n,\omega}\|_{\mathfrak{s}_M} \lesssim_{\mathfrak{s}_M+\mathrm{b}_1} \varepsilon\gamma^{-4}N_{n-1}^{-\mathtt{a}}K_n^{\overline{\sigma}}\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M} \overset{\gamma^{-1}=K_0\leqq K_n}{\lesssim_{s_0+\mathrm{b}_1}} \varepsilon N_{n-1}^{-\mathtt{a}}K_n^{\overline{\sigma}+4}\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M}, \tag{8.30}$$

with $\mathtt{a}$ given as in (8.2). By (5.49), (2.3), (8.4), (8.11), (8.22) and by using (8.21), $\gamma^{-1} = K_0 \leqq K_n$, we get

$$\begin{aligned}\|P_{n,\omega}^\perp\|_{\mathfrak{s}_M} &\lesssim_{\mathfrak{s}_M+\mathrm{b}_1} K_n^{\mu(\mathrm{b})+2\overline{\sigma}-\mathrm{b}_1}\gamma^{-2}(\|\mathcal{F}_\omega(\tilde U_n)\|_{\mathfrak{s}_M+\mathrm{b}_1} + \varepsilon\|\tilde W_n\|_{\mathfrak{s}_M+\mathrm{b}_1}) \\ &\lesssim_{\mathfrak{s}_M+\mathrm{b}_1} K_n^{\mu(\mathrm{b})+3\overline{\sigma}+2-\mathrm{b}_1}(\varepsilon + \|\tilde W_n\|_{\mathfrak{s}_M+\mathrm{b}_1}).\end{aligned} \tag{8.31}$$

ESTIMATE OF R_n. Since $L_n = d_{\iota,\zeta}\mathcal{F}_\omega(\tilde\iota_n(\omega))$ (cf. (8.12)), $d_{\iota,\zeta}\mathcal{F}_\omega(\iota,\zeta)[\widehat\iota,\widehat\zeta] = \omega\cdot\partial_\varphi\widehat\iota - d_\iota X_{\mathcal{H}_\varepsilon}(\breve\iota)[\widehat\iota] - (0,\widehat\zeta,0)$ (cf. (5.1)), and $\mathcal{H}_\varepsilon = \mathcal{N} + \mathcal{P}_\varepsilon$, one has that for any $\widehat U = (\widehat\iota,\widehat\zeta)$,

$$\begin{aligned}L_n\widehat U &= \omega\cdot\partial_\varphi\widehat\iota - d_\iota X_{\mathcal{H}_\varepsilon}\big((\varphi,0,0)+\tilde\iota_n\big)[\widehat\iota] - (0,\widehat\zeta,0) \\ &\overset{(4.13)}{=} \omega\cdot\partial_\varphi\widehat\iota - d_\iota X_{\mathcal{N}}\big((\varphi,0,0)+\tilde\iota_n\big)[\widehat\iota] - d_\iota X_{\mathcal{P}_\varepsilon}\big((\varphi,0,0)+\tilde\iota_n\big)[\widehat\iota] - (0,\widehat\zeta,0)\end{aligned} \tag{8.32}$$

where by (4.13), $d_\iota X_{\mathcal{N}}\big((\varphi,0,0)+\tilde\iota_n\big)[\widehat\iota] = \big(-\Omega_{\mathbb{S}_+}^{kdv}(\nu)[\widehat y],\ 0,\ \partial_x\Omega^{kdv}(D;\nu)[\widehat w]\big)$. By the estimate of $d_\iota X_{\mathcal{P}_\varepsilon}$ of Lemma 4.3, 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), $\gamma^{-1} = K_0 \leqq K_n$, we get

$$\|R_n\|_{\mathfrak{s}_M} \lesssim_{\mathfrak{s}_M+\mathrm{b}_1} K_n^{\mu(\mathrm{b})+3\overline{\sigma}+2-\mathrm{b}_1}(\varepsilon + \|\tilde W_n\|_{\mathfrak{s}_M+\mathrm{b}_1}). \tag{8.33}$$

ESTIMATE OF $\mathcal{F}_\omega(U_{n+1})$. By (8.19), (2.3), (8.21), (8.27), (8.29)–(8.31), (8.33), (8.8), we get (8.23). ESTIMATE OF $W_1 = H_1$. By (8.16) and (8.13) one has

$$\|W_1\|_{\mathfrak{s}_M+\mathrm{b}_1} = \|H_1\|_{\mathfrak{s}_M+\mathrm{b}_1} \lesssim_{\mathfrak{s}_M+\mathrm{b}_1} \gamma^{-2}\|\mathcal{F}_\omega(U_0)\|_{\mathfrak{s}_M+\mathrm{b}_1+\overline{\sigma}} \lesssim_{\mathfrak{s}_M+\mathrm{b}_1} \varepsilon\gamma^{-2} \overset{\gamma^{-1}=K_0}{\lesssim} K_0^2\varepsilon$$

implying the first estimate in (8.24).
ESTIMATE OF $W_{n+1} = \tilde W_n + H_{n+1}$, $n \geqq 1$. The claimed estimates for W_{n+1} in (8.24) follows by (8.25). □

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 $\omega \in \mathcal{G}_{n+1}$, $n \geq 0$,*

$$\|\mathcal{F}_\omega(U_{n+1})\|_{\mathfrak{s}_M}^{\mathrm{Lip}(\gamma)} \leqq C_* \varepsilon K_n^{-\mathtt{a}_1}, \qquad \|W_{n+1}\|_{\mathfrak{s}_M+\mathtt{b}_1}^{\mathrm{Lip}(\gamma)} \leqq C_* K_n^{\mu_1} \varepsilon, \tag{8.34}$$

$$\|H_1\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}, \qquad \|H_{n+1}\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2} K_n^{\mu(\mathtt{b})+2\overline{\sigma}} K_{n-1}^{-\mathtt{a}_1}, \ (n \geqq 1). \tag{8.35}$$

Proof of Lemma 8.3. First note that, by (8.10), $\mathcal{G}_{n+1} \subset \mathcal{G}_n$ and so for any $\omega \in \mathcal{G}_{n+1}$ (8.11) and the inequality in $(\mathcal{P}3)_n$ holds. Then the first inequality in (8.34) follows by (8.23), $(\mathcal{P}2)_n$, $(\mathcal{P}3)_n$, $\gamma^{-1} = K_0 \leqq K_n$, and by (8.3)–(8.6). For $n = 0$ we use also (8.7). Concerning the second inequality in (8.34), note that for $n = 0$, the inequality follows directly from the bound for W_1 in (8.24) since $\mu_1 \geqq 2$ (cf. (8.4)) and $C_* > 0$ is chosen large enough; the second inequality in (8.34) for $n \geqq 1$ 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 $K_0 = \varepsilon^{-\mathfrak{a}}$ large enough (that is, ε small enough). Since $H_1 = W_1$, the first inequality in (8.35) 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 $n \geqq 1$, estimate (8.35) follows by (2.3), (8.26) and (8.11). □

We are now in a position to finish the proof of Theorem 8.1. Denote by $\tilde{H}_{n+1}$ the $\mathrm{Lip}(\gamma)$-extension of $(H_{n+1})_{|\mathcal{G}_{n+1}}$ to the whole set Ω of parameters, provided by the Kirszbraun theorem. Then $\tilde{H}_{n+1}$ satisfies the same bound as H_{n+1} in (8.35) and therefore, by the definition of $\mathtt{a}_2$ in (8.3), the estimate (8.9) holds at $n+1$.

Finally we define $\tilde{W}_{n+1} := \tilde{W}_n + \tilde{H}_{n+1}$ and $\tilde{U}_{n+1} := \tilde{U}_n + \tilde{H}_{n+1}$, which both are defined for all $\omega \in \Omega$. Note that

$$\tilde{U}_{n+1} = U_0 + \tilde{W}_n + \tilde{H}_{n+1} = U_0 + \tilde{W}_{n+1}$$

and that for any $\omega \in \mathcal{G}_{n+1}$, $\tilde{W}_{n+1} = W_{n+1}$, $\tilde{U}_{n+1} = U_{n+1}$. Hence $(\mathcal{P}2)_{n+1}$, $(\mathcal{P}3)_{n+1}$ follow from Lemma 8.3. Moreover by (8.9), which at this point has been proved up to the step $n+1$, we have

$$\|\tilde{W}_{n+1}\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \leqq \sum_{k=1}^{n+1} \|\tilde{H}_k\|_{\mathfrak{s}_M+\mu(\mathtt{b})+\overline{\sigma}}^{\mathrm{Lip}(\gamma)} \leqq C_* \varepsilon\gamma^{-2}$$

and thus also (8.8) holds at the step $n+1$. This completes the proof of Theorem 8.1. □

Corollary 8.4. *Let $\gamma = \varepsilon^{\mathfrak{a}}$ with $\mathfrak{a} \in (0, \mathfrak{a}_0)$, $\mathfrak{a}_0 := 1/\tau_2$ with τ_2 defined as in (8.7), and $K_0 = 1/\gamma$. Then there is $\varepsilon_0 > 0$ so that for any $0 < \varepsilon \leq \varepsilon_0$ the following holds:*

(i) *there exists a function $U_\infty(\omega) = (\breve{\iota}_\infty(\omega), \zeta_\infty(\omega))$, $\omega \in \Omega$, satisfying*

$$\|U_\infty - U_0\|_{\overline{s}}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2}, \quad \|U_\infty - \tilde{U}_n\|_{\overline{s}}^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2} K_n^{-\mathtt{a}_2}, \ n \geqq 1, \tag{8.36}$$

where $\overline{s} := \mathfrak{s}_M + \mu(\mathtt{b}) + \overline{\sigma}$ (with $\mathfrak{s}_M$, $\mu(\mathtt{b})$, $\overline{\sigma}$ fixed in (2.55), (7.6), and respectively, (8.2));

(ii) *for any* ω *in the set*

$$\Omega_\varepsilon := \bigcap_{n \geqq 0} \mathcal{G}_n = \mathcal{G}_0 \cap \bigcap_{n \geqq 1} \boldsymbol{\Omega}^{\gamma}_{n+1}(\tilde{\iota}_{n-1}) \stackrel{(7.61)}{=} \mathcal{G}_0 \cap \Big[\bigcap_{n \geqq 1} \Lambda^{\gamma}_n(\tilde{\iota}_{n-1}) \Big] \cap \Big[\bigcap_{n \geqq 1} \Omega^{\gamma}_n(\tilde{\iota}_{n-1}) \Big], \tag{8.37}$$

the torus embedding $\breve{\iota}_\omega(\varphi)$ *solves* (4.12).

Proof. For any $0 < \varepsilon \leq \varepsilon_0$ with ε_0 small enough, the smallness condition $\varepsilon K_0^{\tau_2} < \delta_0$ in (8.7) holds and Theorem 8.1 applies. By Theorem 8.1-$(\mathcal{P}1)_n$ the sequence $(\tilde{U}_n)_{n \geq 1}$ converges as $n \to \infty$ to a function $U_\infty(\omega)$, satisfying (8.36). By Theorem 8.1-$(\mathcal{P}2)_n$, for any $\omega \in \Omega_\varepsilon$, we have that $\mathcal{F}_\omega(U_\infty(\omega)) = 0$. Formula (5.5) implies that $\zeta_\infty(\omega) = 0$ for any $\omega \in \Omega_\varepsilon$ and item (ii) is proved. □

In order to complete the proof of Theorem 4.1 it only remains to establish the measure estimate (4.10).

8.2. Measure Estimates

The measure estimate (4.10) of Theorem 4.1 for the subset $\Omega_\varepsilon = \cap_{n \geqq 0} \mathcal{G}_n$ defined in (8.37) of nonresonant frequency vectors will be deduced from the measure estimate of Proposition 8.7, using that by Lemmas 8.5 and 8.6, the set $\Omega \setminus \Omega_\varepsilon$ is included in $\Omega \setminus \Omega^{\gamma}_\infty$, where Ω^{γ}_∞ is introduced in Lemma 8.6. The main result of this section is Proposition 8.7. In all of this section, the *assumptions of Theorem* 8.1 *hold with the constants given as in* (8.2)–(8.7) .

In order to prove Lemmata 8.5 and 8.6, we first recall that by (8.10), $\mathcal{G}_0 = \Omega$ and for $n \geq 1$, $\mathcal{G}_n := \mathcal{G}_{n-1} \cap \boldsymbol{\Omega}^{\gamma}_n(\tilde{\iota}_{n-1})$ where $\boldsymbol{\Omega}^{\gamma}_n(\iota) = \Omega^{\gamma}_n(\iota) \cap \Lambda^{\gamma}_n(\iota)$ (cf. (7.61), (7.18), (7.58)).

Lemma 8.5. *For any* $n \geq 0$, *the set*

$$\mathcal{G}_\infty := \mathcal{G}_0 \cap \Big[\bigcap_{n \geqq 1} \Omega^{2\gamma}_n(\iota_\infty) \Big] \cap \Big[\bigcap_{n \geqq 1} \Lambda^{2\gamma}_n(\iota_\infty) \Big] \tag{8.38}$$

is contained in $\mathcal{G}_n$ *and hence* $\mathcal{G}_\infty \subseteq \bigcap_{n \geqq 0} \mathcal{G}_n$.

Proof. We apply the inclusion property (7.28). By (8.36), (5.27), we have, for any $n \geqq 2$,

$$C(S) N^{\tau}_{n-1} \| \iota_\infty - \tilde{\iota}_{n-1} \|_{\mathfrak{s}_M + \mu(\mathtt{b}) + \overline{\sigma}} \leqq C(S) C K^{p\tau - \mathtt{a}_2}_{n-1} \varepsilon \gamma^{-2} \leqq \gamma$$

taking ε small enough, by (8.7) and using $\mathtt{a}_2 \geqq p\tau$ (see (8.3)). For $n = 1$, one also has $C(S) N^{\tau}_0 \| \iota_\infty - \tilde{\iota}_0 \|_{\mathfrak{s}_M + \mu(\mathtt{b}) + \overline{\sigma}} \leqq \gamma$ using the first inequality in (8.36) and recalling that $K_0 = \gamma^{-1}$, $\gamma = \varepsilon^{\mathfrak{a}}$ and $1 - \mathfrak{a}(3 + \tau p) > 0$ (indeed $\mathfrak{a} = 1/\tau_2$ where τ_2 is defined in (8.7) and we take $\overline{\tau} > \tau$ large enough, see Theorem 7.2). Recall also that $S = \mathfrak{s}_M + \mathtt{b}_1$ has been fixed in (8.4). Therefore (7.28) in Theorem 7.2-$(\mathbf{S3})_\nu$ gives $\Omega^{2\gamma}_n(\iota_\infty) \subseteq \Omega^{\gamma}_n(\tilde{\iota}_{n-1})$, $\forall n \geqq 1$. By similar arguments we deduce that $\Lambda^{2\gamma}_n(\iota_\infty) \subseteq \Lambda^{\gamma}_n(\tilde{\iota}_{n-1})$, and the lemma is proved. □

By Theorem 7.2-$(\mathbf{S2})_n$, it follows that for any $j \in \mathbb{S}^\perp$, $\widetilde{\mu}_j^n : \Omega \to \mathbb{R}$, $n \geq 0$, is a Cauchy sequence with respect to the norm $|\cdot|^{\mathrm{Lip}(\gamma)}$. We denote its limit by μ_j^∞,

$$\mu_j^\infty := \lim_{n\to\infty} \widetilde{\mu}_j^n(\iota_\infty), \quad \widetilde{\mu}_j^n(\iota_\infty) \equiv \widetilde{\mu}_j^n(\omega; \iota_\infty), \quad j \in \mathbb{S}^\perp. \tag{8.39}$$

By Theorem 7.2 one has for any $j \in \mathbb{S}^\perp$,

$$\mu_{-j}^\infty = -\mu_j^\infty, \quad |\mu_j^\infty - \widetilde{\mu}_j^n(\iota_\infty)|^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\gamma^{-2} N_{n-1}^{-\mathtt{a}}, \; n \geqq 0. \tag{8.40}$$

Lemma 8.6. *The set Ω_∞^γ is contained in $\mathcal{G}_\infty$ where $\mathcal{G}_\infty$ is defined in* (8.38) *and*

$$\begin{aligned}\Omega_\infty^\gamma := \Big\{\omega \in \mathtt{DC}(4\gamma,\tau) : |\omega\cdot\ell + \mu_j^\infty - \mu_{j'}^\infty| \geqq \frac{4\gamma|j^3 - j'^3|}{\langle\ell\rangle^\tau}, \; \forall(\ell, j, j') \in \mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp\times\mathbb{S}^\perp, \\ |\omega\cdot\ell + \mu_j^\infty| \geqq \frac{4\gamma|j|^3}{\langle\ell\rangle^\tau}, \; \forall(\ell, j) \in \mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp\Big\}.\end{aligned} \tag{8.41}$$

Proof. We have to verify that Ω_∞^γ is contained in each subset on the right hand side of (8.38). Since by (4.4), $\mathtt{DC}(4\gamma,\tau) \subseteq \Omega$ one has that $\Omega_\infty^\gamma \subseteq \Omega \overset{(8.10)}{=} \mathcal{G}_0$. Next we prove that $\Omega_\infty^\gamma \subseteq \Omega_n^{2\gamma}(\iota_\infty)$, $\forall n \geqq 1$. We argue by induction. Assume that $\Omega_\infty^\gamma \subseteq \Omega_n^{2\gamma}(\iota_\infty)$ for some $n \geqq 1$. For all $\omega \in \Omega_\infty^\gamma \subseteq \Omega_n^{2\gamma}(\iota_\infty)$, by (7.16), (8.39), (8.40), we get $|(\widetilde{\mu}_j^n - \widetilde{\mu}_{j'}^n)(\iota_\infty) - (\mu_j^\infty - \mu_{j'}^\infty)| \leqq C\varepsilon\gamma^{-2}N_{n-1}^{-\mathtt{a}}$. Therefore, let $(\ell, j, j') \in \mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp\times\mathbb{S}^\perp$, $|\ell| \leqq N_n$ with $(\ell, j, j') \neq (0, j, j)$ (recall (8.41)). If $j = j'$, then $\ell \neq 0$ and since $\Omega_\infty^\gamma \subseteq \mathtt{DC}(4\gamma,\tau)$ we have

$$|\omega\cdot\ell + \widetilde{\mu}_j^n(\iota_\infty) - \widetilde{\mu}_{j'}^n(\iota_\infty)| = |\omega\cdot\ell| \geqq 4\gamma\langle\ell\rangle^{-\tau}.$$

In case $j \neq j'$, one has

$$\begin{aligned}|\omega\cdot\ell + \widetilde{\mu}_j^n(\iota_\infty) - \widetilde{\mu}_{j'}^n(\iota_\infty)| &\geqq |\omega\cdot\ell + \mu_j^\infty - \mu_{j'}^\infty| - C\varepsilon\gamma^{-2}N_{n-1}^{-\mathtt{a}} \\ &\geqq \frac{4\gamma|j^3 - j'^3|}{\langle\ell\rangle^\tau} - C\varepsilon\gamma^{-2}N_{n-1}^{-\mathtt{a}} \geqq \frac{2\gamma|j^3 - j'^3|}{\langle\ell\rangle^\tau},\end{aligned}$$

provided $\frac{1}{2}C\varepsilon\gamma^{-3}N_{n-1}^{-\mathtt{a}}N_n^\tau \leqq 1$ (note that since $j \neq j'$, $|j^3 - j'^3| \geqq 1$). The latter condition is fullfilled by (7.6), (8.7), by taking $\bar\tau > \tau$ large enough. In conclusion we have proved that $\Omega_\infty^\gamma \subseteq \Omega_{n+1}^{2\gamma}(\iota_\infty)$. Similarly we prove that $\Omega_\infty^\gamma \subseteq \Lambda_n^{2\gamma}(\iota_\infty)$ for all $n \geqq 1$. □

In view of Lemmata 8.5 and 8.6, it suffices to estimate the Lebesgue measure $|\Omega \setminus \Omega_\infty^\gamma|$ of $\Omega \setminus \Omega_\infty^\gamma$ instead of the one of $\Omega \setminus \bigcap_{n\geqq 0}\mathcal{G}_n$.

Proposition 8.7. (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 $|\Omega \setminus \Omega_\infty^\gamma| \lesssim \gamma^{\mathfrak{a}}$.*

Proof. By (8.41), we have

$$\Omega \setminus \Omega_\infty^\gamma = \Omega \setminus \mathtt{DC}(4\gamma,\tau) \;\cup \bigcup_{(\ell,j,j')\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp\times\mathbb{S}^\perp,(\ell,j,j')\neq(0,j,j)} \mathcal{R}_{\ell,j,j'} \;\cup \bigcup_{(\ell,j)\in\mathbb{Z}^{\mathbb{S}_+}\times\mathbb{S}^\perp} \mathcal{Q}_{\ell,j} \tag{8.42}$$

where $\mathcal{R}_{\ell,j,j'}$, $\mathcal{Q}_{\ell,j}$ denote the 'resonant' sets

$$\mathcal{R}_{\ell,j,j'} := \Big\{\omega \in \mathtt{DC}(4\gamma,\tau) \ : \ |\omega\cdot\ell + \mu_j^\infty - \mu_{j'}^\infty| < \frac{4\gamma|j^3 - j'^3|}{\langle\ell\rangle^\tau}\Big\},$$
$$\mathcal{Q}_{\ell,j} := \Big\{\omega \in \mathtt{DC}(4\gamma,\tau) : |\omega\cdot\ell + \mu_j^\infty| < \frac{4\gamma|j|^3}{\langle\ell\rangle^\tau}\Big\}.$$

Note 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 8.7 we shall use the following asymptotic properties of $\mu_j^\infty(\omega)$. For any ω in $\mathtt{DC}(4\gamma,\tau)$, we have $\widetilde{\mu}_j^0(\iota_\infty) = \mu_j^0(\iota_\infty)$ (for simplicity $\mu_j^0(\iota_\infty) \equiv \mu_j^0(\omega;\iota_\infty)$) and we write $\mu_j^\infty(\omega) = \mu_j^0(\iota_\infty) + r_j^\infty(\omega)$, where by (7.2)

$$\mu_j^0(\iota_\infty) = m_3^\infty(\omega)(2\pi j)^3 - m_1^\infty(\omega)2\pi j - q_j(\omega), \quad m_3^\infty := m_3(\iota_\infty), \quad m_1^\infty := m_1(\iota_\infty).$$

On $\mathtt{DC}(4\gamma,\tau)$, the following estimates hold:

$$\begin{aligned} &|m_3^\infty + 1|^{\mathrm{Lip}(\gamma)} \overset{(6.20)}{\lesssim} \varepsilon, \quad |m_1^\infty|^{\mathrm{Lip}(\gamma)} \overset{(6.68)}{\lesssim} \varepsilon\gamma^{-2}, \\ &\sup_{j\in\mathbb{S}^\perp} |j||q_j|^{\mathrm{sup}}, \ \sup_{j\in\mathbb{S}^\perp} |j||q_j|^{\mathrm{lip}} \overset{(7.4)}{\lesssim} 1, \quad |r_j^\infty|^{\mathrm{Lip}(\gamma)} \overset{(8.40)}{\lesssim} \varepsilon\gamma^{-2}. \end{aligned} \tag{8.43}$$

From the latter estimates one infers the following standard lemma see [2, Lemma 5.3]). □

Lemma 8.8. (i) *If $\mathcal{R}_{\ell,j,j'} \neq \emptyset$, then $|j^3 - j'^3| \leqq C\langle\ell\rangle$ for some $C > 0$. In particular one has $j^2 + j'^2 \leqq C\langle\ell\rangle$.*
(ii) *If $\mathcal{Q}_{\ell,j} \neq \emptyset$, then $|j|^3 \leqq C\langle\ell\rangle$ for some $C > 0$.*

Lemma 8.8 can be used to estimate $|\mathcal{R}_{\ell,j,j'}|$ and $|\mathcal{Q}_{\ell,j}|$ for $|\ell|$ sufficiently large.

Lemma 8.9. (i) *If $\mathcal{R}_{\ell,j,j'} \neq \emptyset$, then there exists $C_1 > 0$ with the following property: if $|\ell| \geqq C_1$, 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 $|\ell| \geqq C_1$, then $|\mathcal{Q}_{\ell,j}| \lesssim \gamma|j|^3\langle\ell\rangle^{-(\tau+1)}$.*

Proof of Lemma 8.9. 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 $s \mapsto \phi_{\ell,j,k}(s)$,

$$\phi_{\ell,j,j'}(s) := f_{\ell,j,j'}\big(\bar\omega + s\frac{\ell}{|\ell|}\big), \quad f_{\ell,j,j'}(\omega) := \omega\cdot\ell + \mu_j^\infty(\omega) - \mu_{j'}^\infty(\omega).$$

Using that by Lemma 8.8, $|j^3 - j'^3| \leqq C\langle\ell\rangle$, one infers from (8.43) that for $\varepsilon\gamma^{-2}$ small enough and $|\ell| \geqq C_1$ with C_1 large enough, $|\phi_{\ell,j,j'}(s_2) - \phi_{\ell,j,j'}(s_1)| \geqq \frac{|\ell|}{2}|s_2 - s_1|$. Since $\mathtt{DC}(4\gamma,\tau) \subseteq \Omega$ 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. □

It remains to estimate the Lebesgue measure of the resonant sets $\mathcal{R}_{\ell,j,j'}$ and $\mathcal{Q}_{\ell,j}$ for $|\ell| \leqq C_1$.

Lemma 8.10. *Assume that $|\ell| \leqq C_1$ 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 $|j|, |j'| \leqq C_2$ 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 $|j| \leqq C_2$ and $|\mathcal{Q}_{\ell,j}| \lesssim \gamma^{\mathfrak{a}}$.*

Proof of Lemma 8.10. We only prove item (i) since item (ii) can be proved in a similar way. If $|\ell| \leqq C_1$ and $\mathcal{R}_{\ell,j,j'} \neq \emptyset$, Lemma 8.8-(i) implies that there is a constant C_2 such that $|j|, |j'| \leqq j^2 + j'^2 \leqq C_2$. For $\varepsilon\gamma^{-3}$ small enough one sees, using (8.43), the definition (7.2) of μ_j^0, and the bounds $|\ell| \leqq C_1$, $|j|, |j'| \leqq C_2$, that $|\mu_j^\infty - \omega_j^{kdv}| \lesssim \varepsilon\gamma^{-2} \lesssim \gamma$, implying that for some constant $C_3 > 0$,

$$\mathcal{R}_{\ell,j,j'} \subset \left\{\omega \in \Omega \ : \ |\omega \cdot \ell + \omega_j^{kdv}(\nu(\omega),0) - \omega_{j'}^{kdv}(\nu(\omega),0)| \leqq C_3\gamma\right\}. \tag{8.44}$$

By Lemma 3.9, 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 Proposition 3.1 in [11]), we deduce that the measure of the set on the right hand side of (8.44) is smaller than $\gamma^{\mathfrak{a}}$ for some $\mathfrak{a} \in (0,1)$ and γ small enough. □

We are now in position to finish the proof of Proposition 8.7. By (8.42) and Lemmata 8.9–8.10 we have

$$|\Omega \setminus \Omega_\infty^\gamma| \lesssim \gamma^{\mathfrak{a}} + \gamma \sum_{|\ell| \geqq C_1, |j|,|j'| \leqq C\langle \ell \rangle} \langle \ell \rangle^{-\tau} \lesssim \gamma^{\mathfrak{a}},$$

where we used the assumption that $\tau - 2 > |\mathbb{S}_+|$. □

8.3. Proofs of Theorems 4.1 and 1.1

In this section, we complete the proof of Theorem 4.1 and then derive from it Theorem 1.1.

Proof of Theorem 4.1. In view of Corollary 8.4-(ii) for any $\omega \in \Omega_\varepsilon$, the embedded torus $\breve{\iota}_\omega(\mathbb{T}^{\mathbb{S}_+})$ is invariant under the flow of the Hamiltonian vector field $X_{\mathcal{H}_\varepsilon(\cdot,\nu)}$ and is filled by quasi-periodic solutions with frequency $\omega = -\omega^{kdv}(\nu)$. The bound (4.11) follows by (8.36). The linear stability of the quasi-periodic solution $\breve{\iota}_\omega(\omega t)$ follows by standard arguments as in [1,3,10] since for any $\omega \in \cap_{n \geqq 0}\Omega_n^\gamma(\iota_\infty)$ (cf. (7.30)) we obtain by Theorem 7.3 the complete diagonalization of $\mathrm{L}_0(\iota_\infty)$. It thus remains to prove the measure estimate (4.10) of Theorem 4.1. Since by Lemma 8.5 one has $\mathcal{G}_\infty \subset \Omega_\varepsilon$ and by Lemma 8.6, $\Omega_\infty^\gamma \subseteq \mathcal{G}_\infty$ the claimed estimates of $|\Omega \setminus \Omega_\varepsilon|$ in (4.10) follow from the estimates of $|\Omega \setminus \Omega_\infty^\gamma|$ established in Proposition 8.7. □

Proof of Theorem 1.1. Theorem 1.1 is in fact a reformulation of Theorem 4.1. Choose $\bar{s}$, ε_0, and Ω_ε, $0 < \varepsilon \le \varepsilon_0$, as in Theorem 4.1. Using that $-\omega^{kdv} : \Xi \to \Omega$, $\nu \mapsto -\omega^{kdv}(\nu)$, is a diffeomorphism (cf. (1.13)), we define for any $0 < \varepsilon \le \varepsilon_0$ the set

$$\Xi_\varepsilon := \Big\{ \nu \in \Xi : \nu = (\omega^{kdv})^{-1}(-\omega), \quad \omega \in \Omega_\varepsilon \Big\}.$$

By Theorem 4.1, for any $\nu \in \Xi_\varepsilon$, there exists a torus embedding with lift $\breve{\iota}_\omega : \mathbb{R}^{\mathbb{S}_+} \to \mathbb{R}^{\mathbb{S}_+} \times \mathbb{R}^{\mathbb{S}_+} \times H^{\bar{s}}_\perp(\mathbb{T}_1)$ of the form $\breve{\iota}_\omega(\varphi) = (\varphi, 0, 0) + \iota_\omega(\varphi)$ and $\omega \equiv \omega_\varepsilon(\nu) = -\omega^{kdv}(\nu)$, satisfying (cf. (4.11))

$$\|\iota_\omega\|_{\bar{s}} = O(\varepsilon\gamma^{-2}), \quad \gamma = \varepsilon^{\mathfrak{a}}, \ 0 < \mathfrak{a} \ll 1, \tag{8.45}$$

so that $\breve{\iota}_\omega(\omega t)$ is a linearly stable, quasi-periodic solution of (4.1). By the measure estimate (4.10) one has $\lim_{\varepsilon\to 0} |\Xi \setminus \Xi_\varepsilon| = 0$, which proves (1.16). By the definition of the symplectic diffeomorphism Ψ_ν (cf. Theorem 3.2) and the one of $\mathcal{H}_\varepsilon$ (cf. (4.2)), it then follows that

$$u_\varepsilon(\omega t, x; \nu) := \Psi_\nu\big((\omega t, 0, 0) + \iota_\omega(\omega t)\big) \tag{8.46}$$

is a quasi-periodic solution of the perturbed KdV equation (1.6). Furthermore, by (3.8) in Theorem 3.2 and (3.5), the finite gap solution $t \mapsto q(\omega t, x; \nu)$ of the KdV equation (1.1) (cf. (1.10)) satisfies

$$q(\omega t, x; \nu) = \Psi_\nu(\omega t, 0, 0). \tag{8.47}$$

It then follows from (8.45)–(8.47) that $\|u_\varepsilon(\cdot, \cdot; \nu) - q(\cdot, \cdot; \nu)\|_{\bar{s}} \lesssim \varepsilon^{1-\mathfrak{b}}$ with $\mathfrak{b} = 2\mathfrak{a}$, proving (1.17). □

Acknowledgements. This project was motivated by questions raised by S. Kuksin and V. Zakharov. We would like to thank them for their input. We would also like to thank M. Procesi for very valuable feedback. Part of this work was written during the stay of M. Berti at FIM at ETHZ. We thank FIM for the kind hospitality and support. 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.).

References

1. Baldi, P., Berti, M., Haus, E., Montalto, R.: Time quasi-periodic gravity water waves in finite depth. *Invent. Math.* **214**(2), 739–911, 2018
2. Baldi, P., Berti, M., Montalto, R.: KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation. *Math. Ann.* **359**, 471–536, 2014
3. Baldi, P., Berti, M., Montalto, R.: KAM for autonomous quasi-linear perturbations of KdV. *Ann. Inst. H. Poincaré Analyse Non Lin.* **33**(6), 1589–1638, 2016
4. Baldi, P., Berti, M., Montalto, R.: KAM for autonomous quasi-linear perturbations of mKdV. *Bollettino Unione Matematica Italiana* **9**, 143–188, 2016

5. Bambusi, D., Grébert, B., Maspero, A., Robert, D.: Growth of Sobolev norms for abstract linear Schrödinger equations. *J. Eur. Math. Soc. (JEMS)* (2017). https://doi.org/10.4171/JEMS/10177
6. Berti, M., Biasco, L., Procesi, M.: KAM theory for the Hamiltonian DNLW. *Ann. Sci. Éc. Norm. Supér. (4)* **46**(fascicule 2), 301–373, 2013
7. Berti, M., Biasco, L., Procesi, M.: KAM theory for the reversible derivative wave equation. *Arch. Ration. Mech. Anal*. **212**, 905–955, 2014
8. Berti, M., Bolle, P.: *A Nash–Moser approach to KAM theory*, Fields Institute Communications, special volume "Hamiltonian PDEs and Applications", 255–284, 2015.
9. Berti, M., Kappeler, T., Montalto, R.: Large KAM tori for perturbations of the defocusing NLS equation. *Astérisque* **403**, 2018
10. Berti, M., Montalto, R.: Quasi-periodic standing wave solutions of gravity capillary standing water waves. *Mem. Am. Math. Soc.*, vol. 263, number 1273, 2020.
11. Biasco, L., Coglitore, F.: Periodic orbits accumulating onto elliptic tori for the (N + 1)-body problem. *Celest. Mech. Dyn. Astr*. **101**, 349, 2008
12. Bikbaev, R., Kuksin, S.: On the parametrization of finite-gap solutions by frequency vector and wave number vectors and a theorem by I. Krichever. *Lett. Math. Phys*. **28**, 115–122, 1993
13. Dubrovin, B., Krichever, I., Novikov, S.: Integrable Systems I, in *Dynamical Systems IV*, Encyclopedia of Mathematical Sciences vol. 4 (Eds. Arnold V., Novikov S.) Springer, 173–280, 1990.
14. Eliasson, H., Grébert, B., Kuksin, S.: KAM for the nonlinear beam equation. *Geom. Funct. Anal*. **26**, 1588–1715, 2016
15. Eliasson, H., Kuksin, S.: KAM for non-linear Schrödinger equation. *Ann. Math*. **172**, 371–435, 2010
16. Faddeev, L., Takhtajan, L.: *Hamiltonian Methods in the Theory of Solitons*. Springer, Berlin 1987
17. Feola, R., Procesi, M.: Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations. *J. Differ. Equ*. **259**(7), 3389–3447, 2015
18. Gardner, C., Greene, J., Kruskal, M., Miura, R.: Korteweg–deVries equation and generalization. VI. Methods for exact solution. *Commun. Pure Appl. Math*. **27**, 97–133, 1974
19. Giuliani, F.: Quasi-periodic solutions for quasi-linear generalized KdV equations. *J. Differ. Equ*. **262**(10), 5052–5132, 2017
20. Kappeler, T., Montalto, R.: Normal form coordinates for the KdV equation having expansions in terms of pseudo-differential operators. *Commun. Math. Phys*. **375**, 833–913, 2020. https://doi.org/10.1007/s00220-019-03498-1
21. Kappeler, T., Pöschel, J.: *KdV & KAM*. Springer, Berlin 2003
22. Kappeler, T., Schaad, B., Topalov, P.: Qualitative features of periodic solutions of KdV. *Commun. Partial. Differ. Equ*. **38**(9), 1626–1673, 2013
23. Kappeler, T., Schaad, B., Topalov, P.: Semi-linearity of the nonlinear Fourier transform of the defocusing NLS equation. *Int. Math. Res. Not*. **23**, 7212–7229, 2016
24. Kruskal, M., Zabusky, N.: Interactions of 'solitons" in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett*. **15**, 240–243, 1965
25. Kuksin, S.: A KAM theorem for equations of the Korteweg–de Vries type. *Rev. Math. Phys*. **10**(3), 1–64, 1998
26. Kuksin, S.: *Analysis of Hamiltonian PDEs*. Oxford University Press, Oxford 2000
27. Kuksin, S., Perelman, G.: Vey theorem in infinite dimensions and its application to KdV. *Discrete Contin. Dyn. Syst*. **27**(1), 1–24, 2010
28. Lax, P.: Integrals of nonlinear equations of evolution and solitary waves. *Commun. Pure Appl. Math*. **21**, 467–490, 1968
29. Liu, J., Yuan, X.: A KAM theorem for Hamiltonian partial differential equations with unbounded perturbations. *Commun. Math. Phys* **307**(3), 629–673, 2011

30. Miura, R., Gardner, C., Kruskal, M.: Korteweg–de Vries equation and generalizations. II. Existence of conservation laws and constants of motion. *J. Math. Phys*. **9**, 1204–1209, 1968
31. Stein, E.M.: *Singular Integrals and Differentiability Properties of Functions. Princeton Mathematical Series, No. 30*. Princeton University Press, Princeton 1970
32. Zhang, J., Gao, M., Yuan, X.: KAM tori for reversible partial differential equations. *Nonlinearity* **24**, 1189–1228, 2011

M. Berti
SISSA,
Trieste
Italy.
e-mail: berti@sissa.it

and

T. Kappeler
University of Zurich,
Zurich
Switzerland.
e-mail: thomas.kappeler@math.uzh.ch

and

R. Montalto
University of Milan,
Via Saldini 50,
Milan
20133 Italy.
e-mail: riccardo.montalto@unimi.it

(*Received October 15, 2019 / Accepted November 6, 2020*)
Published online January 4, 2021