CrossMark

Quasi-periodic solutions of forced Kirchhoff equation

Riccardo Montalto

Abstract. In this paper we prove the existence and the stability of small-amplitude quasi-periodic solutions with Sobolev regularity, for the 1-dimensional forced Kirchhoff equation with periodic boundary conditions. This is the first KAM result for a quasi-linear wave-type equation. The main difficulties are: (i) the presence of the highest order derivative in the nonlinearity which does not allow to apply the classical KAM scheme, (ii) the presence of double resonances, due to the double multiplicity of the eigenvalues of $-\partial_{xx}$. The proof is based on a Nash–Moser scheme in Sobolev class. The main point concerns the invertibility of the linearized operator at any approximate solution and the proof of tame estimates for its inverse in high Sobolev norm. To this aim, we conjugate the linearized operator to a 2×2, time independent, block-diagonal operator. This is achieved by using *changes of variables* induced by diffeomorphisms of the torus, *pseudo-differential* operators and a KAM *reducibility* scheme in Sobolev class.

Mathematics Subject Classification. 37K55, 35L72.

Keywords. Kirchhoff equation, Quasi-linear PDEs, Quasi-periodic solutions, Infinite-dimensional dynamical systems, KAM for PDEs, Nash–Moser theory.

Contents

Birkhäuser

1. Introduction and main results

We consider the Kirchhoff equation in 1-dimension with periodic boundary conditions

$$\partial_{tt} v - \left(1 + \int_{\mathbb{T}} (\partial_x v)^2 \, dx\right) \partial_{xx} v = \delta f(\omega t, x), \quad x \in \mathbb{T}, \qquad (1.1)$$

where $\mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$ is the 1-dimensional torus, $\delta > 0$ is a small parameter, $f \in \mathcal{C}^q(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$ and $\omega \in \Omega \subseteq \mathbb{R}^\nu$, with Ω bounded. Our aim is to prove the existence and the linear stability of small-amplitude quasi-periodic solutions with Sobolev regularity, for δ small enough and for ω in a suitable *Cantor like set* of parameters with asymptotically full Lebesgue measure.

The Kirchhoff equation has been introduced for the first time in 1876 by Kirchhoff, in dimension 1, without forcing term and with Dirichlet boundary conditions, namely

$$\partial_{tt} v - \left(1 + \int_0^\pi (\partial_x v)^2 \, dx\right) \partial_{xx} v = 0, \qquad v(t, 0) = v(t, \pi) = 0, \qquad (1.2)$$

to describe the transversal free vibrations of a clamped string in which the dependence of the tension on the deformation cannot be neglected. It is a quasi-linear PDE, namely the nonlinear part of the equation contains as many derivatives as the linear differential operator. The Cauchy problem for the Kirchhoff equation (also in higher dimension) has been extensively studied, starting from the pioneering paper of Bernstein [9]. Both local and global

existence results have been established for initial data in Sobolev and analytic class, see [2,3,23,24,32,40,42] and the recent survey [41].

Concernig the existence of periodic solutions, Kirchhoff himself observed that the equation (1.2) admits a sequence of *normal modes*, namely solutions of the form $v(t,x) = v_j(t)\sin(jx)$ where $v_j(t)$ is 2π-periodic. Under the presence of the forcing term $f(t,x)$ the *normal modes* do not persist,[1] since, expanding $v(t,x) = \sum_j v_j(t)\sin(jx)$, $f(t,x) = \sum_j f_j(t)\sin(jx)$, all the components $v_j(t)$ are coupled in the integral term $\int_{\mathbb{T}}(\partial_x v)^2\,dx$ and the Eq. (1.2) is equivalent to the infinitely many coupled ODEs

$$v_j''(t) + j^2 v_j(t)\left(1 + \sum_k k^2 |v_k(t)|^2\right) = f_j(t), \qquad j = 1,2,\dots.$$

The existence of periodic solutions for the Kirchhoff equation, also in higher dimension, has been proved by Baldi [4], both for Dirichlet boundary conditions ($v = 0$ on $\partial\Omega$) and for periodic boundary conditions ($\Omega = \mathbb{T}^d$). This result is proven via Nash–Moser method and thanks to the special structure of the nonlinearity (it is diagonal in space), the linearized operator at any approximate solution can be inverted by Neumann series. This approach does not imply the linear stability of the solutions and it does not work in the quasi-periodic case, since the *small divisors* problem is more difficult.

In general, the presence of derivatives in the nonlinearity makes uncertain the existence of global (even not periodic or quasi-periodic) solutions, see for example the non-existence results in [34,37] for the equation $v_{tt} - a(v_x)v_{xx} = 0$, $a > 0$, $a(v) = v^p$, $p \geq 1$, near zero.

Concerning the existence of periodic solutions, the first bifurcation result is due to Rabinowitz [44,45], for fully nonlinear forced wave equations with a small dissipation term

$$v_{tt} - v_{xx} + \alpha v_t + \varepsilon f(t,x,v,v_t,v_x,v_{tt},v_{tx},v_{xx}) = 0, \quad x \in \mathbb{T}, \qquad \alpha \neq 0$$

with frequency $\omega = 1$ (2π-periodic solutions). Then Craig [20] proved the existence of small-amplitude periodic solutions, for a large set of frequencies ω, for the autonomous *pseudo differential* equation

$$\partial_{tt}v - \partial_{xx}v = a(x)v + b(x,|D|^\beta v), \qquad \beta < 1$$

and Bourgain [18] obtained the same result for the equation $\partial_{tt}v - \partial_{xx}v + mv + (\partial_t v)^2 = 0$. The above results are based on a Newton–Nash–Moser scheme and a Lyapunov–Schmidt decomposition.

For the water waves equations, which are fully nonlinear PDEs, we mention the pioneering work of Iooss–Plotnikov–Toland [29] about existence of time periodic standing waves, and of Iooss–Plotnikov [30,31] for 3-dimensional travelling water waves. The key idea is to use diffeomorphisms of the torus $\mathbb{T}^2$ and pseudo-differential operators, in order to conjugate the linearized operator to one with constant coefficients plus a sufficiently smoothing remainder. This is enough to invert the whole linearized operator (at any approximate solution)

---

[1] This is true except in the case where f is uni-modal, i.e. $f(t,x) = f_k(t)\sin(kx)$ for some $k \geq 1$.

by Neumann series. Very recently Baldi [5] has further developed the techniques of [29], proving the existence of periodic solutions for fully nonlinear autonomous, reversible Benjamin–Ono equations. We mention also the recent paper of Alazard and Baldi [1] concerning the existence of periodic standing-wave solutions of the water waves equations with surface tension.

These methods do not work for proving the existence of quasi-periodic solutions and they do not imply the linear stability.

Existence of quasi-periodic solutions of PDEs (that we shall call in a broad sense KAM theory) with unbounded perturbations (the nonlinearity contains derivatives) has been developed by Kuksin [35] for KdV and then Kappeler–Pöschel [33]. The key idea is to work with a variable coefficients normal form along the KAM scheme. The homological equations, arising at each step of the iterative scheme, are solved thanks to the so called Kuksin Lemma, see Chapter 5 in [33]. This approach has been improved by Liu–Yuan [38,39] who proved a stronger version of the Kuksin Lemma and applied it to derivative NLS and Benjamin–Ono equations. These methods apply to dispersive PDEs like KdV, derivative NLS but not to derivative wave equation (DNLW) which contains first order derivatives in the nonlinearity. KAM theory for DNLW equation has been recently developed by Berti–Biasco–Procesi in [10] for Hamiltonian and in [11] for reversible equations. The key ingredient is to provide a sufficiently accurate asymptotic expansion of the perturbed eigenvalues which allows to impose the *Second order Melnikov* conditions. This is achieved by introducing the notion of quasi-Töplitz vector field which has been developed by Procesi–Xu [43] and it is inspired to the Töplitz–Lipschitz property developed by Eliasson–Kuksin in [25,26]. Existence of quasi-periodic solutions can be also proved by imposing only *first order Melnikov conditions*. This method has been developed, for PDEs in higher space dimension, by Bourgain in [16,17,19] for analytic NLS and NLW, extending the result of Craig–Wayne [21] for semilinear 1-dimensional wave equation. This approach is based on the so-called *multiscale analysis* of the linearized equations and it has been recently improved by Berti–Bolle [12,13] for NLW, NLS with differentiable nonlinearity and by Berti–Corsi–Procesi [14] on compact Lie-groups. It is especially convenient in the case of high multiplicity of the eigenvalues, since the second order Melnikov conditions are violated. As a consequence of having imposed only the first order Melnikov conditions, this method does not provide any information about the linear stability of the quasi-periodic solutions, since the linearized equations have variable coefficients. Indeed there are very few results concerning the existence and linear stability of quasi-periodic solutions in the case of multiple eigenvalues. We mention Chierchia–You [22], for analytic 1-dimensional NLW equation with periodic boundary conditions (double eigenvalues) and in higher space dimension Eliasson–Kuksin [26] for analytic NLS.

All the aforementioned KAM results concern *semi-linear* PDEs, namely PDEs in which the order of nonlinear part of the vector field is strictly smaller than the order of the linear part. For quasi-linear (either fully nonlinear) PDEs, the first KAM results have been recently proved by Baldi–Berti–Montalto in

[6–8] for perturbations of Airy, KdV and mKdV equations, by Feola–Procesi [28] for fully nonlinear reversible Schrödinger equation and by Feola [27] for quasi-linear Hamiltonian Schrödinger equation. For the water waves equations with surface tension, the existence of quasi-periodic standing wave solutions has been recently proved by Berti–Montalto [15].

The key analysis of the present paper concerns the linearized operator obtained at any step of the Nash–Moser scheme. The main purpose is to reduce the linearized operator to a 2×2 time independent block diagonal operator. This cannot be achieved by implementing directly a KAM *reducibility* scheme, since the constant coefficients part of the linearized operator has the same order as the non-constant part, implying that the *loss of derivatives* accumulates quadratically along the iterative scheme. In order to overcome this problem, we perform some transformations which reduce the order of the derivatives in the perturbation but not its size. We use *quasi-periodic reparametrization of time* and *pseudo differential operators* to reduce the linearized operator to a diagonal operator plus a one smoothing remainder, see (1.14). At this point we perform a KAM reducibility scheme that reduces quadratically the size of the remainder at each step of the iteration. Note that, because of the double multiplicity of the eigenvalues $|j|^2$, $j \in \mathbb{Z}$ of the operator $-\partial_{xx}$, the linearized operator cannot be completely diagonalized. This problem is overcome by working with a 2×2 block diagonal *normal form* along the iteration, which is obtained by pairing the space Fourier modes j and $-j$. This strategy has been also developed by Feola [27] for quasi-linear Hamiltonian NLS. We will explain more precisely our procedure in Sect. 1.1.

We now state precisely the main results of this paper. Rescaling the variable $v \mapsto \delta^{\frac{1}{3}} v$ and writing the Eq. (1.1) as a first order system, we get the PDE

$$\begin{cases} \partial_t v = p \\ \partial_t p = \left(1 + \varepsilon \int_{\mathbb{T}} (\partial_x v)^2 \, dx\right) \partial_{xx} v + \varepsilon f(\omega t, x), \end{cases} \qquad \varepsilon := \delta^{\frac{2}{3}} \qquad (1.3)$$

which is a Hamiltonian equation of the form

$$\begin{cases} \partial_t v = \nabla_p H(\omega t, v, p) \\ \partial_t p = -\nabla_v H(\omega t, v, p) \end{cases} \qquad (1.4)$$

whose Hamiltonian is

$$H(\omega t, v, p)$$
$$:= \frac{1}{2} \int_{\mathbb{T}} \left(p^2 + (\partial_x v)^2\right) dx + \varepsilon \left(\frac{1}{2} \int_{\mathbb{T}} (\partial_x v)^2 \, dx\right)^2 - \varepsilon \int_{\mathbb{T}} f(\omega t, x) v \, dx. \quad (1.5)$$

In (1.4), $\nabla_p H$ and $\nabla_v H$ denote the L^2-gradients of the Hamiltonian H with respect to p and v.

We look for quasi-periodic solutions $(v(\omega t, x), p(\omega t, x))$, $v, p : \mathbb{T}^\nu \times \mathbb{T} \to \mathbb{R}$ of the Eq. (1.3). This is equivalent to find zeros $(v(\varphi, x), p(\varphi, x))$ of the nonlinear operator

$$F(\varepsilon, \omega, v, p) := \begin{pmatrix} \omega \cdot \partial_\varphi v - p \\ \omega \cdot \partial_\varphi p - \left(1 + \varepsilon \int_{\mathbb{T}} (\partial_x v)^2 \, dx\right) \partial_{xx} v - \varepsilon f \end{pmatrix} \qquad (1.6)$$

in the Sobolev space $H^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2) = H^s(\mathbb{T}^{\nu+1}, \mathbb{R}) \times H^s(\mathbb{T}^{\nu+1}, \mathbb{R})$ where
$H^s(\mathbb{T}^{\nu+1}, \mathbb{R})$

$$:= \left\{ v(\varphi, x) = \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{Z}}} \widehat{v}_j(\ell) e^{\mathrm{i}(\ell \cdot \varphi + jx)} : \|v\|_s^2 := \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{Z}}} \langle \ell, j \rangle^{2s} |\widehat{v}_j(\ell)|^2 < +\infty \right\},$$

$$(1.7)$$

$\langle \ell, j \rangle := \max\{1, |\ell|, |j|\}$, $|\ell| := \max_{i=1,\ldots,\nu} |\ell_i|$. From now on we fix $s_0 := [(\nu + 1)/2] + 1$, where for any real number $x \in \mathbb{R}$, we denote by $[x]$ its integer part, so that for any $s \geq s_0$ the Sobolev space $H^s(\mathbb{T}^{\nu+1})$ is compactly embedded in the continuous functions $\mathcal{C}^0(\mathbb{T}^{\nu+1})$.

We assume that the forcing term $f \in \mathcal{C}^q(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$ has zero average, namely

$$\int_{\mathbb{T}^{\nu+1}} f(\varphi, x) \, d\varphi \, dx = 0. \qquad (1.8)$$

Now, we are ready to state the main Theorems of this paper.

Theorem 1.1. *There exist $q := q(\nu) > 0$, $s := s(\nu) > 0$ such that: for any $f \in \mathcal{C}^q(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$ satisfying the condition (1.8), there exists $\varepsilon_0 = \varepsilon_0(f, \nu) > 0$ small enough such that for all $\varepsilon \in (0, \varepsilon_0)$, there exists a Cantor set $\mathcal{C}_\varepsilon \subseteq \Omega$ of asimptotically full Lebesgue measure i.e.*

$$|\mathcal{C}_\varepsilon| \to |\Omega| \qquad as \qquad \varepsilon \to 0,$$

such that for any $\omega \in \mathcal{C}_\varepsilon$ there exist $v(\varepsilon, \omega), p(\varepsilon, \omega) \in H^s(\mathbb{T}^{\nu+1}, \mathbb{R})$, satisfying

$$\int_{\mathbb{T}^{\nu+1}} v(\varepsilon, \omega)(\varphi, x) \, d\varphi \, dx = \int_{\mathbb{T}^{\nu+1}} p(\varepsilon, \omega)(\varphi, x) \, d\varphi \, dx = 0, \qquad (1.9)$$

such that $F(\varepsilon, \omega, v(\varepsilon, \omega), p(\varepsilon, \omega)) = 0$, where the nonlinear operator F is defined in (1.6) and

$$\|v(\varepsilon, \omega)\|_s, \|p(\varepsilon, \omega)\|_s \to 0 \qquad as \qquad \varepsilon \to 0. \qquad (1.10)$$

Remark 1.1. The condition (1.8) on the forcing term f is an essential requirement to get the above existence result. Indeed, if (1.8) does not hold and if (v, p) solves $F(\varepsilon, \omega, v, p) = 0$, integrating with respect to (φ, x), we get immediately a contraddiction.

Remark 1.2. Note that the Sobolev regularity $s = s(\nu)$ of the quasi-periodic solutions and the regularity $q = q(\nu)$ of the forcing term f are required to be large enough. Actually this is due to the nature of the reduction procedure of the linearized operator developed in Sects. 4 and 5 and it is typical of such KAM results. Actually we have made no efforts to find the optimal regularity exponents.

We now discuss the precise meaning of linear stability. The linearized PDE on a quasi-periodic function $(v(\omega t, x), p(\omega t, x))$, associated to the Eq. (1.3), has the form

$$\begin{cases} \partial_t \widehat{v} = \widehat{p} \\ \partial_t \widehat{p} = a(\omega t) \partial_{xx} \widehat{v} + \mathcal{R}(\omega t)[\widehat{v}] \end{cases} \qquad (1.11)$$

where

$$a(\omega t) := 1 + \varepsilon \int_{\mathbb{T}} (v_x(\omega t, x))^2 \, dx, \tag{1.12}$$

$$\mathcal{R}(\omega t)[\widehat{v}] := -2\varepsilon v_{xx}(\omega t, x) \int_{\mathbb{T}} v_{xx}(\omega t, x)\widehat{v} \, dx. \tag{1.13}$$

In order to state precisely the next Theorem, let us introduce, for any $s \geq 0$, the Sobolev spaces

$$H^s(\mathbb{T}_x, \mathbb{R}) := \left\{ u(x) = \sum_{j \in \mathbb{Z}} u_j e^{ijx} : \|u\|_{H_x^s}^2 := \sum_{j \in \mathbb{Z}} \langle j \rangle^{2s} |u_j|^2 < +\infty \right\},$$

$$H_0^s(\mathbb{T}_x, \mathbb{R}) := \left\{ u \in H^s(\mathbb{T}_x, \mathbb{R}) : \int_{\mathbb{T}} u(x) \, dx = 0 \right\},$$

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

Theorem 1.2. (Linear stability) *There exist $\bar{\mu} > 0$, depending on ν, such that for all $S > s_0 + \bar{\mu}$, there exists $\varepsilon_0 = \varepsilon_0(S, \nu) > 0$ such that: for all $\varepsilon \in (0, \varepsilon_0)$, for all $\mathbf{v} = (v, p) \in H^S(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$ with $\|\mathbf{v}\|_{s_0 + \bar{\mu}} \leq 1$, there exists a Cantor like set $\Omega_\infty(\mathbf{v}) \subset \Omega$ such that, for all $\omega \in \Omega_\infty(\mathbf{v})$, for all $s_0 \leq s \leq S - \bar{\mu}$ the following holds: for any initial datum $(\widehat{v}^{(0)}, \widehat{p}^{(0)}) \in H^s(\mathbb{T}_x, \mathbb{R}) \times H_0^{s-1}(\mathbb{T}_x, \mathbb{R})$ the solution $t \in \mathbb{R} \mapsto (\widehat{v}(t, \cdot), \widehat{p}(t, \cdot)) \in H^s(\mathbb{T}_x, \mathbb{R}) \times H_0^{s-1}(\mathbb{T}_x, \mathbb{R})$ of the Eq. (1.11), with initial datum $\widehat{v}(0, \cdot) = \widehat{v}^{(0)}$, $\widehat{p}(0, \cdot) = \widehat{p}^{(0)}$ is stable, namely*

$$\sup_{t \in \mathbb{R}} \left(\|v(t, \cdot)\|_{H_x^s} + \|p(t, \cdot)\|_{H_x^{s-1}} \right) \leq C(s) \left(\|\widehat{v}^{(0)}\|_{H_x^s} + \|\widehat{p}^{(0)}\|_{H_x^{s-1}} \right).$$

Remark 1.3. Note that the linear stability can be proved only for initial data $\widehat{p}^{(0)}$ with zero-average in x. Indeed, the Eq. (1.11) projected on the zero Fourier mode is the ODE

$$\begin{cases} \dot{v}_0(t) = p_0(t) \\ \dot{p}_0(t) = 0 \end{cases}$$

whose solutions are

$$p_0(t) = p^{(0)}, \qquad v_0(t) = v^{(0)} + p^{(0)}t, \qquad v^{(0)}, p^{(0)} \in \mathbb{R}, \qquad \forall t \in \mathbb{R}.$$

Hence, if $p^{(0)} \neq 0$, $|v_0(t)| \to +\infty$ as $t \to \pm\infty$ and we do not have the stability.

1.1. Ideas of the proof

In this section we explain in detail the main ideas of the proof. Because of the special structure of the nonlinear operator F defined in (1.6), it is convenient to perform the decomposition (3.1), (3.2), in order to split the equation $F(\varepsilon, \omega, v, p) = 0$ into the Eqs. (3.4), (3.5). The Eq. (3.5) arises by projecting the nonlinear operator F on the zero Fourier mode in x and it is a constant coefficients PDE which can be easily solved by imposing a diophantine condition on the frequency vector ω (see Lemma 3.1). Hence, we are reduced to find zeros of the nonlinear operator $\mathcal{F}$ defined in (3.7) which is obtained by restricting F to the space of the functions with zero average in x. Theorems

1.1, 1.2 then follow by Theorem 3.1, which is based on a Nash–Moser iteration on the nonlinear map $\mathcal{F}$ on the scale of Sobolev spaces $H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$, see (2.2). The main issue concerns the invertibility of the linearized operator $\mathcal{L} = \partial_{(u,\psi)}\mathcal{F}(u,\psi)$ in (4.1) at any approximate solution and the proof of tame estimates for its inverse (see Theorem 6.1). This information is obtained by conjugating $\mathcal{L}$ to a 2×2 time-independent block diagonal operator. Such a conjugacy procedure is the content of Sects. 4 and 5.

Regularization of the linearized operator. The goal of Sect. 4 is to reduce the linearized operator $\mathcal{L}$ in (4.1) to the operator $\mathcal{L}_4$ in (4.52) which has the form

$$\mathbf{h} = (h, \overline{h}) \mapsto \omega \cdot \partial_\varphi \mathbf{h} + imT|D|\mathbf{h} + \mathcal{R}_4\mathbf{h}, \tag{1.14}$$

where $m \in \mathbb{R}$ is close to 1, $T := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $|D| = \sqrt{-\partial_{xx}}$ and $\mathcal{R}_4$ is a Hamiltonian (see Sect. 2.2.1) and 1-smoothing operator. More precisely the operator $\mathcal{R}_4$ satisfies $|\mathcal{R}_4|D||_s < +\infty$, see Lemma 4.5, where the *block-decay norm* $|\cdot|_s$ is defined in (2.80). This regularization procedure is splitted in three parts.

1. Symmetrization and complex variables. In Sect. 4.1, we symmetrize the highest order non-constant coefficients term $a(\varphi)\partial_{xx}$ in (4.1), by conjugating $\mathcal{L}$ with the transformation $\mathcal{S}$, defined in (4.9). The conjugated operator $\mathcal{L}_1$, defined in (4.13), has the form

$$\begin{pmatrix} \widehat{u} \\ \widehat{\psi} \end{pmatrix} \mapsto \begin{pmatrix} \omega \cdot \partial_\varphi + a_0(\varphi) & -a_1(\varphi)|D| \\ a_1(\varphi)|D| + \mathcal{R}^{(1)} & \omega \cdot \partial_\varphi - a_0(\varphi) \end{pmatrix} \begin{pmatrix} \widehat{u} \\ \widehat{\psi} \end{pmatrix}$$

where a_1, a_0 are real valued Sobolev functions in $H^s(\mathbb{T}^\nu, \mathbb{R})$, with $a_1 - 1$, $a_0 = O(\varepsilon)$ and $\mathcal{R}^{(2)}$ is an arbitrarily regularizing operator of the form (2.101). In Sect. 4.2, we introduce the complex variables $h = \frac{1}{\sqrt{2}}(\widehat{u} + i\widehat{\psi})$ and the operator $\mathcal{L}_1$ transforms into $\mathcal{L}_2$ defined in (4.25) which has the form

$$\begin{pmatrix} h \\ \overline{h} \end{pmatrix} \mapsto \begin{pmatrix} \left(\omega \cdot \partial_\varphi + ia_1(\varphi)|D| + i\mathcal{R}^{(2)}\right)h + \left(a_0(\varphi) + i\mathcal{R}^{(2)}\right)\overline{h} \\ \text{complex conjugate} \end{pmatrix},$$

with $\mathcal{R}^{(2)} := \frac{1}{2}\mathcal{R}^{(1)}$.

2. Change of variables. In Sect. 4.3, we reduce to constant coefficients the highest order term $ia_1(\varphi)|D|$ in the operator $\mathcal{L}_2$. Note that it depends only on time. This is due to the special structure of the equation, since the nonlinear term is *diagonal* in space. To reduce to constant coefficients $ia_1(\varphi)|D|$, we conjugate $\mathcal{L}_2$ by means of the reparametrization of time $\mathcal{A}h(\varphi, x) := h(\varphi + \omega\alpha(\varphi), x)$ induced by the diffeomorphism of the torus $\mathbb{T}^\nu$, $\varphi \mapsto \varphi + \omega\alpha(\varphi)$. Since ω is diophantine, choosing $\alpha(\varphi)$ as in (4.33), the transformed operator $\mathcal{L}_3$ defined in (4.34) is

$$\begin{pmatrix} h \\ \overline{h} \end{pmatrix} \mapsto \begin{pmatrix} \left(\omega \cdot \partial_\vartheta + im|D| + i\mathcal{R}^{(3)}\right)h + \left(b_0 + i\mathcal{R}^{(3)}\right)\overline{h} \\ \text{complex conjugate} \end{pmatrix}$$

where $m \in \mathbb{R}$ is a constant $m \simeq 1$, $b_0 = O(\varepsilon)$ is a real valued Sobolev function in $H^s(\mathbb{T}^\nu, \mathbb{R})$ and $\mathcal{R}^{(4)}$ is a one-smoothing operator still satisfying the estimates

(4.44). Actually $\mathcal{R}^{(4)}$ is arbitrarily smoothing, since it has the form (2.101), but we only need that it is one-smoothing.

3. Descent method. In Sect. 4.4, we perfom one step of descent method, in order to remove the zero-th order term from the operator $\mathcal{L}_3$. Since the operator $\mathcal{R}^{(3)}$ is already one-smoothing, we need just to remove the multiplication operator $\overline{h} \mapsto b_0(\varphi)\overline{h}$. For this purpose we transform $\mathcal{L}_3$ by means of the symplectic transformations $\mathcal{V} = \exp(iV(\varphi)|D|^{-1})$, $V(\varphi) = \begin{pmatrix} 0 & v(\varphi) \\ -v(\varphi) & 0 \end{pmatrix}$ where $v : \mathbb{T}^\nu \to \mathbb{R}$ is a real valued Sobolev function. Choosing v as in (4.51), the transformed operator $\mathcal{L}_4$ in (4.52) is the sum of a diagonal operator and a 1-smoothing operator $\mathcal{R}_4$, such that $\mathcal{R}_4|D|$ has finite block-decay norm.

2×2-block diagonal reducibility scheme. Once (1.14) has been obtained, we perform a quadratic KAM reducibility scheme which conjugates the operator $\mathcal{L}_4$ to the 2×2 block diagonal operator $\mathcal{L}_\infty$ (see Theorems 5.1, 5.2). The reason for which we cannot completely diagonalize the operator $\mathcal{L}_4$ is the following: since we deal with periodic boundary conditions, the eigenvalues of the operator $m|D|$ are double, therefore the second order Melnikov conditions for the differences $m|j| - m| \pm j|$ are violated. This implies that after the first step of the KAM iteration, the correction to the diagonal part $im|D|$ is an operator of the form $i\widehat{D}$, $\widehat{D} = \mathrm{diag}_{j \in \mathbb{N}}\widehat{\mathbf{D}}_j$, where $\widehat{\mathbf{D}}_j$ is a linear self-adjoint operator $\mathrm{span}\{e^{ijx}, e^{-ijx}\} \to \mathrm{span}\{e^{ijx}, e^{-ijx}\}$ which we identify with the 2×2 self-adjoint matrix of its Fourier coefficients with respect to the basis $\{e^{ijx}, e^{-ijx}\}$. The self-adjointness of the 2×2 blocks is provided by the Hamiltonian structure. In order to deal with these 2×2 block diagonal operators, it is convenient to introduce a 2×2 block representation for linear operators. We develop this formalism in Sect. 2.3. We remark that the problem of the double multiplicity of the eigenvalues has been overcome for the first time by Chierchia–You [22], for analytic semilinear Klein–Gordon equation with periodic boundary condition. We also mention that the 2×2-block diagonal reducibility scheme, adopted in Sect. 5, has been recently developed by Feola [27] for quasi-linear Hamiltonian NLS equation.

One of the main task in the KAM reducibility scheme is to provide, along the iterative scheme, an asymptotic expansion of the perturbed 2×2 blocks of form

$$\begin{pmatrix} m|j| & 0 \\ 0 & m|j| \end{pmatrix} + O(\varepsilon|j|^{-1}). \tag{1.15}$$

This expansion allows to show that the required second order Melnikov non-resonance conditions are fullfilled for a large set of frequencies ω. The asymptotic (1.15) is achieved since the initial remainder $\mathcal{R}_0$ is 1-smoothing and this property is preserved along the reducibility scheme (see (5.17) in Theorem 5.1). This is the reason why we performed the regularization procedure of Sect. 4 up to order $O(|D|^{-1})$. We use the block-decay norm $|\cdot|_s$ (see (2.80)) to estimate the size of the remainders along the iteration. This is convenient since the class

of operators having finite block-decay norm is closed under composition (Lemma 2.7), solution of the homological equation (Lemma 5.1) and projections (Lemma 2.11).

Linear stability. A final comment concerns Theorem 1.2 which is proved in Sect. 9.1. Using the splitting (3.1), (3.2), the linearized equation (1.11) is decoupled into the two systems (9.2), (9.3). The system (9.2) is a constant coefficients ODE which can be solved explicitly, hence it is enough to study the stability for the PDE (9.3), which is obtained by (1.11), restricting the vector field to the zero average functions in x. All the transformations we perform along the reduction procedure of Sects. 4 and 5 are Töplitz in time operators (see Sect. 2.1), hence they can be regarded as time dependent quasi-periodic maps acting on the phase space (functions of x only). Hence, by the procedure of Sects. 4 and 5, the linear equation (9.3), transforms into the PDE (9.4), whose vector field is a time independent 2×2 block-diagonal operator. Thanks to the Hamiltonian structure, such a vector field is skew self-adjoint, implying that all the Sobolev norms of the solutions remain constant for all time. This is enough to deduce the linear stability.

To conclude the introduction, we make some comments concerning possible extensions and generalizations of the results obtained in this paper.

1. It would be interesting to study the existence of quasi-periodic solutions for equations of the form

$$\partial_{tt}u - \partial_{xx}u + mu + \varepsilon f(\omega t, x, u, u_x, u_{xx}) = 0, \quad m \geq 0 \qquad (1.16)$$

   where

$$f : \mathbb{T}^v \times \mathbb{T} \times \mathbb{R}^3 \to \mathbb{R}$$

   is a q-times differentiable function. The analysis of the linearized operator associated to this equation, would be much more complicated, since the highest order term depends also on the space variable x (whereas in our case, it depends only on time). Indeed, the linearized equation associated to (1.16), written as a first order system in *complex coordinates*, will have the form

$$\partial_t h + \mathrm{i}\,(1 + \varepsilon a(\omega t, x))\,(\sqrt{-\partial_{xx} + m})h + \cdots,$$

   where ... stand for terms of order smaller or equal than 0. The reduction to constant coefficients of the term $(1 + \varepsilon a(\omega t, x))\,(\sqrt{-\partial_{xx} + m})$ involves *Pseudo differential operators* and *Fourier integral operators* much more complicated than the ones used in the present paper.

2. Another possible generalization concerns the extension of our results in higher space dimension. It is well known that the reducibility of quasi-periodically forced wave-type equations is a difficult matter in dimension greater or equal than two. The main reason is that, if

$$\mu_j \sim |j| = \sqrt{j_1^2 + \cdots + j_d^2}, \ j = (j_1, \ldots, j_d) \in \mathbb{Z}^d \text{ with } d \geq 2, \text{ the}$$
   second order Melnikov conditions

$$|\omega \cdot \ell + \mu_j - \mu_{j'}| \geq \frac{\gamma}{\langle \ell \rangle^\tau} \qquad (1.17)$$

for any $(\ell, j, j') \in \mathbb{Z}^\nu \times \mathbb{Z}^d \times \mathbb{Z}^d$ with $(\ell, |j|, |j'|) \neq (0, |j|, |j|)$ are violated. Hence, our proof does not cover the higher dimensional case, since, along the KAM-reducibility scheme, we impose second order Melnikov conditions of the form (1.17).

Actually one can perform, even in the higher dimensional setting, the steps 1, 2, 3 of pages 8 and 9, due to the special structure of the equation. Then, in order to prove the invertibility of the linearized operator, one could try to use a *multiscale approach* (developed in [12] for semilinear Klein Gordon equation) in which it is required to impose only the first order Melnikov non resonance conditions. This would be the natural generalization to the quasi-periodic case of [4]. In any case this approach does not provide the linear stability of the quasi-periodic solutions.

2. Functional setting

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

$$u(\varphi, x) = \sum_{j \in \mathbb{Z}} u_j(\varphi) e^{ijx} = \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{Z}}} \widehat{u}_j(\ell) e^{i(\ell \cdot \varphi + jx)}, \qquad (2.1)$$

where

$$u_j(\varphi) := \frac{1}{2\pi} \int_{\mathbb{T}} u(\varphi, x) e^{-ijx} \, dx,$$

$$\widehat{u}_j(\ell) := \frac{1}{(2\pi)^{\nu+1}} \int_{\mathbb{T}^{\nu+1}} u(\varphi, x) e^{-i(\ell \cdot \varphi + jx)} \, d\varphi \, dx.$$

We also consider the space of the L^2 real valued functions that we denote by $L^2(\mathbb{T}^{\nu+1}, \mathbb{R})$, $L^2(\mathbb{T}_x, \mathbb{R})$. We define for any $s \geq 0$ the Sobolev spaces $H^s(\mathbb{T}^{\nu+1}, \mathbb{C})$, $H^s(\mathbb{T}_x, \mathbb{C})$ as

$$H^s(\mathbb{T}^{\nu+1}, \mathbb{C}) := \left\{ u \in L^2(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{C}) : \|u\|_s^2 := \sum_{(\ell, j) \in \mathbb{Z}^\nu \times \mathbb{Z}} \langle \ell, j \rangle^{2s} |\widehat{u}_j(\ell)|^2 < +\infty \right\},$$

$$H^s(\mathbb{T}_x, \mathbb{C}) := \left\{ u \in L^2(\mathbb{T}_x, \mathbb{C}) : \|u\|_{H^s_x}^2 := \sum_{j \in \mathbb{Z}} \langle j \rangle^{2s} |\widehat{u}_j|^2 < +\infty \right\}$$

where $\langle \ell, j \rangle := \max\{1, |\ell|, |j|\}$, $\langle j \rangle := \max\{1, |j|\}$ and for any $\ell \in \mathbb{Z}^\nu$, $|\ell| := \max_{i=1,\dots,\nu} |\ell_i|$. In a similar way, we define the Sobolev spaces of real values functions $H^s(\mathbb{T}^{\nu+1}, \mathbb{R})$, $H^s(\mathbb{T}_x, \mathbb{R})$. When no confusion appears, we simply write $L^2(\mathbb{T}^{\nu+1})$, $L^2(\mathbb{T}_x)$, $H^s(\mathbb{T}^{\nu+1})$, $H^s(\mathbb{T}_x)$. For any $s \geq 0$ we also define

$$H_0^s(\mathbb{T}^{\nu+1}) := \left\{ u \in H^s(\mathbb{T}^{\nu+1}) : \int_{\mathbb{T}} u(\varphi, x) \, dx = 0 \right\}, \qquad (2.2)$$

$$H_0^s(\mathbb{T}_x) := \left\{ u \in H^s(\mathbb{T}_x) : \int_{\mathbb{T}} u(x) \, dx = 0 \right\}. \qquad (2.3)$$

and $L_0^2(\mathbb{T}^{\nu+1}) = H_0^0(\mathbb{T}^{\nu+1})$, $L_0^2(\mathbb{T}_x) = H_0^0(\mathbb{T}_x)$. We define the spaces $H_0^s(\mathbb{T}^{\nu+1}, \mathbb{C}^2) := H_0^s(\mathbb{T}^{\nu+1}, \mathbb{C}) \times H_0^s(\mathbb{T}^{\nu+1}, \mathbb{C})$ and $H_0^s(\mathbb{T}_x, \mathbb{C}^2) := H_0^s(\mathbb{T}_x, \mathbb{C}) \times H_0^s(\mathbb{T}_x, \mathbb{C})$ equipped, respectively, by the norms $\|(h, v)\|_s := \max\{\|h\|_s, \|v\|_s\}$ and $\|(h, v)\|_{H_x^s} := \max\{\|h\|_{H_x^s}, \|v\|_{H_x^s}\}$. Similarly we define $H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2) := H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}) \times H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R})$ and $H_0^s(\mathbb{T}_x, \mathbb{R}^2) := H_0^s(\mathbb{T}_x, \mathbb{R}) \times H_0^s(\mathbb{T}_x, \mathbb{R})$ and the norms are defined as in the complex case.

For a function $f : \Omega_o \to E$, $\omega \mapsto f(\omega)$, where $(E, \|\cdot\|_E)$ is a Banach space and Ω_o is a subset of $\mathbb{R}^\nu$, we define the sup-norm and the lipschitz semi-norm as

$$\|f\|_{E,\Omega_o}^{\sup} := \sup_{\omega \in \Omega_o} \|f(\omega)\|_E, \quad \|f\|_{E,\Omega_o}^{\lip} := \sup_{\substack{\omega_1, \omega_2 \in \Omega_o \\ \omega_1 \neq \omega_2}} \frac{\|f(\omega_1) - f(\omega_2)\|_E}{|\omega_1 - \omega_2|} \quad (2.4)$$

and, for $\gamma > 0$, we define the weighted Lipschitz-norm

$$\|f\|_{E,\Omega_o}^{\mathrm{Lip}(\gamma)} := \|f\|_{E,\Omega_o}^{\sup} + \gamma \|f\|_{E,\Omega_o}^{\lip}. \quad (2.5)$$

To shorten the above notations we simply omit to write Ω_o, namely $\|f\|_E^{\sup} = \|f\|_{E,\Omega_o}^{\sup}$, $\|f\|_E^{\lip} = \|f\|_{E,\Omega_o}^{\lip}$, $\|f\|_E^{\mathrm{Lip}(\gamma)} = \|f\|_{E,\Omega_o}^{\mathrm{Lip}(\gamma)}$ If $f : \Omega_o \to \mathbb{C}$, we simply denote $\|f\|_{\mathbb{C}}^{\mathrm{Lip}(\gamma)}$ by $|f|^{\mathrm{Lip}(\gamma)}$ and if $E = H^s(\mathbb{T}^{\nu+1})$ we simply denote $\|f\|_{H^s}^{\mathrm{Lip}(\gamma)} := \|f\|_s^{\mathrm{Lip}(\gamma)}$. Given two Banach spaces E, F, we denote by $\mathcal{L}(E, F)$ the space of the bounded linear operators $E \to F$. If $E = F$, we simply write $\mathcal{L}(E)$.

Notation: The notation $a \leq_s b$ means that there exists a constant $C(s) > 0$ depending on s such that $a \leq C(s)b$. The constant $C(s)$ may depend also on the data of the problem, namely the number of frequencies ν, the diophantine exponent $\tau > 0$ appearing in the non-resonance conditions, the forcing term f. If the constant C does not depend on s or if $s = s_0 = [(\nu+1)/2] + 1$, we simply write $a < b$.

We recall the classical estimates for the operator $(\omega \cdot \partial_\varphi)^{-1}$ defined as

$$(\omega \cdot \partial_\varphi)^{-1}[1] = 0, \quad (\omega \cdot \partial_\varphi)^{-1}[e^{i\ell \cdot \varphi}] = \frac{1}{i(\omega \cdot \ell)} e^{i\ell \cdot \varphi}, \qquad \forall \ell \neq 0, \quad (2.6)$$

for $\omega \in \Omega_{\gamma,\tau}$, where for $\gamma, \tau > 0$,

$$\Omega_{\gamma,\tau} := \left\{ \omega \in \Omega : |\omega \cdot \ell| \geq \frac{\gamma}{|\ell|^\tau}, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\} \right\}. \quad (2.7)$$

If $h(\cdot; \omega) \in H^{s+2\tau+1}(\mathbb{T}^{\nu+1})$, with $\omega \in \Omega_{\gamma,\tau}$, we have

$$\|(\omega \cdot \partial_\varphi)^{-1} h\|_s \leq \gamma^{-1} \|h\|_{s+\tau}, \qquad \|(\omega \cdot \partial_\varphi)^{-1} h\|_s^{\mathrm{Lip}(\gamma)} \leq \gamma^{-1} \|h\|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)}. \quad (2.8)$$

Denote by $\mathbb{N}$, the set of the strictly positive integer numbers $\mathbb{N} := \{1, 2, 3, \ldots\}$ and we set $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$. Given a function $h \in L_0^2(\mathbb{T}^{\nu+1})$, we can write

$$h(\varphi, x) = \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{Z} \setminus \{0\}}} \widehat{h}_j(\ell) e^{i(\ell \cdot \varphi + jx)} = \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \widehat{\mathbf{h}}_j(\ell, x) e^{i\ell \cdot \varphi}, \quad (2.9)$$

where

$$\widehat{\mathbf{h}}_j(\ell, x) := \widehat{h}_j(\ell) e^{ijx} + \widehat{h}_{-j}(\ell) e^{-ijx}, \qquad \forall j \in \mathbb{N}. \quad (2.10)$$

It is straightforward to see that if $h \in H_0^s(\mathbb{T}^{\nu+1})$, one has

$$\|h\|_s^2 = \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \langle \ell, j \rangle^{2s} \|\widehat{\mathbf{h}}_j(\ell)\|_{L^2}^2. \tag{2.11}$$

We now recall the following classical interpolation result.

Lemma 2.1. *Let $u, v \in H^s(\mathbb{T}^{\nu+1})$ with $s \geq s_0$. Then, there exists an increasing function $s \mapsto C(s)$ such that*

$$\|uv\|_s \leq C(s)\|u\|_s \|v\|_{s_0} + C(s_0)\|u\|_{s_0}\|v\|_s.$$

If $u(\cdot; \omega)$, $v(\cdot; \omega)$, $\omega \in \Omega_o \subseteq \mathbb{R}^\nu$ are ω-dependent families of functions in $H^s(\mathbb{T}^{\nu+1})$, with $s \geq s_0$ then the same estimate holds replacing $\|\cdot\|_s$ by $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$.

Iterating the above inequality one gets that, for some constant $K(s)$, for any $n \geq 0$,

$$\|u^k\|_s \leq K(s)^k \|u\|_{s_0}^{k-1}\|u\|_s \tag{2.12}$$

and if $u(\cdot; \omega) \in H^s$, $s \geq s_0$ is a family of Sobolev functions, the same inequality holds repacing $\|\cdot\|_s$ by $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$.

We also recall the classical Lemmas on the composition operators. Since the variables (φ, x) have the same role, we present it for a generic Sobolev space $H^s(\mathbb{T}^n)$. For any $s \geq 0$ integer, for any domain $A \subseteq \mathbb{R}^n$ we denote by $\mathcal{C}^s(A)$ the space of the s-times continuously differentiable functions equipped by the usual $\|\cdot\|_{\mathcal{C}^s}$ norm. We consider the composition operator

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

The following Lemma holds:

Lemma 2.2. (Composition operator) *Let $f \in \mathcal{C}^{s+1}(\mathbb{T}^n \times \mathbb{R}, \mathbb{R})$, with $s \geq s_0 := [n/2]+1$. If $u \in H^s(\mathbb{T}^n)$, with $\|u\|_{s_0} \leq 1$, then $\|\mathtt{f}(u)\|_s \leq C(s, \|f\|_{\mathcal{C}^s})(1+\|u\|_s)$. If $u(\cdot, \omega) \in H^s(\mathbb{T}^n)$, $\omega \in \Omega_o \subseteq \mathbb{R}^\nu$ is a family of Sobolev functions satisfying $\|u\|_{s_0}^{\mathrm{Lip}(\gamma)} \leq 1$, then, $\|\mathtt{f}(u)\|_s^{\mathrm{Lip}(\gamma)} \leq C(s, \|f\|_{\mathcal{C}^{s+1}})(1+\|u\|_s^{\mathrm{Lip}(\gamma)})$.*

Now we state the tame properties of the composition operator $u(y) \mapsto u(y + p(y))$ induced by a diffeomorphism of the torus $\mathbb{T}^n$. The Lemma below, can be proved as Lemma 2.20 in [15].

Lemma 2.3. (Change of variable) *Let $p := p(\cdot; \omega) : \mathbb{R}^n \to \mathbb{R}^n$, $\omega \in \Omega_o \subset \mathbb{R}^\nu$ be a family of 2π-periodic functions satisfying*

$$\|p\|_{\mathcal{C}^{s_0+1}} \leq 1/2, \quad \|p\|_{s_0}^{\mathrm{Lip}(\gamma)} \leq 1 \tag{2.13}$$

where $s_0 := [n/2] + 1$. Let $g(y) := y + p(y)$, $y \in \mathbb{T}^n$. Then the composition operator

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

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

$$\|Au\|_{s_0} \lesssim_{s_0} \|u\|_{s_0}, \qquad \|Au\|_s \leq C(s)\|u\|_s + C(s_0)\|p\|_s\|u\|_{s_0+1}. \tag{2.14}$$

Moreover, for any family of Sobolev functions $u(\cdot; \omega)$

$$\|Au\|_{s_0}^{\mathrm{Lip}(\gamma)} \leq_{s_0} \|u\|_{s_0+1}^{\mathrm{Lip}(\gamma)}, \tag{2.15}$$

$$\|Au\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|u\|_{s+1}^{\mathrm{Lip}(\gamma)} + \|p\|_s^{\mathrm{Lip}(\gamma)}\|u\|_{s_0+2}^{\mathrm{Lip}(\gamma)}, \quad \forall s > s_0. \tag{2.16}$$

The map g is invertible with inverse $g^{-1}(z) = z + q(z)$ and there exists a constant $\delta := \delta(s_0) \in (0, 1)$ such that, if $\|p\|_{2s_0+2}^{\mathrm{Lip}(\gamma)} \leq \delta$, then

$$\|q\|_s \leq_s \|p\|_s, \qquad \|q\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|p\|_{s+1}^{\mathrm{Lip}(\gamma)}. \tag{2.17}$$

Furthermore, the composition operator $A^{-1}u(z) := u(z + q(z))$ satisfies the estimate

$$\|A^{-1}u\|_s \leq_s \|u\|_s + \|p\|_s\|u\|_{s_0+1}, \quad \forall s \geq s_0 \tag{2.18}$$

and for any family of Sobolev functions $u(\cdot; \omega)$

$$\|A^{-1}u\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|u\|_{s+1}^{\mathrm{Lip}(\gamma)} + \|p\|_{s+1}^{\mathrm{Lip}(\gamma)}\|u\|_{s_0+2}^{\mathrm{Lip}(\gamma)}, \quad \forall s \geq s_0. \tag{2.19}$$

2.1. Töplitz in time linear operators

Let $\mathcal{R} : \mathbb{T}^\nu \mapsto \mathcal{L}(L_0^2(\mathbb{T}_x))$, $\varphi \mapsto \mathcal{R}(\varphi)$, be a φ-dependent family of linear operators acting on $L_0^2(\mathbb{T}_x)$. We regard $\mathcal{R}$ also as an operator (that for simplicity we denote by $\mathcal{R}$ as well) which acts on functions $u \in L_0^2(\mathbb{T}^\nu \times \mathbb{T})$ of space-time, i.e. we consider the operator $\mathcal{R} \in \mathcal{L}(L_0^2(\mathbb{T}^\nu \times \mathbb{T}))$ defined by

$$(\mathcal{R}u)(\varphi, x) := (\mathcal{R}(\varphi)u(\varphi, \cdot))(x).$$

The action of this operator on a function $u \in L_0^2(\mathbb{T}^{\nu+1})$ is given by

$$
\begin{aligned}
\mathcal{R}u(\varphi, x) &= \sum_{\substack{j,j' \in \mathbb{Z}\backslash\{0\}}} \mathcal{R}_j^{j'}(\varphi)u_{j'}(\varphi)e^{\mathrm{i}jx} \\
&= \sum_{\substack{\ell,\ell' \in \mathbb{Z}^\nu \\ j,j' \in \mathbb{Z}\backslash\{0\}}} \mathcal{R}_j^{j'}(\ell - \ell')\widehat{u}_{j'}(\ell')e^{\mathrm{i}(\ell\cdot\varphi + jx)}
\end{aligned} \tag{2.20}
$$

where the space Fourier coefficients $\mathcal{R}_j^{j'}(\varphi)$ and the space-time Fourier coefficients $\mathcal{R}_j^{j'}(\ell)$ of the operator $\mathcal{R}$ are defined as

$$\mathcal{R}_j^{j'}(\varphi) := \frac{1}{2\pi} \int_{\mathbb{T}} \mathcal{R}(\varphi)[e^{\mathrm{i}j'x}]e^{-\mathrm{i}jx}\,dx, \quad \varphi \in \mathbb{T}^\nu, \ j,j' \in \mathbb{Z}\backslash\{0\}, \tag{2.21}$$

$$\mathcal{R}_j^{j'}(\ell) := \frac{1}{(2\pi)^\nu} \int_{\mathbb{T}^\nu} \mathcal{R}_j^{j'}(\varphi)e^{-\mathrm{i}\ell\cdot\varphi}\,d\varphi, \quad \ell \in \mathbb{Z}^\nu, \ j,j' \in \mathbb{Z}\backslash\{0\}. \tag{2.22}$$

We shall identify the operator $\mathcal{R} = \mathcal{R}(\varphi)$ with the infinite-dimensional matrices of its Fourier coefficiens

$$\left(\mathcal{R}_j^{j'}(\varphi)\right)_{j,j' \in \mathbb{Z}\backslash\{0\}}, \quad \left(\mathcal{R}_j^{j'}(\ell - \ell')\right)_{\substack{\ell,\ell' \in \mathbb{Z}^\nu \\ j,j' \in \mathbb{Z}\backslash\{0\}}} \tag{2.23}$$

and we refer to such operators as Töplitz in time operators.

If the map $\varphi \in \mathbb{T}^\nu \mapsto \mathcal{R}(\varphi) \in \mathcal{L}(L_0^2(\mathbb{T}_x))$ is differentiable, given $\omega \in \mathbb{R}^\nu$, we can define the operator $\omega \cdot \partial_\varphi \mathcal{R}$ as

$$\omega \cdot \partial_\varphi \mathcal{R} = \left(\omega \cdot \partial_\varphi \mathcal{R}_j^{j'}(\varphi) \right)_{j,j' \in \mathbb{Z} \setminus \{0\}} = \left(i\omega \cdot (\ell - \ell') \mathcal{R}_j^{j'}(\ell - \ell') \right)_{\substack{\ell, \ell' \in \mathbb{Z}^\nu \\ j, j' \in \mathbb{Z} \setminus \{0\}}} . \tag{2.24}$$

We also define the *commutator* between two Töplitz in time operators $\mathcal{R} = \mathcal{R}(\varphi)$ and $\mathcal{B} = \mathcal{B}(\varphi)$ as $[\mathcal{R}(\varphi), \mathcal{B}(\varphi)] := \mathcal{R}(\varphi)\mathcal{B}(\varphi) - \mathcal{B}(\varphi)\mathcal{R}(\varphi)$, $\varphi \in \mathbb{T}^\nu$.

Given a Töplitz in time operator $\mathcal{R}$, we define the conjugated operator $\overline{\mathcal{R}}$ by

$$\overline{\mathcal{R}} u := \overline{\mathcal{R}\bar{u}}. \tag{2.25}$$

One gets easily that the operator $\overline{\mathcal{R}}$ has the matrix representation

$$\left(\overline{\mathcal{R}_{-j}^{-j'}(\varphi)} \right)_{j,j' \in \mathbb{Z} \setminus \{0\}}, \qquad \varphi \in \mathbb{T}^\nu. \tag{2.26}$$

An operator $\mathcal{R}$ is said to be real if it maps real-valued functions on real valued functions and it is easy to see that $\mathcal{R}$ is real if and only if $\mathcal{R} = \overline{\mathcal{R}}$.

We define also the transpose operator $\mathcal{R}^T = \mathcal{R}(\varphi)^T$ by the relation

$$\langle \mathcal{R}(\varphi)[u], v \rangle_{L_x^2} = \langle u, \mathcal{R}(\varphi)^T[v] \rangle_{L_x^2}, \qquad \forall u, v \in L_0^2(\mathbb{T}_x), \quad \forall \varphi \in \mathbb{T}^\nu \tag{2.27}$$

where

$$\langle u, v \rangle_{L_x^2} := \int_{\mathbb{T}} u(x)v(x), dx, \qquad \forall u, v \in L^2(\mathbb{T}_x). \tag{2.28}$$

Note that the operator $\mathcal{R}^T$ has the matrix representation

$$(\mathcal{R}^T)_j^{j'}(\varphi) = \mathcal{R}_{-j'}^{-j}(\varphi), \quad \forall j, j' \in \mathbb{Z} \setminus \{0\}, \quad \forall \varphi \in \mathbb{T}^\nu. \tag{2.29}$$

An operator $\mathcal{R}$ is said to be symmetric if $\mathcal{R} = \mathcal{R}^T$.

We define also the adjoint operator $\mathcal{R}^* = \mathcal{R}(\varphi)^*$ by

$$(\mathcal{R}(\varphi)[u], v)_{L_x^2} = (u, \mathcal{R}(\varphi)^*[v])_{L_x^2}, \quad \forall u, v \in L_0^2(\mathbb{T}_x), \quad \forall \varphi \in \mathbb{T}^\nu, \tag{2.30}$$

where $(\cdot, \cdot)_{L_x^2}$ is the scalar product on $L^2(\mathbb{T})$, namely

$$(u, v)_{L_x^2} := \langle u, \bar{v} \rangle_{L_x^2} = \int_{\mathbb{T}} u(x)\bar{v}(x) \, dx, \quad \forall u, v \in L^2(\mathbb{T}_x). \tag{2.31}$$

An operator $\mathcal{R}$ is said to be self-adjoint if $\mathcal{R} = \mathcal{R}^*$. It is easy to see that $\mathcal{R}^* = \overline{\mathcal{R}}^T$ and its matrix representation is given by

$$(\mathcal{R}^*)_j^{j'}(\varphi) = \overline{\mathcal{R}_{j'}^j(\varphi)}, \quad \forall j, j' \in \mathbb{Z} \setminus \{0\}, \quad \forall \varphi \in \mathbb{T}^\nu.$$

In the following we also deal with smooth families of real operators $\varphi \mapsto G(\varphi) \in \mathcal{L}(L_0^2(\mathbb{T}_x, \mathbb{R}^2))$, of the form

$$G(\varphi) := \begin{pmatrix} A(\varphi) & B(\varphi) \\ C(\varphi) & D(\varphi) \end{pmatrix}, \qquad \varphi \in \mathbb{T}^\nu \tag{2.32}$$

where $A(\varphi), B(\varphi), C(\varphi), D(\varphi) \in \mathcal{L}(L_0^2(\mathbb{T}_x, \mathbb{R}))$, for all $\varphi \in \mathbb{T}^\nu$. Actually G may be regarded as an operator in $\mathcal{L}(L_0^2(\mathbb{T}^{\nu+1}, \mathbb{R}^2))$, according to the fact that A, B, C, D are Töplitz in time operators. By (2.27), the transpose operator G^T with respect to the bilinear form

$$\langle (u_1, \psi_1), (u_2, \psi_2) \rangle_{L_x^2} := \langle u_1, u_2 \rangle_{L_x^2} + \langle \psi_1, \psi_2 \rangle_{L_x^2}, \tag{2.33}$$

$\forall (u_1, \psi_1), (u_2, \psi_2) \in L_0^2(\mathbb{T}_x, \mathbb{R}^2)$, is given by

$$G^T = \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix}. \tag{2.34}$$

Then it is easy to verify that G is symmetric, i.e. $G = G^T$ if and only if $A = A^T$, $B = C^T$, $D = D^T$. It is also convenient to regard the real operator G in the complex variables

$$z := \frac{u + \mathrm{i}\psi}{\sqrt{2}}, \qquad \overline{z} = \frac{u - \mathrm{i}\psi}{\sqrt{2}}, \tag{2.35}$$

$$u = \frac{z + \overline{z}}{\sqrt{2}}, \qquad \psi = \frac{z - \overline{z}}{\mathrm{i}\sqrt{2}}. \tag{2.36}$$

The transformed operator $\mathcal{R}$ has the form

$$\mathcal{R} = \begin{pmatrix} R_1 & R_2 \\ R_2 & R_1 \end{pmatrix}, \tag{2.37}$$

$$R_1 := \frac{A + D - \mathrm{i}(B - C)}{2}, \quad R_2 := \frac{A - D + \mathrm{i}(B + C)}{2}.$$

Note that the operator $\mathcal{R}$ satisfies

$$\mathcal{R} : \mathbf{L}_0^2(\mathbb{T}^{\nu+1}) \to \mathbf{L}_0^2(\mathbb{T}^{\nu+1}), \quad \mathcal{R}(\varphi) : \mathbf{L}_0^2(\mathbb{T}_x) \to \mathbf{L}_0^2(\mathbb{T}_x), \quad \forall \varphi \in \mathbb{T}^\nu \tag{2.38}$$

where $\mathbf{L}_0^2(\mathbb{T}^{\nu+1})$, resp. $\mathbf{L}_0^2(\mathbb{T}_x)$ are the real subspaces of $L_0^2(\mathbb{T}^{\nu+1}, \mathbb{C}^2)$, resp. $L_0^2(\mathbb{T}_x, \mathbb{C}^2)$ defined as

$$\mathbf{L}_0^2(\mathbb{T}^{\nu+1}) := \left\{ (z, \overline{z}) : z \in L_0^2(\mathbb{T}^{\nu+1}, \mathbb{C}) \right\}, \tag{2.39}$$

$$\mathbf{L}_0^2(\mathbb{T}_x) := \left\{ (z, \overline{z}) : z \in L_0^2(\mathbb{T}_x, \mathbb{C}) \right\}. \tag{2.40}$$

For the sequel, we also introduce for any $s \geq 0$, the real subspaces of $H_0^s(\mathbb{T}^{\nu+1}, \mathbb{C}^2)$ and $H_0^s(\mathbb{T}_x, \mathbb{C}^2)$

$$\mathbf{H}_0^s(\mathbb{T}^{\nu+1}) := H_0^s(\mathbb{T}^{\nu+1}, \mathbb{C}^2) \cap \mathbf{L}_0^2(\mathbb{T}^{\nu+1}), \tag{2.41}$$

$$\mathbf{H}_0^s(\mathbb{T}_x) := H_0^s(\mathbb{T}_x, \mathbb{C}^2) \cap \mathbf{L}_0^2(\mathbb{T}_x). \tag{2.42}$$

2.2. Hamiltonian formalism

We define the symplectic form $\mathcal{W}$ as

$$\mathcal{W}[\mathbf{u}_1, \mathbf{u}_2] := \langle \mathbf{u}_1, J\mathbf{u}_2 \rangle_{L_x^2}, \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{2.43}$$

for all $\mathbf{u_1}, \mathbf{u_2} \in L_0^2(\mathbb{T}_x, \mathbb{R}^2)$.

Definition 2.1. A φ-dependent linear vector field $X(\varphi) : L_0^2(\mathbb{T}_x, \mathbb{R}^2) \to L_0^2(\mathbb{T}_x, \mathbb{R}^2)$ is Hamiltonian, if $X(\varphi) = JG(\varphi)$, where J is given in (2.43) and the operator G is symmetric. The operator

$$\mathcal{L} = \omega \cdot \partial_\varphi \mathbb{I}_2 - JG(\varphi) : H_0^1(\mathbb{T}^{\nu+1}, \mathbb{R}^2) \to L_0^2(\mathbb{T}^{\nu+1}, \mathbb{R}^2), \quad \mathbb{I}_2 := \begin{pmatrix} \mathrm{Id}_0 & 0 \\ 0 & \mathrm{Id}_0 \end{pmatrix}$$

where $\mathrm{Id}_0 : L_0^2(\mathbb{T}^{\nu+1}) \to L_0^2(\mathbb{T}^{\nu+1})$ is the identity, is called Hamiltonian operator.

Definition 2.2. A φ-dependent map $\Phi(\varphi) : L_0^2(\mathbb{T}_x, \mathbb{R}^2) \to L_0^2(\mathbb{T}_x, \mathbb{R}^2)$ is symplectic if for any $\varphi \in \mathbb{T}^\nu$, for any $\mathbf{u}_1, \mathbf{u}_2 \in L_0^2(\mathbb{T}_x, \mathbb{R}^2)$,

$$\mathcal{W}[\Phi(\varphi)\mathbf{u}_1, \, \Phi(\varphi)\mathbf{u}_2] = \mathcal{W}[\mathbf{u}_1, \mathbf{u}_2],$$

or equivalently $\Phi(\varphi)^T J \Phi(\varphi) = J$, for all $\varphi \in \mathbb{T}^\nu$.

Under a symplectic transformation $\Phi = \Phi(\varphi)$, assuming that the map $\varphi \in \mathbb{T}^\nu \mapsto \Phi(\varphi) \in \mathcal{L}(L_0^2(\mathbb{T}_x, \mathbb{R}^2))$ is differentiable, a linear Hamiltonian operator $\mathcal{L} = \omega \cdot \partial_\varphi \mathbb{I}_2 - JG(\varphi)$ transforms into the operator $\mathcal{L}_+ = \Phi^{-1}\mathcal{L}\Phi = \omega \cdot \partial_\varphi \mathbb{I}_2 - JG_+(\varphi)$ with

$$G_+(\varphi) := \Phi(\varphi)^T G(\varphi)\Phi(\varphi) + \Phi(\varphi)^T J\omega \cdot \partial_\varphi \Phi(\varphi). \tag{2.44}$$

Note that for all $\varphi \in \mathbb{T}^\nu$, $G_+(\varphi)$ is symmetric, because $G(\varphi)$ is symmetric and $\omega \cdot \partial_\varphi [\Phi(\varphi)^T] J\Phi(\varphi) + \Phi(\varphi)^T J\omega \cdot \partial_\varphi \Phi(\varphi) = 0$ for all $\varphi \in \mathbb{T}^\nu$ and then $\mathcal{L}_+$ is still a Hamiltonian operator. Actually the conjugation (2.44) can be interpreted also from a dynamical point of view. Indeed, consider the quasi-periodically forced linear Hamiltonian PDE

$$\partial_t \mathbf{h} = JG(\omega t)\mathbf{h}, \quad t \in \mathbb{R}, \quad \omega \in \mathbb{R}^\nu. \tag{2.45}$$

Under the change of coordinates $\mathbf{h} = \Phi(\omega t)\mathbf{v}$, the above PDE is transformed into the equation

$$\partial_t \mathbf{v} = JG_+(\omega t)\mathbf{v} \tag{2.46}$$

which is still a linear Hamiltonian PDE.

2.2.1. Hamiltonian formalism in complex coordinates. In this section we explain how the real Hamiltonian structure described above, reads in the complex coordinates introduced in (2.35), (2.36). According to (2.37), under the change of coordinates (2.35), (2.36), a linear Hamiltonian vector field $JG(\varphi)$, transforms into

$$\mathcal{R}(\varphi) = -\mathrm{i}\begin{pmatrix} R_1(\varphi) & R_2(\varphi) \\ -\overline{R_2(\varphi)} & -\overline{R_1(\varphi)} \end{pmatrix}, \tag{2.47}$$

where the operators $R_i = R_i(\varphi)$, $i = 1, 2$ are defined as

$$R_1 := \frac{A + D - \mathrm{i}B + \mathrm{i}B^T}{2}, \quad R_2 := \frac{A - D + \mathrm{i}B + \mathrm{i}B^T}{2} \tag{2.48}$$

(recall that the operator $\overline{R}$ is defined in (2.25)). Note that the operators $R_1(\varphi)$, $R_2(\varphi)$ are linear operators acting on complex valued L^2 functions $L_0^2(\mathbb{T}_x, \mathbb{C})$, moreover since $G(\varphi)$ is symmetric, $A(\varphi) = A(\varphi)^T$, $B(\varphi) = C(\varphi)^T$, $D(\varphi) = D(\varphi)^T$, then it turns out that

$$R_1(\varphi) = R_1(\varphi)^*, \qquad R_2(\varphi) = R_2(\varphi)^T, \qquad \forall \varphi \in \mathbb{T}^\nu. \tag{2.49}$$

Since the operator $\mathcal{R}$ in (2.4) has the form (2.37), it satisfies the property (2.38). Furthermore, one has that $\mathcal{R}(\varphi)$ is the linear Hamiltonian vector field associated to the real Hamiltonian

$$\mathcal{H}(\mathbf{z}) := \langle \mathbf{G}(\varphi)[\mathbf{z}], \, \mathbf{z} \rangle_{L_x^2}, \quad \mathbf{G}(\varphi) := \begin{pmatrix} \overline{R_2(\varphi)} & \overline{R_1(\varphi)} \\ R_1(\varphi) & R_2(\varphi) \end{pmatrix}, \tag{2.50}$$

namely

$$\mathcal{H}(z, \overline{z}) = \int_{\mathbb{T}} R_1(\varphi)[z]\overline{z}\, dx + \frac{1}{2}\int_{\mathbb{T}} R_2(\varphi)[z]z\, dx + \frac{1}{2}\int_{\mathbb{T}} \overline{R_2(\varphi)}[\overline{z}]\,\overline{z}\, dx. \quad (2.51)$$

Indeed, $\mathbf{G}(\varphi)$ is symmetric, since by (2.49), $\overline{R_1}^T = R_1^* = R_1$ and $R_1^T = \overline{R}_1$, then

$$\mathcal{R}(\varphi)[\mathbf{z}] = -\mathrm{i}J\nabla_{\mathbf{z}}\mathcal{H}(\mathbf{z}) = -\mathrm{i}J\mathbf{G}(\varphi)[\mathbf{z}], \quad \mathbf{z} \in \mathbf{L}_0^2(\mathbb{T}_x), \quad (2.52)$$

where $\nabla_{\mathbf{z}}\mathcal{H} := (\nabla_z\mathcal{H}, \nabla_{\overline{z}}\mathcal{H})$ with

$$\nabla_z\mathcal{H} = \frac{1}{\sqrt{2}}(\nabla_\eta\mathcal{H} - \mathrm{i}\nabla_\psi\mathcal{H}), \quad \nabla_{\overline{z}}\mathcal{H} := \overline{\nabla_z\mathcal{H}} = \frac{1}{\sqrt{2}}(\nabla_\eta\mathcal{H} + \mathrm{i}\nabla_\psi\mathcal{H})$$

(recall (2.35), (2.40)). The symplectic form $\mathcal{W}$ in (2.43), reads in the complex coordinates (2.35) as

$$\Gamma[\mathbf{z}_1, \mathbf{z}_2] = \mathrm{i}\int_{\mathbb{T}}(z_1\overline{z}_2 - \overline{z}_1 z_2)\, dx = \mathrm{i}\langle\mathbf{z}_1,\, J\mathbf{z}_2\rangle_{L_x^2}, \quad \forall\mathbf{z}_1, \mathbf{z}_2 \in \mathbf{L}_0^2(\mathbb{T}_x). \quad (2.53)$$

Definition 2.3. Let $\Phi_i = \Phi_i(\varphi)$, $\varphi \in \mathbb{T}^\nu$, $i = 1, 2$ be φ-dependent families of linear operators $L_0^2(\mathbb{T}_x, \mathbb{C}) \to L_0^2(\mathbb{T}_x, \mathbb{C})$. We say that the map

$$\Phi(\varphi) = \begin{pmatrix} \Phi_1(\varphi) & \Phi_2(\varphi) \\ \overline{\Phi_2(\varphi)} & \overline{\Phi_1(\varphi)} \end{pmatrix}, \qquad \varphi \in \mathbb{T}^\nu$$

is symplectic if

$$\Gamma[\Phi(\varphi)[\mathbf{z}_1], \Phi(\varphi)[\mathbf{z}_2]] = \Gamma[\mathbf{z}_1, \mathbf{z}_2], \quad \forall\mathbf{z}_1, \mathbf{z}_2 \in \mathbf{L}_0^2(\mathbb{T}_x), \quad \forall\varphi \in \mathbb{T}^\nu$$

or equivalently $\Phi(\varphi)^T J\Phi(\varphi) = J$, for all $\varphi \in \mathbb{T}^\nu$.

It is well known that if $\mathcal{R}(\varphi)$ is an operator of the form (2.4), (2.49), namely by (2.52) $\mathcal{R}(\varphi)$ is a linear Hamiltonian vector field associated to the quadratic Hamiltonian $\mathcal{H}$ in (2.51), the operators $\exp(\pm\mathcal{R}(\varphi))$ are symplectic maps.

Definition 2.4. If $\mathcal{R}(\varphi)$ is a Hamiltonian vector field like in (2.4), (2.49), we define the Hamiltonian operator in complex coordinates as

$$\mathcal{L} = \omega\cdot\partial_\varphi\mathbb{I}_2 - \mathcal{R}(\varphi) = \omega\cdot\partial_\varphi\mathbb{I}_2 + \mathrm{i}J\mathbf{G}(\varphi) : \mathbf{H}_0^1(\mathbb{T}^{\nu+1}) \to \mathbf{L}_0^2(\mathbb{T}^{\nu+1}).$$

Under the action of a smooth family of symplectic map $\Phi(\varphi)$, $\varphi \in \mathbb{T}^\nu$, a Hamiltonian operator $\mathcal{L}$ transforms into the Hamiltonian operator $\mathcal{L}_+ = \Phi^{-1}\mathcal{L}\Phi = \omega\cdot\partial_\varphi\mathbb{I}_2 + \mathrm{i}J\mathbf{G}_+(\varphi)$ where

$$\mathbf{G}_+(\varphi) := \Phi(\varphi)^T\mathbf{G}(\varphi)\Phi(\varphi) + \Phi(\varphi)^T J\omega\cdot\partial_\varphi\Phi(\varphi), \quad \forall\varphi \in \mathbb{T}^\nu.$$

Note that the operator $\mathbf{G}_+(\varphi)$ is symmetric and it has the same form as $\mathbf{G}(\varphi)$ in (2.50). Arguing as in (2.45), (2.46), under the transformation $\mathbf{v} = \Phi(\omega t)\mathbf{h}$, the PDE

$$\partial_t\mathbf{h} = -\mathrm{i}J\mathbf{G}(\omega t)\mathbf{h}, \qquad \omega \in \mathbb{R}^\nu, \quad t \in \mathbb{R}, \quad (2.54)$$

transforms into the PDE

$$\partial_t\mathbf{h} = -\mathrm{i}J\mathbf{G}_+(\omega t)\mathbf{h}. \quad (2.55)$$

In the following, we will consider also quasi-periodic reparametrizations of time, namely operators of the form

$$\mathcal{A}\mathbf{h}(\varphi, x) = \mathbf{h}(\varphi + \omega\alpha(\varphi), x),$$

where $\alpha : \mathbb{T}^\nu \to \mathbb{R}$ is a sufficiently smooth function and such that $\|\alpha\|_{\mathcal{C}^1}$ is sufficiently small. The transformation $\mathcal{A}$ is invertible and its inverse $\mathcal{A}^{-1}$ has the form

$$\mathcal{A}^{-1}\mathbf{h}(\vartheta, x) = \mathbf{h}(\vartheta + \omega\tilde{\alpha}(\vartheta), x)$$

where $\vartheta \mapsto \vartheta + \omega\tilde{\alpha}(\vartheta)$ is the inverse diffeomorphism of $\varphi \mapsto \varphi + \omega\alpha(\varphi)$. The conjugated operator is $\mathcal{A}^{-1}\mathcal{L}\mathcal{A} = \rho\mathcal{L}_+$, where $\mathcal{L}_+ = \omega \cdot \partial_\varphi + \mathrm{i}J\mathbf{G}_+(\vartheta)$ with

$$\rho(\vartheta) := \mathcal{A}^{-1}[1 + \omega \cdot \partial_\varphi\alpha](\vartheta), \quad \mathbf{G}_+(\vartheta) := \frac{1}{\rho(\vartheta)}\mathbf{G}(\vartheta + \omega\tilde{\alpha}(\vartheta)). \tag{2.56}$$

Note that $\mathcal{L}_+$ is still a Hamiltonian operator. From a dynamical point of view, under the reparametrization of time

$$\tau = t + \alpha(\omega t), \qquad t = \tau + \tilde{\alpha}(\omega\tau),$$

setting $\mathbf{v}(t) := A(\omega t)\mathbf{h} := \mathbf{h}(t + \alpha(\omega t), x)$, the PDE (2.54) is transformed into

$$\partial_\tau\mathbf{v} = -\mathrm{i}J\mathbf{G}_+(\omega\tau)\mathbf{v}. \tag{2.57}$$

2.3. 2×2 block representation of linear operators

We may regard a Töplitz in time operator given by (2.20) as a 2×2 block matrix

$$\left(\mathbf{R}_j^{j'}(\ell - \ell')\right)_{\substack{\ell, \ell' \in \mathbb{Z}^\nu \\ j, j' \in \mathbb{N}}}, \tag{2.58}$$

where for all $\ell \in \mathbb{Z}^\nu$, $j, j' \in \mathbb{N}$ the 2×2 matrix $\mathbf{R}_j^{j'}(\ell)$ is defined by

$$\mathbf{R}_j^{j'}(\ell) := \begin{pmatrix} \mathcal{R}_j^{j'}(\ell) & \mathcal{R}_j^{-j'}(\ell) \\ \mathcal{R}_{-j}^{j'}(\ell) & \mathcal{R}_{-j}^{-j'}(\ell) \end{pmatrix}. \tag{2.59}$$

The 2×2 matrix $\mathbf{R}_j^{j'}(\ell)$ can be regarded as a linear operator in $\mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$, where for all $j \in \mathbb{N}$, the two dimensional space $\mathbf{E}_j$ is defined as

$$\mathbf{E}_j := \mathrm{span}\{e^{\mathrm{i}jx}, e^{-\mathrm{i}jx}\}. \tag{2.60}$$

Note that for any $j \in \mathbb{N}$, the finite dimensional space $\mathbf{E}_j$ is the eigenspace of the operator $-\partial_{xx}$ corresponding to the eigenvalue j^2. We identify the space $\mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ of the linear operators from $\mathbf{E}_{j'}$ onto $\mathbf{E}_j$ with the space of the 2×2 matrices of their Fourier coefficients, namely

$$\mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j) \simeq \left\{ M = \left(M_k^{k'}\right)_{\substack{k=\pm j \\ k'=\pm j'}} \right\} \simeq \mathrm{Mat}(2 \times 2). \tag{2.61}$$

Indeed if $M \in \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$, its action is given by

$$Mu(x) = \sum_{\substack{k=\pm j \\ k'=\pm j'}} M_k^{k'} u_{k'} e^{\mathrm{i}kx}, \quad \forall u \in \mathbf{E}_{j'}, \quad u(x) = u_{j'}e^{\mathrm{i}j'x} + u_{-j'}e^{-\mathrm{i}j'x}.$$

$$\tag{2.62}$$

If $j = j'$, we use the notation $\mathcal{L}(\mathbf{E}_j) = \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ and we denote by $\mathbf{I}_j$ the identity operator on the space $\mathbf{E}_j$, namely

$$\mathbf{I}_j : \mathbf{E}_j \to \mathbf{E}_j, \qquad u \mapsto u. \tag{2.63}$$

According to (2.9), (2.58), (2.62), we may write the action of a Töplitz in time operator on a function $h(\varphi, x)$ as

$$\mathcal{R}h(\varphi, x) = \sum_{\substack{\ell, \ell' \in \mathbb{Z}^\nu \\ j, j' \in \mathbb{N}}} \mathbf{R}_j^{j'}(\ell - \ell')[\widehat{\mathbf{h}}_{j'}(\ell')]e^{\mathrm{i}\ell \cdot \varphi}. \tag{2.64}$$

We denote by $[\mathcal{R}]$ the 2×2 block-diagonal part of the operator $\mathcal{R}$, namely

$$[\mathcal{R}] := \mathrm{diag}_{j \in \mathbb{N}} \mathbf{R}_j^j(0) \tag{2.65}$$

and its action on a function $h(\varphi, x)$ is given by

$$[\mathcal{R}]h(\varphi, x) = \sum_{\ell \in \mathbb{Z}^\nu, \, j \in \mathbb{N}} \mathbf{R}_j^j(0)[\widehat{\mathbf{h}}_j(\ell)]e^{\mathrm{i}\ell \cdot \varphi}.$$

If $\mathbf{R}_j^{j'}(\ell) = 0$, for any $(\ell, j, j') \neq (0, j, j)$, we have $\mathcal{R} = [\mathcal{R}]$ and we refer to such operators as 2×2 block-diagonal operators.

For any $M \in \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$, we define the transpose operator $M^T \in \mathcal{L}(\mathbf{E}_j, \mathbf{E}_{j'})$ by

$$(M^T)_k^{k'} := M_{-k'}^{-k}, \quad k = \pm j', \quad k' = \pm j, \tag{2.66}$$

the conjugate operator $\overline{M} \in \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ by

$$(\overline{M})_k^{k'} := \overline{M_{-k}^{-k'}}, \quad k = \pm j, \quad k' = \pm j', \tag{2.67}$$

the adjoint operator $M^* \in \mathcal{L}(\mathbf{E}_j, \mathbf{E}_{j'})$ as

$$M^* := \overline{M}^T. \tag{2.68}$$

Given an operator $A \in \mathcal{L}(\mathbf{E}_j)$, we define its trace as

$$\mathrm{Tr}(A) := A_j^j + A_{-j}^{-j}. \tag{2.69}$$

It is easy to check that if $A, B \in \mathcal{L}(\mathbf{E}_j)$, then

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA). \tag{2.70}$$

For all $j, j' \in \mathbb{N}$, the space $\mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ defined in (2.61), is a Hilbert space equipped by the inner product given for any $X, Y \in \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ by

$$\langle X, Y \rangle := \mathrm{Tr}(XY^*). \tag{2.71}$$

This scalar product induces the L^2-norm

$$\|X\| := \sqrt{\mathrm{Tr}(XX^*)} = \left(\sum_{\substack{|k|=j \\ |k'|=j'}} |X_k^{k'}|^2 \right)^{\frac{1}{2}}. \tag{2.72}$$

Actually all the norms on the finite dimensional space $\mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ are equivalent.

Given a linear operator $\mathbf{L} : \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j) \to \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$, we denote by $\|\mathbf{L}\|_{\mathrm{Op}(j,j')}$ its operatorial norm, when the space $\mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ is equipped by the L^2-norm (2.72), namely

$$\|\mathbf{L}\|_{\mathrm{Op}(j,j')} := \sup\{\|\mathbf{L}(M)\| : M \in \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j), \quad \|M\| \leq 1\}. \qquad (2.73)$$

We denote by $\mathbf{I}_{j,j'}$ the identity operator on $\mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$, namely

$$\mathbf{I}_{j,j'} : \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j) \to \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j), \qquad X \mapsto X. \qquad (2.74)$$

For any operator $A \in \mathcal{L}(\mathbf{E}_j)$ we denote by $M_L(A) : \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j) \to \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ the linear operator defined for any $X \in \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ as

$$M_L(A)X := AX. \qquad (2.75)$$

Similarly, given an operator $B \in \mathcal{L}(\mathbf{E}_{j'})$, we denote by $M_R(B) : \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j) \to \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ the linear operator defined for any $X \in \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ as

$$M_R(B)X := XB. \qquad (2.76)$$

The following elementary estimates hold:

$$\|M_L(A)\|_{\mathrm{Op}(j,j')} \leq \|A\|, \quad \|M_R(B)\|_{\mathrm{Op}(j,j')} \leq \|B\|. \qquad (2.77)$$

For any $j \in \mathbb{N}$, we denote by $\mathcal{S}(\mathbf{E}_j)$, the set of the self-adjoint operators form $\mathbf{E}_j$ onto itself, namely

$$\mathcal{S}(\mathbf{E}_j) := \{A \in \mathcal{L}(\mathbf{E}_j) : A = A^*\}, \qquad (2.78)$$

which we identify with the set of the 2×2 self-adjoint matrices. Furthermore, for any $A \in \mathcal{L}(\mathbf{E}_j)$ denote by $\mathrm{spec}(A)$ the spectrum of A. The following Lemma can be proved by using elementary arguments from linear algebra, hence the proof is omitted.

Lemma 2.4. *Let $A \in \mathcal{S}(\mathbf{E}_j)$, $B \in \mathcal{S}(\mathbf{E}_{j'})$, then the following holds:*

(i)   *The operators $M_L(A)$, $M_R(B)$ defined in (2.75), (2.76) are self-adjoint operators with respect to the scalar product defined in (2.71).*

(ii)   *The spectrum of the operator $M_L(A) \pm M_R(B)$ satisfies*

$$\mathrm{spec}\,(M_L(A) \pm M_R(B)) = \{\lambda \pm \mu : \lambda \in \mathrm{spec}(A), \quad \mu \in \mathrm{spec}(B)\}.$$

We finish this section by recalling some well known facts concerning linear self-adjoint operators on finite dimensional Hilbert spaces. Let $\mathcal{H}$ be a finite dimensional Hilbert space of dimension n equipped by the inner product $(\cdot, \cdot)_{\mathcal{H}}$. For any self-adjoint operator $A : \mathcal{H} \to \mathcal{H}$, we order its eigenvalues as

$$\mathrm{spec}(A) := \{\lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_n(A)\}. \qquad (2.79)$$

Lemma 2.5. *Let $\mathcal{H}$ be a Hilbert space of dimension n. Then the following holds:*

(i)   *Let $A_1, A_2 : \mathcal{H} \to \mathcal{H}$ be self-adjoint operators. Then their eigenvalues, ranked as in (2.79), satisfy the Lipschitz property*

$$|\lambda_k(A_1) - \lambda_k(A_2)| \leq \|A_1 - A_2\|_{\mathcal{L}(\mathcal{H})}, \qquad \forall k = 1, \ldots, n.$$

(ii)   *Let $A = \eta \mathrm{Id}_{\mathcal{H}} + B$, where $\eta \in \mathbb{R}$, $\mathrm{Id}_{\mathcal{H}} : \mathcal{H} \to \mathcal{H}$ is the identity and $B : \mathcal{H} \to \mathcal{H}$ is selfadjoint. Then*

$$\lambda_k(A) = \eta + \lambda_k(B), \qquad \forall k = 1, \ldots, n.$$

(iii)   *Let $A : \mathcal{H} \to \mathcal{H}$ be self-adjoint and assume that $\mathrm{spec}(A) \subset \mathbb{R} \backslash \{0\}$. Then A is invertible and its inverse satisfies*

$$\|A^{-1}\|_{\mathcal{L}(\mathcal{H})} = \frac{1}{\min_{k=1,\dots,n} |\lambda_k(A)|}.$$

2.4. Block-decay norm for linear operators

In this Section, we introduce the block-decay norm for linear operators. Given a Töplitz in time operator $\mathcal{R}$ as in (2.20), recalling its 2×2 block representation (2.58), (2.59), we define its block-decay norm as

$$|\mathcal{R}|_s := \sup_{j' \in \mathbb{N}} \left(\sum_{\ell \in \mathbb{Z}^\nu, \, j \in \mathbb{N}} \langle \ell, j - j' \rangle^{2s} \|\mathbf{R}_j^{j'}(\ell)\|^2 \right)^{\frac{1}{2}}, \qquad (2.80)$$

where $\| \cdot \|$ is defined in (2.72). For a family of Töplitz in time operators $\mathcal{R} = \mathcal{R}(\omega) \in \mathcal{L}\left(H_0^s(\mathbb{T}^{\nu+1})\right)$, $\omega \in \Omega_o$, given $\gamma > 0$, we define the norm

$$|\mathcal{R}|_s^{\mathrm{Lip}(\gamma)} := |\mathcal{R}|_s^{\sup} + \gamma |\mathcal{R}|_s^{\mathrm{lip}}, \qquad (2.81)$$

$$|\mathcal{R}|_s^{\sup} := \sup_{\omega \in \Omega_o} |\mathcal{R}(\omega)|_s, \quad |\mathcal{R}|_s^{\mathrm{lip}} := \sup_{\substack{\omega_1, \omega_2 \in \Omega_o \\ \omega_1 \neq \omega_2}} \frac{|\mathcal{R}(\omega_1) - \mathcal{R}(\omega_2)|_s}{|\omega_1 - \omega_2|}.$$

For families of linear operators $\mathcal{R}(\omega)$, $\omega \in \Omega_o$ of the form

$$\mathcal{R} = \begin{pmatrix} \mathcal{R}_1 & \mathcal{R}_2 \\ \overline{\mathcal{R}_2} & \overline{\mathcal{R}_1} \end{pmatrix}, \qquad (2.82)$$

where $\mathcal{R}_1, \mathcal{R}_2 \in \mathcal{L}\left(H_0^s(\mathbb{T}^{\nu+1})\right)$ are Töplitz in time operators of the form (2.20), we define

$$|\mathcal{R}|_s := \max\{|\mathcal{R}_1|_s, |\mathcal{R}_2|_s\}, \; |\mathcal{R}|_s^{\mathrm{Lip}(\gamma)} := \max\{|\mathcal{R}_1|_s^{\mathrm{Lip}(\gamma)}, |\mathcal{R}_2|_s^{\mathrm{Lip}(\gamma)}\}. \quad (2.83)$$

In the following, we state some properties of this norm. We prove such properties for operators $\mathcal{R} \in \mathcal{L}\left(H_0^s(\mathbb{T}^{\nu+1})\right)$. If $\mathcal{R}$ is an operator of the form (2.82) then the same statements hold with the obvious modifications. To state the following lemma we need the following definition. For all $m \in \mathbb{R}$ we define the operator $|D|^m$ as

$$|D|^m(1) = 0, \qquad |D|^m(e^{\mathrm{i}jx}) = |j|^m e^{\mathrm{i}jx} \quad \forall j \neq 0. \qquad (2.84)$$

Lemma 2.6. (i)   *The norm $|\cdot|_s$ is increasing, namely $|\mathcal{R}|_s \leq |\mathcal{R}|_{s'}$, for $s \leq s'$.*
(ii) $|\mathcal{R}|_s \leq |\mathcal{R}|D\||_s$ *and the operator $\omega \cdot \partial_\varphi \mathcal{R}$ (see (2.24)) satisfies $|\omega \cdot \partial_\varphi \mathcal{R}|_s \leq |\mathcal{R}|_{s+1}$.*
(iii)   *For any $j \in \mathbb{N}$, the 2×2 block $\mathbf{R}_j^j(0)$ (see (2.59)) satisfies $\sup_{j \in \mathbb{N}} \|\mathbf{R}_j^j(0)\| \lessdot |\mathcal{R}|_0$, where $\|\cdot\|$ is defined in (2.72). Moreover the operator $[\mathcal{R}]$ defined by (2.65), satisfies $|[\mathcal{R}]|_s \leq |\mathcal{R}|_s$.*
(iv)   *Items (i)–(iii) hold, replacing $|\cdot|_s$ by $|\cdot|_s^{\mathrm{Lip}(\gamma)}$ and $\|\cdot\|$ by $\|\cdot\|^{\mathrm{Lip}(\gamma)}$.*

Proof. The proof is elementary. It follows directly by the definitions (2.80), (2.81), hence we omit it. $\qquad \square$

Lemma 2.7. *Let $\mathcal{R}$, $\mathcal{B}$ be operators of the form* (2.20)*. Then*

$$|\mathcal{RB}|_s \leq_s |\mathcal{R}|_s |\mathcal{B}|_{s_0} + |\mathcal{R}|_{s_0} |\mathcal{B}|_s. \tag{2.85}$$

If $\mathcal{R} = \mathcal{R}(\omega)$, $\mathcal{B} = \mathcal{B}(\omega)$ are Lipschitz with respect to the parameter $\omega \in \Omega_o \subseteq \Omega$, then the same estimate holds replacing $|\cdot|_s$ by $|\cdot|_s^{\mathrm{Lip}(\gamma)}$.

Proof. According to the matrix representations (2.58), (2.59), the operator $\mathcal{RB}$ has the 2×2 block representation

$$\mathcal{RB} = \left([\mathbf{RB}]_j^{j'}(\ell - \ell') \right)_{\substack{\ell,\ell' \in \mathbb{Z}^\nu \\ j,j' \in \mathbb{N}}}, \quad [\mathbf{RB}]_j^{j'}(\ell) = \sum_{\substack{\ell' \in \mathbb{Z}^\nu \\ k \in \mathbb{N}}} \mathbf{R}_j^k(\ell - \ell') \mathbf{B}_k^{j'}(\ell').$$

By the Cauchy Schwartz inequality, $\|\mathbf{R}_j^k(\ell - \ell')\mathbf{B}_k^{j'}(\ell')\| \leq \|\mathbf{R}_j^k(\ell - \ell')\| \|\mathbf{B}_k^{j'}(\ell')\|$, then for all $j' \in \mathbb{N}$, we get

$$\sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \langle \ell, j - j' \rangle^{2s} \| [\mathbf{RB}]_j^{j'}(\ell) \|^2 \leq \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \left(\sum_{\substack{\ell' \in \mathbb{Z}^\nu \\ k \in \mathbb{N}}} \langle \ell, j - j' \rangle^s \| \mathbf{R}_j^k(\ell - \ell') \| \| \mathbf{B}_k^{j'}(\ell') \| \right)^2. \tag{2.86}$$

Using that $\langle \ell, j - j' \rangle^s \leq_s \langle \ell - \ell', j - k \rangle^s + \langle \ell', k - j' \rangle^s$, we get (2.86) $\leq_s (A) + (B)$ where

$$(A) := \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \left(\sum_{\substack{\ell' \in \mathbb{Z}^\nu \\ k \in \mathbb{N}}} \langle \ell - \ell', j - k \rangle^s \| \mathbf{R}_j^k(\ell - \ell') \| \| \mathbf{B}_k^{j'}(\ell') \| \right)^2, \tag{2.87}$$

$$(B) := \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \left(\sum_{\substack{\ell' \in \mathbb{Z}^\nu \\ k \in \mathbb{N}}} \langle \ell', k - j' \rangle^s \| \mathbf{R}_j^k(\ell - \ell') \| \| \mathbf{B}_k^{j'}(\ell') \| \right)^2. \tag{2.88}$$

By the Cauchy–Schwartz inequality, using that since $s_0 = [(\nu + 1)/2] + 1 > (\nu + 1)/2$, the series $\sum_{\ell' \in \mathbb{Z}^\nu, k \in \mathbb{N}} \langle \ell', k - j' \rangle^{-2s_0} = C(s_0)$, one has

$$(A) \lesssim \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \sum_{\substack{\ell' \in \mathbb{Z}^\nu \\ k \in \mathbb{N}}} \langle \ell - \ell', j - k \rangle^{2s} \| \mathbf{R}_j^k(\ell - \ell') \|^2 \langle \ell', k - j' \rangle^{2s_0} \| \mathbf{B}_k^{j'}(\ell') \|^2$$

$$\lesssim \sum_{\substack{\ell' \in \mathbb{Z}^\nu \\ k \in \mathbb{N}}} \langle \ell', k - j' \rangle^{2s_0} \| \mathbf{B}_k^{j'}(\ell') \|^2 \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \langle \ell - \ell', j - k \rangle^{2s} \| \mathbf{R}_j^k(\ell - \ell') \|^2$$

$$\lesssim \sup_{j' \in \mathbb{N}} \sum_{\substack{\ell' \in \mathbb{Z}^\nu \\ k \in \mathbb{N}}} \langle \ell', k - j' \rangle^{2s_0} \| \mathbf{B}_k^{j'}(\ell') \|^2 \sup_{k \in \mathbb{N}} \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \langle \ell - \ell', j - k \rangle^{2s} \| \mathbf{R}_j^k(\ell - \ell') \|^2$$

$$\overset{(2.80)}{\lesssim} |\mathcal{R}|_s^2 |\mathcal{B}|_{s_0}^2.$$

By similar arguments, one gets $(B) \lesssim |\mathcal{R}|_{s_0}^2 |\mathcal{B}|_s^2$ and hence the claimed estimate follows by taking the supremum over $j' \in \mathbb{N}$ in (2.86). The estimate in $|\cdot|_s^{\mathrm{Lip}(\gamma)}$, follows by applying the estimate (2.85) to

$$\frac{\mathcal{R}(\omega_1)\mathcal{B}(\omega_1)-\mathcal{R}(\omega_2)\mathcal{B}(\omega_2)}{\omega_1-\omega_2}=\frac{(\mathcal{R}(\omega_1)-\mathcal{R}(\omega_2))\mathcal{B}(\omega_1)}{\omega_1-\omega_2}+\frac{\mathcal{R}(\omega_2)(\mathcal{B}(\omega_1)-\mathcal{B}(\omega_2))}{\omega_1-\omega_2}$$

and passing to the sup for $\omega_1,\omega_2\in\Omega_o$ with $\omega_1\neq\omega_2$. $\qquad\square$

For all $n\geq 1$, iterating the estimate of Lemma 2.7 we get

$$|\mathcal{R}^n|_{s_0}\leq[C(s_0)]^{n-1}|\mathcal{R}|_{s_0}^n,\quad|\mathcal{R}^n|_s\leq nC(s)^n|\mathcal{R}|_{s_0}^{n-1}|\mathcal{R}|_s,\ \forall s\geq s_0,\quad(2.89)$$

for some constant $C(s)>0$, and the same bounds also hold for the norm $|\cdot|_s^{\mathrm{Lip}(\gamma)}$ if $\mathcal{R}=\mathcal{R}(\omega)$ is Lipschitz with respect to the parameter ω.

Lemma 2.8. *Let $\mathcal{R}$ satisfy $|\mathcal{R}|_s<+\infty$, with $s\geq s_0$. Then for all $u\in H_0^s(\mathbb{T}^{\nu+1})$, the following estimate holds*

$$\|\mathcal{R}u\|_s\leq_s|\mathcal{R}|_s\|u\|_{s_0}+|\mathcal{R}|_{s_0}\|u\|_s.$$

If $\mathcal{R}=\mathcal{R}(\omega)$, $u=u(\cdot,\omega)$ are Lipschitz with respect to the parameter $\omega\in\Omega_o\subseteq\mathbb{R}^\nu$, then the same estimate holds replacing $|\cdot|_s$ by $|\cdot|_s^{\mathrm{Lip}(\gamma)}$ and $\|\cdot\|_s$ by $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$.

Proof. The proof is similar to the one of Lemma 2.7, hence it is omitted. $\qquad\square$

Lemma 2.9. *Let $a\in H^s(\mathbb{T}^\nu)$. Then the multiplication operator $\mathcal{R}:h(\varphi,x)\mapsto a(\varphi)h(\varphi,x)$ satisfies $|\mathcal{R}|_s\lessdot\|a\|_s$. If $a=a(\cdot;\omega)$ is a Lipschitz family in $H^s(\mathbb{T}^\nu)$, then the same estimate holds, replacing $\|\cdot\|_s$ by $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$ and $|\cdot|_s$ by $|\cdot|_s^{\mathrm{Lip}(\gamma)}$.*

Proof. The operator $\mathcal{R}$ admits the 2×2-block representation

$$\mathcal{R}=\left(\mathbf{R}_j^j(\ell-\ell')\right)_{\substack{\ell,\ell'\in\mathbb{Z}^\nu\\j\in\mathbb{N}}},\quad\mathbf{R}_j^j(\ell):=\widehat{a}(\ell)\mathbf{I}_j,\quad\forall\ell\in\mathbb{Z}^\nu,\quad\forall j\in\mathbb{N}$$

(recall (2.63)). Since $\|\mathbf{I}_j\|=\sqrt{2}$, by (2.80), one has $|\mathcal{R}|_s\lessdot\|a\|_s$. The estimate for $|\mathcal{R}|_s^{\mathrm{Lip}(\gamma)}$ follows similarly. $\qquad\square$

Lemma 2.10. *Let $\Phi=\exp(\Psi)$ with $\Psi:=\Psi(\omega)$, depending in a Lipschitz way on the parameter $\omega\in\Omega_o\subseteq\mathbb{R}^\nu$, such that $|\Psi|D|\|_{s_0}^{\mathrm{Lip}(\gamma)}\leq 1$, $|\Psi|D|\|_s^{\mathrm{Lip}(\gamma)}<+\infty$, with $s\geq s_0$. Then*

$$|(\Phi^{\pm 1}-\mathrm{Id})|D|\|_s\leq_s|\Psi|D|\|_s,\quad|(\Phi^{\pm 1}-\mathrm{Id})|D|\|_s^{\mathrm{Lip}(\gamma)}\leq_s|\Psi|D|\|_s^{\mathrm{Lip}(\gamma)}.\quad(2.90)$$

The differential $\partial_\Psi\Phi$ of the map $\Psi\mapsto\Phi^{\pm 1}=\exp(\pm\Psi)$ in the direction $\widehat{\Psi}$ satisfies for any $|\Psi|_{s_0}\leq 1$ the estimate

$$|\partial_\Psi\Phi^{\pm 1}[\widehat{\Psi}]|_s\leq_s\left(|\widehat{\Psi}|_s+|\Psi|_s|\widehat{\Psi}|_{s_0}\right).\quad(2.91)$$

Moreover the map $\Phi_{\geq 2}=\Phi-\mathrm{Id}-\Psi$, satisfies

$$|\Phi_{\geq 2}|D|\|_s\leq_s|\Psi|D|\|_s|\Psi|D|\|_{s_0},\quad(2.92)$$

$$|\Phi_{\geq 2}|D|\|_s^{\mathrm{Lip}(\gamma)}\leq_s|\Psi|D|\|_s^{\mathrm{Lip}(\gamma)}|\Psi|D|\|_{s_0}^{\mathrm{Lip}(\gamma)},\quad(2.93)$$

$$|\partial_\Psi\Phi_{\geq 2}[\widehat{\Psi}]|_s\leq_s\left(|\Psi|_{s_0}|\widehat{\Psi}|_s+|\Psi|_s|\widehat{\Psi}|_{s_0}\right).\quad(2.94)$$

Proof. Let us prove the estimate (2.90) for Φ. We write

$$\Phi - \mathrm{Id} = \Psi + \sum_{k \geq 2} \frac{\Psi^k}{k!}.$$

For any $k \geq 2$ one has

$$
\begin{aligned}
|\Psi^k|D||_s & \overset{Lemma\,2.7}{\leq} C(s) \left(|\Psi^{k-1}|_s |\Psi|D||_{s_0} + |\Psi^{k-1}|_{s_0} |\Psi|D||_s \right) \\
& \overset{(2.89)}{\leq_s} (k-1)C(s)^k \left(|\Psi|_s |\Psi|_{s_0}^{k-2} |\Psi|D||_{s_0} + |\Psi|_{s_0}^{k-1} |\Psi|D||_s \right) \\
& \overset{Lemma\,2.6-(ii)}{\leq} 2(k-1)C(s)^k |\Psi|D||_s |\Psi|D||_{s_0}^{k-1} \\
& \overset{|\Psi|D||_{s_0} \leq 1}{\leq_s} 2(k-1)C(s)^k |\Psi|D||_s.
\end{aligned}
\tag{2.95}
$$

Hence

$$|(\Phi - \mathrm{Id})|D||_s \overset{(2.95)}{\leq} |\Psi|D||_s \left(1 + 2 \sum_{k \geq 2} \frac{(k-1)C(s)^k}{k!} \right) \leq_s |\Psi|D||_s.$$

The same inequatity holds for the inverse $\Phi^{-1} = \exp(-\Psi)$.

Now let us prove the estimate (2.91). For any $k \geq 1$, one has that

$$\partial_\Psi (\Psi^k)[\widehat{\Psi}] = \sum_{i+j=k-1} \Psi^i \widehat{\Psi} \Psi^j.$$

For any $i + j = k - 1$

$$
\begin{aligned}
&|\Psi^i \widehat{\Psi} \Psi^j|_s \\
& \overset{Lemma\,2.7}{\leq} C(s)^2 \left(|\Psi^i|_s |\widehat{\Psi}|_{s_0} |\Psi^j|_{s_0} + |\Psi^i|_{s_0} |\widehat{\Psi}|_s |\Psi^j|_{s_0} + |\Psi^i|_{s_0} |\widehat{\Psi}|_{s_0} |\Psi^j|_s \right) \\
& \overset{(2.89)}{\leq} 2kC(s)^{k+1} \left(|\widehat{\Psi}|_s + |\Psi|_s |\widehat{\Psi}|_{s_0} \right).
\end{aligned}
\tag{2.96}
$$

Hence

$$
|\partial_\Psi \Phi[\widehat{\Psi}]|_s \leq \sum_{k \geq 1} \frac{|\partial_\Psi (\Psi^k)[\widehat{\Psi}]|_s}{k!} \overset{(2.96)}{\leq} \sum_{k \geq 1} \frac{2kC(s)^{k+1}}{k!} \left(|\widehat{\Psi}|_s + |\Psi|_s |\widehat{\Psi}|_{s_0} \right)
$$

$$
\leq_s |\widehat{\Psi}|_s + |\Psi|_s |\widehat{\Psi}|_{s_0},
\tag{2.97}
$$

which is the estimate (2.91). The estimates (2.92)–(2.94), can be proved arguing as above, using that $\Phi_{\geq 2} = \sum_{k \geq 2} \frac{\Psi^k}{k!}$. $\qquad \square$

Given $N \in \mathbb{N}$, we define the smoothing operator $\Pi_N \mathcal{R}$, for any operator $\mathcal{R}$ as in (2.20)

$$
(\Pi_N \mathcal{R})_j^{j'} (\ell - \ell') := \begin{cases} \mathcal{R}_j^{j'} (\ell - \ell') & \text{if } |\ell - \ell'| \leq N \\ 0 & \text{otherwise,} \end{cases}
\tag{2.98}
$$

or equivalently, using the block-matrix representation (2.58), (2.59)

$$(\Pi_N \mathbf{R})_j^{j'}(\ell - \ell') := \begin{cases} \mathbf{R}_j^{j'}(\ell - \ell') & \text{if } |\ell - \ell'| \leq N \\ 0 & \text{otherwise.} \end{cases} \tag{2.99}$$

Lemma 2.11. *The operator* $\Pi_N^\perp := \text{Id} - \Pi_N$ *satisfies*

$$|\Pi_N^\perp \mathcal{R}|_s \leq N^{-\mathtt{b}} |\mathcal{R}|_{s+\mathtt{b}}, \quad |\Pi_N^\perp \mathcal{R}|_s^{\text{Lip}(\gamma)} \leq N^{-\mathtt{b}} |\mathcal{R}|_{s+\mathtt{b}}^{\text{Lip}(\gamma)}, \quad \mathtt{b} \geq 0, \tag{2.100}$$

where in the second inequality $\mathcal{R}$ *is Lipschitz with respect to the parameter* $\omega \in \Omega_o \subseteq \mathbb{R}^\nu$.

Proof. The proof follows easily by the definitions (2.80), (2.81) and hence it is omitted. $\qquad\square$

Lemma 2.12. *Let us define the operator*

$$\mathcal{R}h(\varphi, x) := q(\varphi, x) \int_{\mathbb{T}} g(\varphi, x) h(\varphi, x)\, dx, \quad q, g \in H_0^s(\mathbb{T}^{\nu+1}), \quad s \geq s_0. \tag{2.101}$$

Then

$$|\mathcal{R}|_s \leq_s \|g\|_{s_0} \|q\|_s + \|g\|_s \|q\|_{s_0}. \tag{2.102}$$

Moreover if the functions g *and* q *are Lipschitz with respect to the parameter* $\omega \in \Omega_o \subseteq \mathbb{R}^\nu$, *then the same estimate holds replacing* $|\cdot|_s$ *by* $|\cdot|_s^{\text{Lip}(\gamma)}$ *and* $\|\cdot\|_s$ *by* $\|\cdot\|_s^{\text{Lip}(\gamma)}$.

Proof. A direct calculation shows that for all $\ell \in \mathbb{Z}^\nu$ and for all $k, k' \in \mathbb{Z} \setminus \{0\}$

$$\mathcal{R}_k^{k'}(\ell) = \sum_{\ell' \in \mathbb{Z}^\nu} \widehat{q}_k(\ell - \ell') \widehat{g}_{-k'}(\ell').$$

Using definition (2.72) we get that for all $\ell \in \mathbb{Z}^\nu$, $j, j' \in \mathbb{N}$,

$$\|\mathbf{R}_j^{j'}(\ell)\|^2 = \sum_{\substack{k=\pm j \\ k'=\pm j'}} |\mathcal{R}_k^{k'}(\ell)|^2 \leq \left(\sum_{\substack{k=\pm j \\ k'=\pm j'}} |\mathcal{R}_k^{k'}(\ell)| \right)^2$$

$$\leq \left(\sum_{\ell' \in \mathbb{Z}^\nu} \sum_{\substack{k=\pm j \\ k'=\pm j'}} |\widehat{q}_k(\ell - \ell')||\widehat{g}_{-k'}(\ell')| \right)^2.$$

$$\lessdot \left(\sum_{\ell' \in \mathbb{Z}^\nu} \|\widehat{\mathbf{q}}_j(\ell - \ell')\|_{L^2} \|\widehat{\mathbf{g}}_{j'}(\ell')\|_{L^2} \right)^2 \tag{2.103}$$

where the last inequality holds, since, recalling (2.10), for any $k = \pm j$, $k' = \pm j'$, $|\widehat{q}_k(\ell - \ell')| \leq \|\widehat{\mathbf{q}}_j(\ell - \ell')\|_{L^2}$ and $|\widehat{g}_{-k'}(\ell')| \leq \|\widehat{\mathbf{g}}_{j'}(\ell')\|_{L^2}$. Now for all $j' \in \mathbb{N}$,

$$\sum_{\substack{\ell\in\mathbb{Z}^\nu \\ j\in\mathbb{N}}} \langle \ell, j-j'\rangle^{2s} \|\mathbf{R}_j^{j'}(\ell)\|^2$$

$$\overset{(2.103)}{\leq} \sum_{\substack{\ell\in\mathbb{Z}^\nu \\ j\in\mathbb{N}}} \left(\sum_{\ell'\in\mathbb{Z}^\nu} \langle \ell, j-j'\rangle^s \|\widehat{\mathbf{q}}_j(\ell-\ell')\|_{L^2} \|\widehat{\mathbf{g}}_{j'}(\ell')\|_{L^2} \right)^2. \qquad (2.104)$$

Using that $\langle \ell, j-j'\rangle^s \leq_s \langle \ell-\ell', j\rangle^s + \langle \ell', j'\rangle^s$, one gets $(2.104) \leq_s (A) + (B)$, where

$$(A) := \sum_{\substack{\ell\in\mathbb{Z}^\nu \\ j\in\mathbb{N}}} \left(\sum_{\ell'\in\mathbb{Z}^\nu} \langle \ell-\ell', j\rangle^s \|\widehat{\mathbf{q}}_j(\ell-\ell')\|_{L^2} \|\widehat{\mathbf{g}}_{j'}(\ell')\|_{L^2} \right)^2, \qquad (2.105)$$

$$(B) := \sum_{\substack{\ell\in\mathbb{Z}^\nu \\ j\in\mathbb{N}}} \left(\sum_{\ell'\in\mathbb{Z}^\nu} \langle \ell', j'\rangle^s \|\widehat{\mathbf{q}}_j(\ell-\ell')\|_{L^2} \|\widehat{\mathbf{g}}_{j'}(\ell')\|_{L^2} \right)^2. \qquad (2.106)$$

By the Cauchy–Schwartz inequality, using that $\sum_{\ell'\in\mathbb{Z}^\nu} \langle \ell'\rangle^{-2s_0} = C(s_0)$ (recall that $s_0 = [(\nu+1)/2] + 1 > (\nu+1)/2$), one gets

$$(A) \leq_s \sum_{\substack{\ell\in\mathbb{Z}^\nu \\ j\in\mathbb{N}}} \sum_{\ell'\in\mathbb{Z}^\nu} \langle \ell-\ell', j\rangle^{2s} \|\widehat{\mathbf{q}}_j(\ell-\ell')\|_{L^2}^2 \langle \ell'\rangle^{2s_0} \|\widehat{\mathbf{g}}_{j'}(\ell')\|_{L^2}^2 \overset{(2.11)}{\leq_s} \|q\|_s \|g\|_{s_0}.$$

By similar arguments one can prove that $(B) \leq_s \|q\|_{s_0} \|g\|_s$ and the claimed estimate follows by taking the sup over $j' \in \mathbb{N}$ in (2.104). The Lipschitz estimates follow by applying (2.102) to

$$\frac{\mathcal{R}(\omega_1) - \mathcal{R}(\omega_2)}{\omega_1 - \omega_2} = \frac{q(\omega_1) - q(\omega_2)}{\omega_1 - \omega_2} \langle \cdot, g(\omega_1)\rangle_{L_x^2} + g(\omega_2) \left\langle \frac{g(\omega_1) - g(\omega_2)}{\omega_1 - \omega_2}, \cdot \right\rangle_{L_x^2}$$

and passing to the sup for $\omega_1, \omega_2 \in \Omega_o$ with $\omega_1 \neq \omega_2$. $\qquad\square$

As we already mentioned, a Töplitz in time operator $\mathcal{R}$ in (2.20) may be regarded as φ-dependent family acting on the space of functions depending only on the x-variable

$$\mathcal{R}(\varphi) = \left(\mathcal{R}_j^{j'}(\varphi) \right)_{j,j'\in\mathbb{Z}\setminus\{0\}},$$

and it admits the block representation

$$\mathcal{R}(\varphi) = \left(\mathbf{R}_j^{j'}(\varphi) \right)_{j,j'\in\mathbb{N}}, \quad \mathbf{R}_j^{j'}(\varphi)$$

$$= \begin{pmatrix} \mathcal{R}_j^{j'}(\varphi) & \mathcal{R}_j^{-j'}(\varphi) \\ \mathcal{R}_{-j}^{j'}(\varphi) & \mathcal{R}_{-j}^{-j'}(\varphi) \end{pmatrix} \quad \forall \varphi \in \mathbb{T}^\nu, \quad \forall j, j' \in \mathbb{N}.$$

The 2×2 matrix $\mathbf{R}_j^{j'}(\varphi)$ may be regarded as a linear operator in $\mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$, given by

$$\mathbf{R}_j^{j'}(\varphi)[u] = \sum_{\substack{k=\pm j \\ k'=\pm j'}} \mathcal{R}_k^{k'}(\varphi) u_{k'} e^{ikx}, \qquad \forall u(x) = u_{j'} e^{ij'x} + u_{-j'} e^{-ij'x} \in \mathbf{E}_{j'}.$$

For the operator $\mathcal{R}(\varphi)$, we denote by $|\mathcal{R}(\varphi)|_{s,x}$ the block-decay norm (only with respect to the x-variable)

$$|\mathcal{R}(\varphi)|_{s,x} := \sup_{j' \in \mathbb{N}} \left(\sum_{j \in \mathbb{Z}} \langle j - j' \rangle^{2s} \| \mathbf{R}_j^{j'}(\varphi) \|^2 \right)^{\frac{1}{2}}. \tag{2.107}$$

If $\mathcal{R}$ is an operator of the form (2.82), we define

$$|\mathcal{R}(\varphi)|_{s,x} := \max\{|\mathcal{R}_1(\varphi)|_{s,x}, |\mathcal{R}_2(\varphi)|_{s,x}\}. \tag{2.108}$$

The following Lemma holds:

Lemma 2.13. *Let $\mathcal{R}$ be a Töplitz in time operator. Then the following holds:*
(i) Let $s \geq 1$. If for any $\varphi \in \mathbb{T}^\nu$, $|\mathcal{R}(\varphi)|_{s,x} < +\infty$, then for any $u \in H_0^s(\mathbb{T}_x)$

$$\|\mathcal{R}(\varphi)u\|_{H_x^s} \leq_s |\mathcal{R}(\varphi)|_{1,x} \|u\|_{H_x^s} + |\mathcal{R}(\varphi)|_{s,x} \|u\|_{H_x^1}.$$

(ii) $|\mathcal{R}(\varphi)|_{s,x} \lessdot |\mathcal{R}|_{s+s_0}$.

Proof. The proof of item (i) is similar to the one of Lemma 2.8, hence it is omitted. Item (ii) follows since, expanding $\mathbf{R}_j^{j'}(\varphi) = \sum_{\ell \in \mathbb{Z}^\nu} \mathbf{R}_j^{j'}(\ell) e^{\mathrm{i}\ell \cdot \varphi}$, applying the Cauchy–Schwartz inequality and using that $\sum_{\ell \in \mathbb{Z}^\nu} \langle \ell \rangle^{-2s_0} = C(s_0)$, one has that for all $j' \in \mathbb{N}$,

$$\sum_{j \in \mathbb{N}} \langle j - j' \rangle^{2s} \| \mathbf{R}_j^{j'}(\varphi) \|^2 \lessdot \sum_{\substack{\ell \in \mathbb{Z}^\nu \\ j \in \mathbb{N}}} \langle j - j' \rangle^{2s} \langle \ell \rangle^{2s_0} \| \mathbf{R}_j^{j'}(\ell) \|^2 \lessdot |\mathcal{R}|_{s+s_0},$$

which implies the claimed estimate passing to the supremum on $j' \in \mathbb{N}$. $\square$

3. A reduction on the zero mean value functions

For any function $u \in L^2(\mathbb{T})$, we define

$$\pi_0 u := \frac{1}{2\pi} \int_{\mathbb{T}} u(x)\, dx, \qquad \pi_0^\perp := \mathrm{Id} - \pi_0 \tag{3.1}$$

and

$$\Pi_0 := \begin{pmatrix} \pi_0 & 0 \\ 0 & \pi_0 \end{pmatrix}, \qquad \Pi_0^\perp := \begin{pmatrix} \pi_0^\perp & 0 \\ 0 & \pi_0^\perp \end{pmatrix}. \tag{3.2}$$

Given a function $v \in H^s(\mathbb{T}^{\nu+1})$, $v(\varphi, x) = \sum_{j \in \mathbb{Z}} v_j(\varphi) e^{\mathrm{i}jx}$, we write

$$v(\varphi, x) = v_0(\varphi) + v_\perp(\varphi, x),$$
$$v_0(\varphi) := \pi_0 v(\varphi, x), \quad v_\perp(\varphi, x) := \pi_0^\perp v(\varphi, x) = \sum_{j \neq 0} v_j(\varphi) e^{\mathrm{i}jx}. \tag{3.3}$$

Then according to the splitting (3.3), applying the projection Π_0, $\Pi_0^\perp$ to the nonlinear map F defined in (1.6) and setting $u := \pi_0^\perp v$, $\psi := \pi_0^\perp p$, the equation $F(v, p) = F(\varepsilon, \omega, v, p) = 0$ is decomposed in

$$\begin{cases} \omega \cdot \partial_\varphi u - \psi = 0 \\ \omega \cdot \partial_\varphi \psi - \left(1 + \varepsilon \int_{\mathbb{T}} |\partial_x u|^2 \, dx\right) \partial_{xx} u - \varepsilon f_\perp = 0, \end{cases} \tag{3.4}$$

$$\begin{cases} \omega \cdot \partial_\varphi v_0 - p_0 = 0 \\ \omega \cdot \partial_\varphi p_0 - \varepsilon f_0 = 0 \end{cases} \tag{3.5}$$

(we have used that $\partial_x v = \partial_x v_\perp = \partial_x u$ in (3.4)). The above two systems are completely decoupled, hence they can be solved separately. In the next lemma, we solve explicitly the second system (3.5). We use the hypothesis (1.8) on the forcing term $f(\varphi, x)$.

Lemma 3.1. Let $\gamma, \tau > 0$ and $q > 2\tau$. Then, for all $\omega \in \Omega_{\gamma,\tau}$ (see (2.7)), there exists a solution $v_0(\cdot; \omega, \varepsilon), p_0(\cdot; \omega, \varepsilon) \in H^{q-2\tau}(\mathbb{T}^\nu, \mathbb{R})$ of the system (3.5) with $\int_{\mathbb{T}^\nu} p_0(\varphi) \, d\varphi = \int_{\mathbb{T}^\nu} v_0(\varphi) \, d\varphi = 0$ and satisfying the estimates

$$\|v_0\|_s \lessdot \varepsilon \gamma^{-2} \|f\|_{s+2\tau}, \quad \|p_0\|_s \lessdot \varepsilon \gamma^{-1} \|f\|_{s+\tau}, \qquad \forall 0 \le s \le q - 2\tau. \tag{3.6}$$

Proof. Since

$$\int_{\mathbb{T}^\nu} f_0(\varphi) \, d\varphi = \int_{\mathbb{T}^{\nu+1}} f(\varphi, x) \, d\varphi \, dx \overset{(1.8)}{=} 0,$$

the second equation in (3.5) can be solved by taking $p_0 := \varepsilon(\omega \cdot \partial_\varphi)^{-1} f_0$ where, since $\omega \in \Omega_{\gamma,\tau}$, the operator $(\omega \cdot \partial_\varphi)^{-1}$ is well defined by (2.6). Then we can solve the second equation in (3.5) by defining $v_0 := (\omega \cdot \partial_\varphi)^{-1} p_0 = \varepsilon(\omega \cdot \partial_\varphi)^{-2} f_0$. Clearly $\int_{\mathbb{T}^\nu} v_0(\varphi) \, d\varphi = \int_{\mathbb{T}^\nu} p_0(\varphi) \, d\varphi = 0$ and the claimed estimates follow by applying (2.8). □

In all the rest of the paper, we will study the equation (3.4) on the zero mean value functions in x. We will find zeros of the nonlinear operator $\mathcal{F}(\varepsilon, \omega, \cdot) : H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2) \to H_0^{s-2}(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$ (recall (2.2)), defined as

$$\mathcal{F}(\varepsilon, \omega, u, \psi) := \begin{pmatrix} \omega \cdot \partial_\varphi u - \psi \\ \omega \cdot \partial_\varphi \psi - \left(1 + \varepsilon \int_{\mathbb{T}} (\partial_x u)^2 \, dx\right) \partial_{xx} u - \varepsilon f_\perp \end{pmatrix}. \tag{3.7}$$

Note that, setting $\mathbf{u} := (u, \psi)$, $\mathcal{F}(\mathbf{u}) = \mathcal{F}(\varepsilon, \omega, \mathbf{u})$, one has

$$\mathcal{F}(\mathbf{u}) = \omega \cdot \partial_\varphi \mathbf{u} - J\nabla_{\mathbf{u}}\mathcal{H}_\varepsilon(\mathbf{u}), \qquad J := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{3.8}$$

where $J\nabla_{\mathbf{u}}\mathcal{H}_\varepsilon$ is the Hamiltonian vector field

$$J\nabla_{\mathbf{u}}\mathcal{H}_\varepsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \nabla_u \mathcal{H}_\varepsilon \\ \nabla_\psi \mathcal{H}_\varepsilon \end{pmatrix} = \begin{pmatrix} -\nabla_\psi \mathcal{H}_\varepsilon \\ \nabla_u \mathcal{H}_\varepsilon \end{pmatrix}$$

generated by the Hamiltonian

$$\mathcal{H}_\varepsilon(u, \psi) := \frac{1}{2} \int_{\mathbb{T}} \left(\psi^2 + (\partial_x u)^2\right) \, dx + \varepsilon \left(\frac{1}{2} \int_{\mathbb{T}} (\partial_x u)^2 \, dx\right)^2 - \varepsilon \int_{\mathbb{T}} f_\perp u \, dx, \tag{3.9}$$

defined on the phase space $H_0^1(\mathbb{T}_x, \mathbb{R}) \times L_0^2(\mathbb{T}_x, \mathbb{R})$. The Hamiltonian $\mathcal{H}_\varepsilon$ is simply the restriction of the Hamiltonian H in (1.5) to the space of the functions with zero average in x. We look for the zeros of (3.7) by means of an implicit function Theorem of Nash–Moser type. The Theorem 1.1 will be deduced by Lemma 3.1 and by the following Theorem

Theorem 3.1. *There exist $q := q(\nu) > 0$, $s := s(\nu) > 0$ such that: for any $f \in \mathcal{C}^q(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$, there exists $\varepsilon_0 = \varepsilon_0(\nu, f) > 0$ small enough such that for all $\varepsilon \in (0, \varepsilon_0)$, there exists a Cantor set $\mathcal{C}_\varepsilon \subseteq \Omega$ of asimptotically full Lebesgue measure i.e.*

$$|\mathcal{C}_\varepsilon| \to |\Omega| \qquad as \qquad \varepsilon \to 0,$$

such that for any $\omega \in \mathcal{C}_\varepsilon$ there exists $\mathbf{u}(\varepsilon, \omega) = (u(\varepsilon, \omega), \psi(\varepsilon, \omega)) \in H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$ satisfying $\mathcal{F}(\varepsilon, \omega, \mathbf{u}(\varepsilon, \omega)) = 0$ where the nonlinear operator $\mathcal{F}$ is defined in (3.7) and

$$\|\mathbf{u}(\varepsilon, \omega)\|_s \to 0 \qquad as \qquad \varepsilon \to 0.$$

Theorem 3.1 is based on a Nash–Moser iterative scheme implemented in Sect. 7. he key ingredient in the proof—which also implies the linear stability of the quasi-periodic solutions—is the *reducibility* of the linear operator $\mathcal{L} = \mathcal{L}(\mathbf{u}) = \partial_\mathbf{u}\mathcal{F}(\mathbf{u})$ obtained by linearizing (3.7) at any approximate (or exact) solution $\mathbf{u} = (u, \psi)$. This is the content of Sects. 4 and 5. The proof of the invertibility of $\mathcal{L}$ and the tame estimates for its inverse is provided in Sect. 6. The measure estimate of the set $\mathcal{C}_\varepsilon$ of the *good parameters* is provided in Sect. 8.

4. Regularization of the linearized operator

For any family $\omega \in \Omega_o(\mathbf{u}) \mapsto \mathbf{u}(\cdot; \omega) := (u(\cdot; \omega), \psi(\cdot; \omega)) \in H_0^S(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$, we consider the linearized operator $\mathcal{L} = \mathcal{L}(\mathbf{u}) = \mathcal{L}(\omega, \mathbf{u}(\omega)) := \partial_\mathbf{u}\mathcal{F}(\varepsilon, \omega, \mathbf{u}(\omega)) : H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2) \to H_0^{s-2}(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$ for $2 \le s \le S - 2$ (recall (2.2)). It has the form

$$\mathcal{L}[\widehat{u}, \widehat{\psi}] := \begin{pmatrix} \omega \cdot \partial_\varphi \widehat{u} - \widehat{\psi} \\ \omega \cdot \partial_\varphi \widehat{\psi} - a(\varphi)\partial_{xx}\widehat{u} + \mathcal{R}\widehat{u} \end{pmatrix} \tag{4.1}$$

where

$$a(\varphi) := 1 + \varepsilon \int_{\mathbb{T}} (\partial_x u(\varphi, x))^2 \, dx, \qquad \mathcal{R}\widehat{u} := 2\varepsilon\partial_{xx}u \int_{\mathbb{T}} (\partial_{xx}u)\, \widehat{u} \, dx. \tag{4.2}$$

Along this section, we will always assume the following hypothesis, which will be verified along the Nash–Moser nonlinear iteration of Section 7.

- ASSUMPTION. The function $\mathbf{u} := (u, \psi)$ depends in a Lipschitz way on the parameter $\omega \in \Omega_o := \Omega_o(\mathbf{u}) \subset \Omega_{\gamma,\tau}$ with $\gamma \in (0, 1)$, $\tau > 0$ (recall (2.7)) and for some $\mu := \mu(\tau, \nu) > 0$, for some $S \ge s_0 + \mu$, the map $\omega \in \Omega_o(\mathbf{u}) \mapsto \mathbf{u}(\cdot; \omega) \in H_0^S(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$ satisfies

$$\|\mathbf{u}\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \le 1 \quad \text{and} \quad \varepsilon\gamma^{-1} \ll 1 \tag{4.3}$$

  where we recall that $s_0 := [(\nu+1)/2] + 1$, so that $H^{s_0}(\mathbb{T}^{\nu+1})$ is compactly embedded in $\mathcal{C}^0(\mathbb{T}^{\nu+1})$. We remark that in Sects. 4–7, the constant $\tau > 0$ is independent from the number of frequencies ν. It will be fixed as a function of ν only in Sect. 8 for the measure estimates (see (8.2)).

The function a and the operator $\mathcal{R}$ in (4.2) depend only on the first component u of the function $\mathbf{u} = (u, \psi)$. We denote by $\partial_u a[h]$, $\partial_u \mathcal{R}[h]$ their derivatives with recpect to u in the direction h.

Note that, since $a(\varphi)$ is a real valued function and $\mathcal{R}$ is symmetric, the operator $\mathcal{L}$ is Hamiltonian in the sense of the definition 2.1. Let us give some estimates on a and $\mathcal{R}$ defined in (4.2).

Lemma 4.1. *Assume* (4.3), *with* $\mu = 2$. *Then for any* $s_0 \leq s \leq S - 2$ *the following holds:*

$$\|a - 1\|_s \leq_s \varepsilon \|u\|_{s+1}, \qquad \|a - 1\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon \|u\|_{s+1}^{\mathrm{Lip}(\gamma)}, \tag{4.4}$$

$$\|\partial_u a[h]\|_s \leq_s \varepsilon(\|h\|_{s+1} + \|u\|_{s+1}\|h\|_{s_0+1}). \tag{4.5}$$

The operator $\mathcal{R}$ *in* (4.2) *has the form* (2.101), *with* q *and* g *satisfying the estimates*

$$\|q\|_s \leq_s \varepsilon \|u\|_{s+2}, \quad \|g\|_s \leq_s \|u\|_{s+2}, \tag{4.6}$$

$$\|q\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon \|u\|_{s+2}^{\mathrm{Lip}(\gamma)}, \quad \|g\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|u\|_{s+2}^{\mathrm{Lip}(\gamma)}, \tag{4.7}$$

$$\|\partial_u q[h]\|_s \leq_s \varepsilon \|h\|_{s+2}, \quad \|\partial_u g[h]\|_s \leq_s \|h\|_{s+2}. \tag{4.8}$$

Proof. The estimates (4.4), (4.5) follow by the definition (4.2) and by the interpolation Lemma 2.1, using the condition (4.3). The estimates (4.6)–(4.8) follow since $\mathcal{R}$ is an operator of the form (2.101), with $q := 2\varepsilon\partial_{xx}u$ and $g := \partial_{xx}u$. $\qquad\square$

Notation. In the following, with a slight abuse of notations, for any function $a(\varphi)$, we simply denote by $a = a(\varphi)$, the multiplication operator $h(\varphi, x) \mapsto a(\varphi)h(\varphi, x)$, acting on the space of functions with zero average in x.

4.1. Symplectic symmetrization of the highest order

We start by symmetrizing the highest order of the operator

$$\mathcal{L} = \begin{pmatrix} \omega \cdot \partial_\varphi & -1 \\ -a\partial_{xx} + \mathcal{R} & \omega \cdot \partial_\varphi \end{pmatrix}.$$

Let us consider the transformation

$$\mathcal{S} = \mathcal{S}(\varphi) := \begin{pmatrix} \beta(\varphi)|D|^{-\frac{1}{2}} & 0 \\ 0 & \beta(\varphi)^{-1}|D|^{\frac{1}{2}} \end{pmatrix} \tag{4.9}$$

where $\beta : \mathbb{T}^\nu \to \mathbb{R}$ is a Sobolev function close to 1 to be determined (recall also the definition (2.84)). The inverse of the operator $\mathcal{S}$ (acting on Sobolev spaces of zero average functions in x) is given by

$$\mathcal{S}^{-1} = \mathcal{S}(\varphi)^{-1} := \begin{pmatrix} \beta(\varphi)^{-1}|D|^{\frac{1}{2}} & 0 \\ 0 & \beta(\varphi)|D|^{-\frac{1}{2}} \end{pmatrix}. \tag{4.10}$$

Using that for any function $a = a(\varphi)$ depending only on time, the commmutators $[a, |D|^m] = 0$, $[a, \mathcal{R}] = 0$ where $\mathcal{R}$ is defined in (4.2) and since $-\partial_{xx} = |D|^2$, we have

$$\mathcal{S}^{-1}\mathcal{L}\mathcal{S} = \begin{pmatrix} \omega \cdot \partial_\varphi + \beta^{-1}(\omega \cdot \partial_\varphi \beta) & -\beta^{-2}|D| \\ a\beta^2|D| + \beta^2|D|^{-\frac{1}{2}}\mathcal{R}|D|^{-\frac{1}{2}} & \omega \cdot \partial_\varphi + \beta\omega \cdot \partial_\varphi(\beta^{-1}) \end{pmatrix}. \tag{4.11}$$

We choose $\beta(\varphi)$ so that $\beta^{-2}(\varphi) = a(\varphi)\beta^2(\varphi)$, namely we define

$$\beta(\varphi) := \frac{1}{[a(\varphi)]^{\frac{1}{4}}}. \tag{4.12}$$

Since $\beta\omega \cdot \partial_\varphi(\beta^{-1}) = -\beta^{-1}\omega \cdot \partial_\varphi\beta$, we get that

$$\mathcal{L}_1 := \mathcal{S}^{-1}\mathcal{L}\mathcal{S} = \begin{pmatrix} \omega \cdot \partial_\varphi + a_0 & -a_1|D| \\ a_1|D| + \mathcal{R}^{(1)} & \omega \cdot \partial_\varphi - a_0 \end{pmatrix}, \tag{4.13}$$

where

$$a_0 := \frac{\omega \cdot \partial_\varphi\beta}{\beta}, \quad a_1 := \sqrt{a}, \quad \mathcal{R}^{(1)} := \beta^2|D|^{-\frac{1}{2}}\mathcal{R}|D|^{-\frac{1}{2}}. \tag{4.14}$$

Since $\beta(\varphi)$ is a real-valued function and the operators $|D|^{\pm\frac{1}{2}}$ are real operators, the operator $\mathcal{S}$ is real. A direct verification shows that it is also symplectic (see Definition 2.2). Hence the transformed operator $\mathcal{L}_1$ is still real and Hamiltonian (see Definition 2.1). Now we give some estimates on the coefficients of the operator $\mathcal{L}_1$.

Lemma 4.2. *Assume* (4.3), *with* $\mu = 2$. *Then for any* $s_0 \leq s \leq S - 2$ *the following holds: the maps* $\mathcal{S}^{\pm 1} : H_0^{s+\frac{1}{2}}(\mathbb{T}^{\nu+1}, \mathbb{R}^2) \to H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$ *satisfy the estimates*

$$\|\mathcal{S}^{\pm 1}\mathbf{h}\|_s \leq_s \|\mathbf{h}\|_{s+\frac{1}{2}} + \|u\|_{s+1}\|\mathbf{h}\|_{s_0+\frac{1}{2}}, \quad \mathbf{h} \in H^{s+\frac{1}{2}}(\mathbb{T}^{\nu+1}, \mathbb{R}^2). \tag{4.15}$$

For any family $\mathbf{h}(\cdot; \omega) \in H_0^{s+\frac{1}{2}}(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$, $\omega \in \Omega_o$,

$$\|\mathcal{S}^{\pm 1}\mathbf{h}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|\mathbf{h}\|_{s+\frac{1}{2}}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+1}^{\mathrm{Lip}(\gamma)}\|\mathbf{h}\|_{s_0+\frac{1}{2}}^{\mathrm{Lip}(\gamma)}. \tag{4.16}$$

The functions a_0, a_1 *defined in* (4.14) *satisfy the estimates*

$$\|a_1 - 1\|_s, \|a_0\|_s \leq_s \varepsilon(1 + \|u\|_{s+2}), \tag{4.17}$$

$$\|a_1 - 1\|_s^{\mathrm{Lip}(\gamma)}, \|a_0\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon(1 + \|u\|_{s+2}^{\mathrm{Lip}(\gamma)}), \tag{4.18}$$

$$\|\partial_u a_k[h]\|_s \leq_s \varepsilon(\|h\|_{s+2} + \|u\|_{s+2}\|h\|_{s_0+2}), \quad k = 0, 1. \tag{4.19}$$

The remainder $\mathcal{R}^{(1)}$ *in* (4.14) *has the form* (2.101), *with* $q = q_1$, $g = g_1$ *satisfying the estimates*

$$\|q_1\|_s \leq_s \varepsilon\|u\|_{s+2}, \quad \|g_1\|_s \leq_s \|u\|_{s+2}, \tag{4.20}$$

$$\|q_1\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon\|u\|_{s+2}^{\mathrm{Lip}(\gamma)}, \quad \|g_1\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|u\|_{s+2}^{\mathrm{Lip}(\gamma)}, \tag{4.21}$$

$$\|\partial_u q_1[h]\|_s \leq_s \varepsilon(\|h\|_{s+2} + \|u\|_{s+2}\|h\|_{s_0+2}), \quad \|\partial_u g_1[h]\|_s \leq_s \|h\|_{s+2}. \tag{4.22}$$

Proof. The estimates (4.15)–(4.19) follow by the definitions (4.9), (4.10), (4.12), (4.14), by the estimates (4.4) and by Lemmata 2.1, 2.2. Let us prove the estimates (4.20)–(4.22). By (4.2), (4.14), using that $|D|^{-\frac{1}{2}}$ is symmetric, one has that $\mathcal{R}^{(1)}h = q_1 \int_{\mathbb{T}} g_1 h \, dx$ with

$$q_1 := 2\varepsilon\beta^2(|D|^{-\frac{1}{2}}\partial_{xx}u), \quad g_1 := |D|^{-\frac{1}{2}}\partial_{xx}u. \tag{4.23}$$

One can estimate the function β in (4.12) by using Lemma 2.2 and the estimate (4.4). Applying the interpolation Lemma 2.1, the claimed estimates follow. $\square$

Lemma 4.3. *The operators $\mathcal{S}^{\pm 1}$ defined in* (4.9), (4.10) *can be regarded as an operator acting on the Sobolev space of the functions in x, namely for any $s \geq 1$, for any $\varphi \in \mathbb{T}^{\nu}$,*

$$\mathcal{S}(\varphi) \in \mathcal{L}\left(H_0^{s-\frac{1}{2}}(\mathbb{T}_x, \mathbb{R}^2), H_0^s(\mathbb{T}_x, \mathbb{R}) \times H_0^{s-1}(\mathbb{T}_x, \mathbb{R})\right),$$

$$\mathcal{S}(\varphi)^{-1} \in \mathcal{L}\left(H_0^s(\mathbb{T}_x, \mathbb{R}) \times H_0^{s-1}(\mathbb{T}_x, \mathbb{R}), H_0^{s-\frac{1}{2}}(\mathbb{T}_x, \mathbb{R}^2)\right).$$

Proof. By the definition of the function $\beta(\varphi)$ in (4.12), using the estimate (4.4) on $a(\varphi)$, the Lemma 2.2 and the ansatz (4.3), one gets $\|\beta^{\pm 1}\|_{L^{\infty}(\mathbb{T}^{\nu})} \ll 1$. Moreover $\||D|^{\frac{1}{2}}h\|_{H_x^s} \leq \|h\|_{H_x^{s+\frac{1}{2}}}, \quad \||D|^{-\frac{1}{2}}h\|_{H_x^s} \leq \|h\|_{H_s^{s-\frac{1}{2}}}$ and then the Lemma follows. $\qquad\square$

4.2. Complex variables

Now we consider the complex variables $z := \frac{\widehat{u} + i\widehat{\psi}}{\sqrt{2}}$ introduced in (2.35), (2.36) in order to write the operator $\mathcal{L}_1$ defined in (4.13) in complex coordinates. More precisely, we consider the transformations

$$\mathcal{B} := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ \frac{1}{i} & -\frac{1}{i} \end{pmatrix} \qquad \mathcal{B}^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \tag{4.24}$$

and we get that the conjugated operator $\mathcal{L}_2 := \mathcal{B}^{-1}\mathcal{L}_1\mathcal{B}$ is given by

$$\mathcal{L}_2 = \begin{pmatrix} \omega \cdot \partial_{\varphi} + ia_1|D| + i\mathcal{R}^{(2)} & a_0 + i\mathcal{R}^{(2)} \\ a_0 - i\mathcal{R}^{(2)} & \omega \cdot \partial_{\varphi} - ia_1|D| - i\mathcal{R}^{(2)} \end{pmatrix}, \tag{4.25}$$

with $\mathcal{R}^{(2)} := \frac{\mathcal{R}^{(1)}}{2}$. Since a_1 and a_0 are real valued functions and $\mathcal{R}^{(1)}$ (and then $\mathcal{R}^{(2)}$) is symmetric and real, the operator $\mathcal{L}_2$ is a Hamiltonian operator in complex coordinates, in the sense of the Definition 2.4. Note that the transformations $\mathcal{B}^{\pm 1}$ satisfy for all $s \geq 0$

$$\mathcal{B} : \mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2), \quad \mathcal{B}^{-1} : H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1}), \quad (4.26)$$

$$\mathcal{B} : \mathbf{H}_0^s(\mathbb{T}_x) \to H_0^s(\mathbb{T}_x, \mathbb{R}^2), \quad \mathcal{B}^{-1} : H_0^s(\mathbb{T}_x, \mathbb{R}^2) \to \mathbf{H}_0^s(\mathbb{T}_x) \tag{4.27}$$

where we recall that the real subspace $\mathbf{H}_0^s(\mathbb{T}^{\nu+1})$, resp. $\mathbf{H}_0^s(\mathbb{T}_x)$ of $H_0^s(\mathbb{T}^{\nu+1}, \mathbb{C}^2)$, resp. $H_0^s(\mathbb{T}_x, \mathbb{C}^2)$, is defined in (2.41).

4.3. Change of variables

The aim of this section is to reduce to constant coefficients the highest order term $a_1(\varphi)|D|$ in the operator $\mathcal{L}_2$ defined in (4.25). In order to do this, let us consider a diffeomorphism of the torus $\mathbb{T}^{\nu}$ of the form

$$\varphi \in \mathbb{T}^{\nu} \mapsto \varphi + \omega\alpha(\varphi) \in \mathbb{T}^{\nu},$$

where $\alpha : \mathbb{T}^{\nu} \to \mathbb{R}$ has to be determined. This diffeomorphism of the torus induces on the space of functions $h(\varphi, x)$ a linear operator

$$(\mathcal{A}h)(\varphi, x) := h(\varphi + \omega\alpha(\varphi), x), \tag{4.28}$$

whose inverse has the form

$$\mathcal{A}^{-1}h(\vartheta, x) := h(\vartheta + \omega\tilde{\alpha}(\vartheta), x), \tag{4.29}$$

where $\vartheta \to \vartheta + \omega \tilde{\alpha}(\vartheta)$ is the inverse diffeomorphism of $\varphi \to \varphi + \omega \alpha(\varphi)$. One has

$$\mathcal{A}^{-1}(\omega \cdot \partial_\varphi)\mathcal{A} = \mathcal{A}^{-1}[1 + \omega \cdot \partial_\varphi \alpha]\omega \cdot \partial_\vartheta, \quad \mathcal{A}^{-1}|D|\mathcal{A} = |D|, \quad \mathcal{A}^{-1}a\mathcal{A} = \mathcal{A}^{-1}[a]$$

where we recall that a denotes the multiplication operator $h \to ah$. Recalling that $\mathbb{I}_2 := \begin{pmatrix} \mathrm{Id}_0 & 0 \\ 0 & \mathrm{Id}_0 \end{pmatrix}$, where $\mathrm{Id}_0 : L_0^2 \to L_0^2$ is the identity and defining

$$\rho := \mathcal{A}^{-1}[1 + \omega \cdot \partial_\varphi \alpha], \tag{4.30}$$

we get

$$\mathcal{A}^{-1}\mathbb{I}_2\mathcal{L}_2\mathcal{A}\mathbb{I}_2$$
$$= \begin{pmatrix} \rho\omega \cdot \partial_\vartheta + \mathrm{i}\mathcal{A}^{-1}[a_1]|D| + \mathrm{i}\mathcal{A}^{-1}\mathcal{R}^{(2)}\mathcal{A} & \mathcal{A}^{-1}[a_0] + \mathrm{i}\mathcal{A}^{-1}\mathcal{R}^{(2)}\mathcal{A} \\ \mathcal{A}^{-1}[a_0] - \mathrm{i}\mathcal{A}^{-1}\mathcal{R}^{(2)}\mathcal{A} & \rho\omega \cdot \partial_\vartheta - \mathrm{i}\mathcal{A}^{-1}[a_1]|D| - \mathrm{i}\mathcal{A}^{-1}\mathcal{R}^{(2)}\mathcal{A} \end{pmatrix}.$$

We want to choose the function α so that the coefficient ρ in front of $\omega \cdot \partial_\vartheta$ is proportional to the coefficient $\mathcal{A}^{-1}[a_1]$ in front of the operator $|D|$. To this aim it is enough to solve the equation

$$m\left(1 + \omega \cdot \partial_\varphi \alpha(\varphi)\right) = a_1(\varphi) \quad m \in \mathbb{R}. \tag{4.31}$$

Integrating on $\mathbb{T}^\nu$ we fix the value of m as

$$m := \frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} a_1(\varphi)\,d\varphi \tag{4.32}$$

and then, since $\omega \in \Omega_o \subseteq \Omega_{\gamma,\tau}$, recalling the definitions (2.6), (2.7) we get

$$\alpha(\varphi) = (\omega \cdot \partial_\varphi)^{-1}\left[\frac{a_1}{m} - 1\right](\varphi). \tag{4.33}$$

Note that, since the function a_1 is real valued, m is real and then α is a real valued function. We have $\mathcal{A}^{-1}\mathbb{I}_2\mathcal{L}_2\mathcal{A}\mathbb{I}_2 = \rho\mathcal{L}_3$, with

$$\mathcal{L}_3 := \begin{pmatrix} \omega \cdot \partial_\vartheta + \mathrm{i}m|D| + \mathrm{i}\mathcal{R}^{(3)} & b_0 + \mathrm{i}\mathcal{R}^{(3)} \\ b_0 - \mathrm{i}\mathcal{R}^{(3)} & \omega \cdot \partial_\vartheta - \mathrm{i}m|D| - \mathrm{i}\mathcal{R}^{(3)} \end{pmatrix}, \tag{4.34}$$

$$b_0 := \rho^{-1}\mathcal{A}^{-1}[a_0], \qquad \mathcal{R}^{(3)} := \rho^{-1}\mathcal{A}^{-1}\mathcal{R}^{(2)}\mathcal{A}. \tag{4.35}$$

Note that the operator $\mathcal{L}_3$ is still Hamiltonian in the sense of the definition 2.4, since $m \in \mathbb{R}$, $|D|$ is a symmetric real operator, b_0 is a real valued function and $\mathcal{R}^{(3)}$ is a real and symmetric operator, implying that $(\mathcal{R}^{(3)})^* = \overline{(\mathcal{R}^{(3)})^T} = \mathcal{R}^{(3)}$.

Lemma 4.4. *There exists a constant $\sigma = \sigma(\tau, \nu) > 2$ such that if (4.3) holds with $\mu = \sigma$, then for all $s_0 \leq s \leq S - \sigma$ the following estimates hold:*

$$|m - 1| \lessdot \varepsilon, \quad |m - 1|^{\mathrm{Lip}(\gamma)} \lessdot \varepsilon, \quad |\partial_u m[h]| \lessdot \varepsilon\|h\|_{s_0+2}. \tag{4.36}$$

The transformations $\mathcal{A}^{\pm 1} : H_0^s(\mathbb{T}^{\nu+1}, \mathbb{C}) \to H_0^s(\mathbb{T}^{\nu+1}, \mathbb{C})$ satisfy

$$\|\mathcal{A}^{\pm 1}h\|_s \leq_s \|h\|_s + \|u\|_{s+\sigma}\|h\|_{s_0+1}, \tag{4.37}$$

$$\|\mathcal{A}^{\pm 1}h\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|h\|_{s+1}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|h\|_{s_0+2}^{\mathrm{Lip}(\gamma)}, \tag{4.38}$$

$$\|\partial_u(\mathcal{A}^{\pm 1}h)g\|_s \leq_s \varepsilon\gamma^{-1}\left(\|h\|_{s+\sigma}\|g\|_{s_0+\sigma} + \|h\|_{s_0+\sigma}\|g\|_{s+\sigma}\right.$$
$$\left. + \|u\|_{s+\sigma}\|h\|_{s_0+\sigma}\|g\|_{s_0+\sigma}\right). \tag{4.39}$$

The function ρ defined in (4.30) satisfies

$$\|\rho^{\pm 1} - 1\|_s \leq_s \varepsilon(1 + \|u\|_{s+\sigma}), \qquad \|\rho^{\pm 1} - 1\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon(1 + \|u\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}), \quad (4.40)$$

$$\|\partial_u \rho^{\pm 1}[h]\|_s \leq_s \varepsilon \left(\|h\|_{s+\sigma} + \|h\|_{s+\sigma}\|h\|_{s_0+\sigma}\right). \qquad (4.41)$$

The function b_0 defined in (4.35) satisfies

$$\|b_0\|_s \leq_s \varepsilon(1 + \|u\|_{s+\sigma}), \quad \|b_0\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon(1 + \|u\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}), \qquad (4.42)$$

$$\|\partial_u b_0[h]\|_s \leq_s \varepsilon \left(\|h\|_{s+\sigma} + \|u\|_{s+\sigma}\|h\|_{s_0+\sigma}\right). \qquad (4.43)$$

The remainder $\mathcal{R}^{(3)}$ defined in (4.35) satisfies the estimates

$$|\mathcal{R}^{(3)}|D|\|_s \leq_s \varepsilon(1 + \|u\|_{s+\sigma}), \; |\mathcal{R}^{(3)}|D|\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon(1 + \|u\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}), \quad (4.44)$$

$$|\partial_u \mathcal{R}^{(3)}[h]|_s \leq_s \varepsilon \left(\|h\|_{s+\sigma} + \|u\|_{s+\sigma}\|h\|_{s_0+\sigma}\right). \qquad (4.45)$$

Proof. The estimates (4.36) follow by the formula (4.32) and using the estimates (4.17), (4.19). The transformation $\mathcal{A}$ has been also used in [6–8,15,28]. The proof of the estimates (4.37)–(4.41) can be done by using Lemma 2.3 as in these papers. For a detailed proof see for instance [28], pp. 25–26. Let us prove the estimates (4.44), (4.45). One has

$$\mathcal{R}^{(3)}h = \rho^{-1}\mathcal{A}^{-1}\mathcal{R}^{(2)}\mathcal{A}h \overset{(4.25)}{=} \frac{1}{2}\rho^{-1}\mathcal{A}^{-1}\mathcal{R}^{(1)}\mathcal{A}h \overset{(4.23)}{=} q_3 \int_{\mathbb{T}} g_3\, h\, dx,$$

with

$$q_3 := \frac{1}{2}\rho^{-1}\mathcal{A}^{-1}(q_1), \qquad g_3 := \mathcal{A}^{-1}(g_1).$$

Therefore, the functions q_3 and g_3, can be estimated by using (4.20), (4.22), (4.37)–(4.41) and Lemma 2.1. The estimates in (4.44) then follow by applying Lemma 2.12. The estimate for $\partial_u \mathcal{R}^{(3)}[h]$ follows by differentiating the expression of $\mathcal{R}^{(3)}$, q_3, g_3 given above and applying again Lemma 2.12. $\qquad \square$

4.4. Descent method

Introducing the notation

$$T := \begin{pmatrix} \mathrm{Id}_0 & 0 \\ 0 & -\mathrm{Id}_0 \end{pmatrix} \qquad (4.46)$$

we can write the operator $\mathcal{L}_3$ in (4.34) as

$$\mathcal{L}_3 = \omega \cdot \partial_\varphi \mathbb{I}_2 + imT|D| + B_0 + \mathcal{R}_3, \qquad (4.47)$$

where

$$B_0(\varphi) := \begin{pmatrix} 0 & b_0(\varphi) \\ b_0(\varphi) & 0 \end{pmatrix}, \quad \mathcal{R}_3 := i \begin{pmatrix} \mathcal{R}^{(3)} & \mathcal{R}^{(3)} \\ -\mathcal{R}^{(3)} & -\mathcal{R}^{(3)} \end{pmatrix}. \qquad (4.48)$$

Our aim is to eliminate from the operator $\mathcal{L}_3$ the terms of order $|D|^0$, namely, since $\mathcal{R}^{(3)}$ is an operator of the form (2.101) (then arbitrarily regularizing), we only need to remove the multiplication operator by the matrix valued function $B_0(\varphi)$.

For this purpose, we consider the operator

$$\mathcal{V} = \mathcal{V}(\varphi) := \exp\left(iV(\varphi)|D|^{-1}\right), \quad V(\varphi) := \begin{pmatrix} 0 & v(\varphi) \\ -v(\varphi) & 0 \end{pmatrix}, \qquad (4.49)$$

where $v : \mathbb{T}^\nu \to \mathbb{R}$ is a real valued function to be determined. Note that $\mathcal{V}$ is symplectic, since $iV(\varphi)|D|^{-1}$ is a Hamiltonian vector field. We write

$$\mathcal{V} = \mathbb{I}_2 + iV|D|^{-1} + \mathcal{V}_{\geq 2}, \qquad \mathcal{V}_{\geq 2} := \sum_{k \geq 2} \frac{i^k}{k!} V^k |D|^{-k},$$

hence

$$\mathcal{L}_3 \mathcal{V} = \mathcal{V}\left(\omega \cdot \partial_\varphi \mathbb{I}_2 + imT|D|\right) + [imT|D|, iV|D|^{-1}] + B_0 + B_0(\mathcal{V} - \mathbb{I}_2)$$
$$+ [imT|D|, \mathcal{V}_{\geq 2}] + i\omega \cdot \partial_\varphi(\mathcal{V} - \mathbb{I}_2) + \mathcal{R}_3 \mathcal{V}. \qquad (4.50)$$

The term of order $|D|^0$ is given by

$$[imT|D|, iV(\varphi)|D|^{-1}] + B_0(\varphi) = \begin{pmatrix} 0 & -2mv(\varphi) + b_0(\varphi) \\ -2mv(\varphi) + b_0(\varphi) & 0 \end{pmatrix}.$$

In order to remove it, we choose

$$v(\varphi) := \frac{b_0(\varphi)}{2m} \qquad (4.51)$$

and we get

$$\mathcal{L}_4 := \mathcal{V}^{-1}\mathcal{L}_3\mathcal{V} = \omega \cdot \partial_\varphi \mathbb{I}_2 + imT|D| + \mathcal{R}_4, \qquad (4.52)$$
$$\mathcal{R}_4 := \mathcal{V}^{-1}\left(B_0(\varphi)(\mathcal{V} - \mathbb{I}_2) + [imT|D|, \mathcal{V}_{\geq 2}] + \omega \cdot \partial_\varphi(\mathcal{V} - \mathbb{I}_2)\right) + \mathcal{V}^{-1}\mathcal{R}_3\mathcal{V}. \qquad (4.53)$$

Note that, since $\mathcal{L}_3$ is Hamiltonian and $\mathcal{V}$ is symplectic, we have that $\mathcal{L}_4$ is still a Hamiltonian operator. In the next lemma we provide some estimates on the transformation $\mathcal{V}$ and on the remainder $\mathcal{R}_4$.

Lemma 4.5. *There exists $\overline{\sigma} = \overline{\sigma}(\tau, \nu) > \sigma > 0$, where σ is the loss of derivatives in Lemma 4.4, such that if (4.3) holds with $\mu = \overline{\sigma}$, then for any $s_0 \leq s \leq S - \overline{\sigma}$, $\mathcal{V}^{\pm 1} : \mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1})$ (recall (2.41)) and the following estimates hold:*

$$|(\mathcal{V}^{\pm 1} - \mathbb{I}_2)|D||_s \lesssim_s \varepsilon(1 + \|u\|_{s+\overline{\sigma}}), \qquad (4.54)$$

$$|(\mathcal{V}^{\pm 1} - \mathbb{I}_2)|D||_s^{\mathrm{Lip}(\gamma)} \lesssim_s \varepsilon(1 + \|u\|_{s+\overline{\sigma}}^{\mathrm{Lip}(\gamma)}), \qquad (4.55)$$

$$|\partial_u \mathcal{V}^{\pm 1} h|_s \lesssim_s \varepsilon(\|h\|_{s+\overline{\sigma}} + \|u\|_{s+\overline{\sigma}}\|h\|_{s_0+\overline{\sigma}}), \qquad (4.56)$$

$$|\mathcal{R}_4|D||_s \lesssim_s \varepsilon(1 + \|u\|_{s+\overline{\sigma}}), \quad |\mathcal{R}_4|D||_s^{\mathrm{Lip}(\gamma)} \lesssim_s \varepsilon(1 + \|u\|_{s+\overline{\sigma}}^{\mathrm{Lip}(\gamma)}), \qquad (4.57)$$

$$|\partial_u \mathcal{R}_4[h]|_s \lesssim_s \varepsilon(\|h\|_{s+\overline{\sigma}} + \|u\|_{s+\overline{\sigma}}\|h\|_{s_0+\overline{\sigma}}). \qquad (4.58)$$

Proof. Proof of (4.54)–(4.56). By Lemma 2.9 one has

$$|(V|D|^{-1})|D||_s = \|V\|_s \leq \|v\|_s \overset{(4.36),(4.42)}{\lesssim_s} \varepsilon(1 + \|u\|_{s+\sigma}). \qquad (4.59)$$

By (4.3) we have that $|(V|D|^{-1})|D||_{s_0} = \|V\|_{s_0} \lessdot \varepsilon \leq 1$, for ε small enough, then Lemma 2.10 can be applied and the claimed estimate (4.54) follows. The estimate (4.56) follows by applying the estimate (2.91) and using that by (4.51), (4.36), (4.43), $||D|^{-1}|_s \leq 1$

$$|\partial_u V|D|^{-1}|_s \lessdot \|\partial_u v[h]\|_s \lesssim_s \varepsilon(\|h\|_{s+\sigma} + \|u\|_{s+\sigma}\|h\|_{s_0+\sigma}).$$

PROOF OF (4.57), (4.58). The claimed estimates follow by the definition (4.53), by applying Lemmata 2.6, 2.7, 2.9, 2.10. and by the estimates (4.42)–(4.45) and (4.54)–(4.56).   □

Lemma 4.6. *Assume* (4.3) *with* $\mu = \overline{\sigma} + s_0$. *Then for any* $s_0 \leq s \leq S - \overline{\sigma} - s_0$ *for any* $\varphi \in \mathbb{T}^\nu$, $\mathcal{V}^{\pm 1}(\varphi) : \mathbf{H}_0^s(\mathbb{T}_x) \to \mathbf{H}_0^s(\mathbb{T}_x)$ *(recall* (2.42)*) and*

$$|\mathcal{V}^{\pm 1}(\varphi)|_{s,x} \leq_s 1 + \|u\|_{s+\overline{\sigma}+s_0}.$$

Proof. The claimed estimate follows by applying Lemma 2.13-(ii) and by the estimates (4.54).   □

5. 2×2 block-diagonal reduction

The goal of this section is to block-diagonalize the linear Hamiltonian operator $\mathcal{L}_4$ obtained in (4.52). We are going to perform an iterative Nash–Moser reducibility scheme for the linear Hamiltonian operator

$$\mathcal{L}_0 := \mathcal{L}_4 = \omega \cdot \partial_\varphi \mathbb{I}_2 + \mathcal{D}_0 + \mathcal{R}_0 : \mathbf{H}_0^1(\mathbb{T}^{\nu+1}) \to \mathbf{L}_0^2(\mathbb{T}^{\nu+1}), \qquad (5.1)$$

where

$$\mathcal{D}_0 = \mathrm{i} \begin{pmatrix} \mathcal{D}_0^{(1)} & 0 \\ 0 & -\mathcal{D}_0^{(1)} \end{pmatrix}, \quad \mathcal{D}_0^{(1)} := m|D| = \mathrm{diag}_{j \in \mathbb{Z} \setminus \{0\}} m|j| \qquad (5.2)$$

and $\mathcal{R}_0 := \mathcal{R}_4$ is a Hamiltonian operator of the form

$$\mathcal{R}_0 = \mathrm{i} \begin{pmatrix} \mathcal{R}_0^{(1)} & \mathcal{R}_0^{(2)} \\ -\overline{\mathcal{R}}_0^{(2)} & -\overline{\mathcal{R}}_0^{(1)} \end{pmatrix}, \quad \mathcal{R}_0^{(1)} = \left(\mathcal{R}_0^{(1)} \right)^*, \quad \mathcal{R}_0^{(2)} = (\mathcal{R}_0^{(2)})^T \qquad (5.3)$$

satisfying, by (4.57), for any $s_0 \leq s \leq S - \overline{\sigma}$ the estimates

$$|\mathcal{R}_0|D|\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon(1 + \|u\|_{s+\overline{\sigma}}), \ |\partial_u \mathcal{R}_0[h]|_s \leq_s \varepsilon \left(\|h\|_{s+\overline{\sigma}} + \|u\|_{s+\overline{\sigma}} \|h\|_{s_0+\overline{\sigma}} \right), \tag{5.4}$$

where $\overline{\sigma}$ is the loss of derivatives given in Lemma 4.5. We define

$$N_{-1} := 1, \quad N_\nu := N_0^{\chi^\nu} \ \forall \nu \geq 0, \quad \chi := 3/2 \qquad (5.5)$$

(then $N_{\nu+1} = N_\nu^\chi$, $\forall \nu \geq 0$) and

$$\mathtt{a} := 6\tau + 4, \quad \mathtt{b} := \mathtt{a} + 1. \qquad (5.6)$$

We assume that (4.3) holds with $\mu = \overline{\sigma} + \mathtt{b}$, so that by (5.4)

$$|\mathcal{R}_0|D|\|_{s_0+\mathtt{b}}^{\mathrm{Lip}(\gamma)} \ll \varepsilon, \qquad |\partial_u \mathcal{R}_0[h]|_{s_0+\mathtt{b}} \ll \varepsilon \|h\|_{s_0+\overline{\sigma}+\mathtt{b}}. \qquad (5.7)$$

For the reducibility Theorem below, we use the 2×2 block representation of linear operators, given in Sect. 2.3. According to (2.58) and recalling also (2.65), the operator $\mathcal{D}_0^{(1)}$ can be written as

$$\mathcal{D}_0^{(1)} = \mathrm{diag}_{j \in \mathbb{N}} mj \mathbf{I}_j, \qquad (5.8)$$

where $\mathbf{I}_j : \mathbf{E}_j \to \mathbf{E}_j$ is the identity, $\mathbf{E}_j = \mathrm{span}\{e^{\mathrm{i}jx}, e^{-\mathrm{i}jx}\}$ is the two dimensional space (2.60) and the real constant m satisfies the estimates (4.36). We also recall the definition of the space $\mathcal{S}(\mathbf{E}_j)$ given in (2.78) which is isomorphic to the space of the 2×2 self-adjoint matrices, the definition of the norm

$\|\cdot\|_{\mathrm{Op}(j,j')}$ given in (2.73), the identity $\mathbf{I}_{j,j'}$ in (2.74), the definition of $M_L(A)$ in (2.75) and the definition of $M_R(B)$ in (2.76). Now we are ready to state the following.

Theorem 5.1. (KAM reducibility) *Let $\gamma \in (0,1)$ and $\tau > 0$. Assume (4.3) with $\mu = \bar{\sigma} + \mathsf{b}$ and with $S \geq s_0 + \bar{\sigma} + \mathsf{b}$. There exist $N_0 = N_0(S, \tau, \nu) > 0$ large enough, $\delta_0 = \delta_0(S, \tau, \nu) \in (0,1)$ small enough, such that, if*

$$\varepsilon\gamma^{-1} \leq \delta_0 \tag{5.9}$$

then:

$(\mathbf{S1})_\nu$ *For all $\nu \geq 0$, there exists an operator*

$$\mathcal{L}_\nu := \omega \cdot \partial_\varphi \mathbb{I}_2 + \mathcal{D}_\nu + \mathcal{R}_\nu \tag{5.10}$$

$$\mathcal{D}_\nu = \mathrm{i} \begin{pmatrix} \mathcal{D}_\nu^{(1)} & 0 \\ 0 & -\overline{\mathcal{D}}_\nu^{(1)} \end{pmatrix}, \qquad \mathcal{D}_\nu^{(1)} := \mathrm{diag}_{j\in\mathbb{N}} \mathbf{D}_j^\nu, \tag{5.11}$$

$$\mathbf{D}_j^\nu := \mathbf{D}_j^\nu(\omega) = \mathbf{D}_j^0(\omega) + \widehat{\mathbf{D}}_j^\nu(\omega), \quad \mathbf{D}_j^0 := mj\mathbf{I}_j, \quad \forall j \in \mathbb{N}, \tag{5.12}$$

(with $\widehat{\mathbf{D}}_j^0 = 0$) defined for all $\omega \in \Omega_\nu^\gamma(\mathbf{u})$, where $\Omega_0^\gamma(\mathbf{u}) := \Omega_o = \Omega_o(\mathbf{u})$, and for $\nu \geq 1$, $\Omega_\nu^\gamma = \Omega_\nu^\gamma(\mathbf{u})$ is defined by

$$\Omega_\nu^\gamma := \left\{ \omega \in \Omega_{\nu-1}^\gamma : \|\mathbf{A}_{\nu-1}^-(\ell, j, j')^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{\langle\ell\rangle^\tau}{\gamma\langle j - j'\rangle}, \forall(\ell, j, j') \neq (0, j, j), \right.$$

$$|\ell| \leq N_{\nu-1}, \|\mathbf{A}_{\nu-1}^+(\ell, j, j')^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{\langle\ell\rangle^\tau}{\gamma\langle j + j'\rangle},$$

$$\left. \forall(\ell, j, j'), |\ell| \leq N_{\nu-1} \right\}, \tag{5.13}$$

where the operators $\mathbf{A}_{\nu-1}^\pm(\ell, j, j') : \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j) \to \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ are defined by

$$\mathbf{A}_{\nu-1}^-(\ell, j, j') := \omega \cdot \ell \mathbf{I}_{j,j'} + M_L(\mathbf{D}_j^{\nu-1}) - M_R(\mathbf{D}_{j'}^{\nu-1}), \tag{5.14}$$

$$\mathbf{A}_{\nu-1}^+(\ell, j, j') := \omega \cdot \ell \mathbf{I}_{j,j'} + M_L(\mathbf{D}_j^{\nu-1}) + M_R(\overline{\mathbf{D}}_{j'}^{\nu-1}). \tag{5.15}$$

For $\nu \geq 0$, for all $j \in \mathbb{N}$, the 2×2 self-adjoint block $\widehat{\mathbf{D}}_j^\nu \in \mathcal{S}(\mathbf{E}_j)$ satisfies

$$\|\widehat{\mathbf{D}}_j^\nu\|^{\mathrm{Lip}(\gamma)} \lessdot \varepsilon j^{-1} \quad \forall j \in \mathbb{N}. \tag{5.16}$$

The Hamiltonian remainder $\mathcal{R}_\nu : \mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1})$ satisfies $\forall s \in [s_0, S - \bar{\sigma} - \mathsf{b}]$,

$$|\mathcal{R}_\nu|D|\|_s^{\mathrm{Lip}(\gamma)} \leq \frac{|\mathcal{R}_0|D|\|_{s+\mathsf{b}}^{\mathrm{Lip}(\gamma)}}{N_{\nu-1}^{\mathsf{a}}}, \ |\mathcal{R}_\nu|D|\|_{s+\mathsf{b}}^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_0|D|\|_{s+\mathsf{b}}^{\mathrm{Lip}(\gamma)} N_{\nu-1}. \tag{5.17}$$

Moreover, for any $\nu \geq 1$,

$$\mathcal{L}_\nu = \Phi_{\nu-1}^{-1} \mathcal{L}_{\nu-1} \Phi_{\nu-1}, \quad \Phi_{\nu-1} := \exp(\Psi_{\nu-1}), \tag{5.18}$$

where $\Psi_{\nu-1}$ is Hamiltonian, $\Phi_{\nu-1}$ is symplectic and they satisfy $\Psi_{\nu-1}$, $\Phi_{\nu-1}^{\pm 1} : \mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1})$,

$$|\Psi_{\nu-1}|_s^{\mathrm{Lip}(\gamma)}, |\Psi_{\nu-1}|D|\|_s^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_0|D|\|_{s+\mathsf{b}}^{\mathrm{Lip}(\gamma)} \gamma^{-1} N_{\nu-1}^{2\tau+1} N_{\nu-2}^{-\mathsf{a}}. \tag{5.19}$$

$(\mathbf{S2})_\nu$ *For all $j \in \mathbb{N}$, there exists a Lipschitz extension to the whole parameter space Ω_o, $\widetilde{\mathbf{D}}_j^\nu(\cdot) : \Omega_o \to \mathcal{S}(\mathbf{E}_j)$ of $\mathbf{D}_j^\nu(\cdot) : \Omega_\nu^\gamma \to \mathcal{S}(\mathbf{E}_j)$ satisfying, for $\nu \geq 1$,*

$$\|\widetilde{\mathbf{D}}_j^\nu - \widetilde{\mathbf{D}}_j^{\nu-1}\|^{\mathrm{Lip}(\gamma)} \lessdot j^{-1}|\mathcal{R}_{\nu-1}|D\|_{s_0}^{\mathrm{Lip}(\gamma)} \lessdot N_{\nu-2}^{-\mathtt{a}}\varepsilon j^{-1}. \tag{5.20}$$

$(\mathbf{S3})_\nu$ *Let $\mathbf{u}_i(\omega) = (u_i(\omega), \psi_i(\omega))$, $i = 1, 2$ be Lipschitz families of Sobolev functions in $H^{s_0+\bar{\sigma}+\mathtt{b}}(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$, defined for $\omega \in \Omega_o$ satisfying (4.3) with $\mu = \bar{\sigma} + \mathtt{b}$. Then there exists a constant $K_0 > 0$ such that, for $\nu \geq 0$, $\forall \omega \in \Omega_\nu^{\gamma_1}(\mathbf{u}_1) \cap \Omega_\nu^{\gamma_2}(\mathbf{u}_2)$, with $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$,*

$$|\mathcal{R}_\nu(u_1) - \mathcal{R}_\nu(u_2)|_{s_0} \leq K_0 N_{\nu-1}^{-\mathtt{a}}\varepsilon\|u_1 - u_2\|_{s_0+\bar{\sigma}+\mathtt{b}}, \tag{5.21}$$

$$|\mathcal{R}_\nu(u_1) - \mathcal{R}_\nu(u_2)|_{s_0+\mathtt{b}} \leq K_0 N_{\nu-1}\varepsilon\|u_1 - u_2\|_{s_0+\bar{\sigma}+\mathtt{b}}. \tag{5.22}$$

Moreover, for $\nu \geq 1$, $\forall j \in \mathbb{N}$,

$$\left\|\left(\widehat{\mathbf{D}}_j^\nu(u_2) - \widehat{\mathbf{D}}_j^\nu(u_1)\right) - \left(\widehat{\mathbf{D}}_j^{\nu-1}(u_2) - \widehat{\mathbf{D}}_j^{\nu-1}(u_1)\right)\right\|$$
$$\leq |\mathcal{R}_{\nu-1}(u_2) - \mathcal{R}_{\nu-1}(u_1)|_{s_0}, \tag{5.23}$$

$$\|\widehat{\mathbf{D}}_j^\nu(u_2) - \widehat{\mathbf{D}}_j^\nu(u_1)\| \lessdot \varepsilon\|u_1 - u_2\|_{s_0+\bar{\sigma}+\mathtt{b}}. \tag{5.24}$$

$(\mathbf{S4})_\nu$ *Let $\mathbf{u}_1, \mathbf{u}_2$ like in $(\mathbf{S3})_\nu$ and $0 < \rho < \gamma/2$. For all $\nu \geq 0$*

$$\varepsilon K_1 N_{\nu-1}^\tau \|u_1 - u_2\|_{s_0+\bar{\sigma}+\mathtt{b}}^{\sup} \leq \rho \quad \Longrightarrow \quad \Omega_\nu^\gamma(\mathbf{u}_1) \subseteq \Omega_\nu^{\gamma-\rho}(\mathbf{u}_2), \tag{5.25}$$

where K_1 is a suitable constant depending on τ and ν.

5.1. Proof of Theorem 5.1

PROOF OF $(\mathbf{Si})_0$, $i = 1, \ldots, 4$. Properties (5.10)–(5.17) in $(\mathbf{S1})_0$ hold by (5.1)–(5.3) with $\mathbf{D}_j^0$ defined in (5.12) and $\widehat{\mathbf{D}}_j^0(\omega) = 0$ (for (5.17) recall that $N_{-1} := 1$, see (5.5)). Moreover, since m is a real function, $\mathbf{D}_j^0$ is self-adjoint. Then there is nothing else to verify.

$(\mathbf{S2})_0$ holds, since the function $m(\omega) = m(\omega, u(\omega))$ is already defined for all $\omega \in \Omega_o = \Omega_o(\mathbf{u})$.

$(\mathbf{S3})_0$ follows by the estimate (5.7) and by the mean value Theorem, by taking $K_0 > 0$ large enough.

$(\mathbf{S4})_0$ is trivial because, by definition, $\Omega_0^\gamma(\mathbf{u}_1) := \Omega_o =: \Omega_0^{\gamma-\rho}(\mathbf{u}_2)$.

5.2. The reducibility step

We now describe the inductive step, showing how to define a symplectic transformation $\Phi_\nu := \exp(\Psi_\nu)$ so that $\mathcal{L}_{\nu+1} = \Phi_\nu^{-1}\mathcal{L}_\nu\Phi_\nu$ has the desired properties. To simplify notations, in this section we drop the index ν and we write $+$ for $\nu + 1$. At each step of the iteration we have a Hamiltonian operator

$$\mathcal{L} = \omega \cdot \partial_\varphi \mathbb{I}_2 + \mathcal{D} + \mathcal{R} \tag{5.26}$$

where

$$\mathcal{D} := \mathrm{i}\begin{pmatrix} \mathcal{D}^{(1)} & 0 \\ 0 & -\overline{\mathcal{D}}^{(1)} \end{pmatrix}, \qquad \mathcal{D}^{(1)} := \mathrm{diag}_{j \in \mathbb{N}} \mathbf{D}_j, \tag{5.27}$$

$\mathbf{D}_j \in \mathcal{S}(\mathbf{E}_j)$, $\forall j \in \mathbb{N}$ (recall the definiton (2.78)) and $\mathcal{R}$ is a Hamiltonian operator, namely it has the form

$$\mathcal{R} = \mathrm{i} \begin{pmatrix} \mathcal{R}^{(1)} & \mathcal{R}^{(2)} \\ -\overline{\mathcal{R}}^{(2)} & -\overline{\mathcal{R}}^{(1)} \end{pmatrix}, \qquad \mathcal{R}^{(1)} = (\mathcal{R}^{(1)})^*, \qquad \mathcal{R}^{(2)} = (\mathcal{R}^{(2)})^T. \quad (5.28)$$

Let us consider a transformation

$$\Phi := \exp(\Psi), \qquad \Psi := \mathrm{i} \begin{pmatrix} \Psi^{(1)} & \Psi^{(2)} \\ -\overline{\Psi}^{(2)} & -\overline{\Psi}^{(1)} \end{pmatrix}, \qquad (5.29)$$

with $\Psi^{(1)} = (\Psi^{(1)})^*$, $\Psi^{(2)} = (\Psi^{(2)})^T$. Writing

$$\Phi = \mathbb{I}_2 + \Psi + \Phi_{\geq 2}, \quad \Phi_{\geq 2} := \sum_{k \geq 2} \frac{\Psi^k}{k!}, \qquad (5.30)$$

we have

$$\begin{aligned} \mathcal{L}\Phi &= \Phi\left(\omega \cdot \partial_\varphi \mathbb{I}_2 + \mathcal{D}\right) + \left(\omega \cdot \partial_\varphi \Psi + [\mathcal{D}, \Psi] + \Pi_N \mathcal{R}\right) + \Pi_N^\perp \mathcal{R} \\ &\quad + \omega \cdot \partial_\varphi \Phi_{\geq 2} + [\mathcal{D}, \Phi_{\geq 2}] + \mathcal{R}(\Phi - I). \end{aligned} \qquad (5.31)$$

We want to determine the operator Ψ so that

$$\omega \cdot \partial_\varphi \Psi + [\mathcal{D}, \Psi] + \Pi_N \mathcal{R} = [\mathcal{R}], \qquad (5.32)$$

where, recalling the notation (2.65),

$$[\mathcal{R}] := \mathrm{i} \begin{pmatrix} [\mathcal{R}^{(1)}] & 0 \\ 0 & -[\mathcal{R}^{(1)}] \end{pmatrix}, \qquad [\mathcal{R}^{(1)}] := \mathrm{diag}_{j \in \mathbb{N}}(\mathbf{R}^{(1)})_j^j(0). \qquad (5.33)$$

We recall that, according to (2.58), the operator $(\mathbf{R}^{(1)})_j^j(0)$, $j \in \mathbb{N}$ is identified with its 2×2 matrix representation

$$(\mathbf{R}^{(1)})_j^j(0) = \left((\mathcal{R}^{(1)})_k^{k'}(0)\right)_{k,k'=\pm j}.$$

Since $\mathcal{R}^{(1)}$ is self-adjoint, all the 2×2 blocks $(\mathbf{R}^{(1)})_j^j(0)$ are self-adjoint and then also $[\mathcal{R}^{(1)}]$ is self-adjoint.

Lemma 5.1. (Homological equation) *For all $\omega \in \Omega_{\nu+1}^\gamma$ (see (5.13)), there exists a solution Ψ of the homological equation (5.32), which is Hamiltonian and satisfies*

$$|\Psi|D||_s^{\mathrm{Lip}(\gamma)} \ll N^{2\tau+1} \gamma^{-1} |\mathcal{R}|D||_s^{\mathrm{Lip}(\gamma)}. \qquad (5.34)$$

Moreover if $\gamma/2 \leq \gamma_1, \gamma_2 \leq 2\gamma$, and if $\mathbf{u}_i(\omega) = (u_i(\omega), \psi_i(\omega)) \in H^{s_0+\overline{\sigma}+\mathtt{b}}(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$, $i = 1, 2$ are Lipschitz families, then for all $s \in [s_0, s_0 + \mathtt{b}]$, for all $\omega \in \Omega_{\nu+1}^{\gamma_1}(\mathbf{u}_1) \cap \Omega_{\nu+1}^{\gamma_2}(\mathbf{u}_2)$

$$\begin{aligned} |\Psi(u_1) - &\Psi(u_2)|_s \\ &\leq_s N^{2\tau+1}\gamma^{-1}\left(|\mathcal{R}(u_1)|_s\|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}} + |\mathcal{R}(u_1) - \mathcal{R}(u_2)|_s\right). \end{aligned} \qquad (5.35)$$

Proof. Recalling (5.28), (5.29), the Eq. (5.32) is splitted in the two equations

$$i\omega \cdot \partial_\varphi \Psi^{(1)} + [\Psi^{(1)}, \mathcal{D}^{(1)}] + i\Pi_N \mathcal{R}^{(1)} = i[\mathcal{R}^{(1)}], \qquad (5.36)$$

$$i\omega \cdot \partial_\varphi \Psi^{(2)} - (\mathcal{D}^{(1)}\Psi^{(2)} + \Psi^{(2)}\overline{\mathcal{D}}^{(1)}) + i\Pi_N \mathcal{R}^{(2)} = 0. \qquad (5.37)$$

Using the decomposition (2.58), the Eqs. (5.36), (5.37) become, for all $j, j' \in \mathbb{N}$, $\ell \in \mathbb{Z}^\nu$ such that $|\ell| \leq N$,

$$\omega \cdot \ell (\Psi^{(1)})_j^{j'}(\ell) + \mathbf{D}_j (\Psi^{(1)})_j^{j'}(\ell) - (\Psi^{(1)})_j^{j'}(\ell)\mathbf{D}_{j'} = i(\mathbf{R}^{(1)})_j^{j'}(\ell) - i[\mathbf{R}^{(1)}]_j^{j'}, \qquad (5.38)$$

$$\omega \cdot \ell (\Psi^{(2)})_j^{j'}(\ell) + \mathbf{D}_j (\Psi^{(2)})_j^{j'}(\ell) + (\Psi^{(2)})_j^{j'}(\ell)\overline{\mathbf{D}}_{j'} = i(\mathbf{R}^{(2)})_j^{j'}(\ell). \qquad (5.39)$$

By the Definitions (5.14), (5.15), the Eqs. (5.38), (5.39) can be written in the form

$$\mathbf{A}^-(\ell, j, j')(\Psi^{(1)})_j^{j'}(\ell) = i(\mathbf{R}^{(1)})_j^{j'}(\ell) - i[\mathbf{R}^{(1)}]_j^{j'},$$

$$\mathbf{A}^+(\ell, j, j')(\Psi^{(2)})_j^{j'}(\ell) = i(\mathbf{R}^{(2)})_j^{j'}(\ell).$$

Then, since $\omega \in \Omega_{\nu+1}^\gamma$, we can define, $\forall(\ell, j, j') \in \mathbb{Z}^\nu \times \mathbb{N} \times \mathbb{N}, \quad (\ell, j, j') \neq (0, j, j), \quad |\ell| \leq N$,

$$(\Psi^{(1)})_j^{j'}(\ell) = i\mathbf{A}^-(\ell, j, j')^{-1}(\mathbf{R}^{(1)})_j^{j'}(\ell), \qquad (5.40)$$

with the normalization $(\Psi^{(1)})_j^j(0) = 0$, and $\forall(\ell, j, j') \in \mathbb{Z}^\nu \times \mathbb{N} \times \mathbb{N}, \quad |\ell| \leq N$,

$$(\Psi^{(2)})_j^{j'}(\ell) = i\mathbf{A}^+(\ell, j, j')^{-1}(\mathbf{R}^{(2)})_j^{j'}(\ell). \qquad (5.41)$$

Since

$$\|\mathbf{A}^-(\ell, j, j')^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{\langle \ell \rangle^\tau}{\gamma \langle j - j' \rangle}, \quad \|\mathbf{A}^+(\ell, j, j')^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{\langle \ell \rangle^\tau}{\gamma \langle j + j' \rangle},$$

(recall (2.73)) we get immediately that

$$\|(\Psi^{(i)})_j^{j'}(\ell)\| \leq N^\tau \gamma^{-1} \|(\mathbf{R}^{(i)})_j^{j'}(\ell)\|, \quad i = 1, 2. \qquad (5.42)$$

Now, let $\omega_1, \omega_2 \in \Omega_{\nu+1}^\gamma$. As a notation for any function $f = f(\omega)$, we write $\Delta_\omega f := f(\omega_1) - f(\omega_2)$.
By (5.40), one has

$$\Delta_\omega (\Psi^{(1)})_j^{j'}(\ell) = i\left\{\Delta_\omega \mathbf{A}^-(\ell, j, j')^{-1}\right\}(\mathbf{R}^{(1)})_j^{j'}(\ell; \omega_1)$$
$$+ i\mathbf{A}^-(\ell, j, j'; \omega_2)^{-1}\left\{\Delta_\omega (\mathbf{R}^{(1)})_j^{j'}(\ell)\right\}. \qquad (5.43)$$

The second term in the above formula satisfies

$$\|\mathbf{A}^-(\ell, j, j'; \omega_2)^{-1}\left\{\Delta_\omega (\mathbf{R}^{(1)})_j^{j'}(\ell)\right\}\| \leq N^\tau \gamma^{-1}\|\Delta_\omega (\mathbf{R}^{(1)})_j^{j'}(\ell)\|, \qquad (5.44)$$

hence it remains to estimate only the first term in (5.43). We have

$$\Delta_\omega \mathbf{A}^-(\ell, j, j')^{-1} \qquad (5.45)$$
$$= -\mathbf{A}^-(\ell, j, j'; \omega_1)^{-1}\left\{\Delta_\omega \mathbf{A}^-(\ell, j, j')\right\}\mathbf{A}^-(\ell, j, j'; \omega_2)^{-1},$$

therefore

$$\|\Delta_\omega \mathbf{A}^-(\ell, j, j')^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{N^{2\tau}}{\gamma^2 \langle j - j'\rangle^2}\|\Delta_\omega \mathbf{A}^-(\ell, j, j')\|_{\mathrm{Op}(j,j')}. \qquad (5.46)$$

Moreover

$$\Delta_\omega \mathbf{A}^-(\ell, j, j') = (\omega_1 - \omega_2) \cdot \ell\, \mathbf{I}_{j,j'} + M_L(\Delta_\omega \mathbf{D}_j) - M_R(\Delta_\omega \mathbf{D}_{j'}) \qquad (5.47)$$

and using that, by (5.12), (5.16)

$$\mathbf{D}_j(\omega) = m(\omega)\, j\mathbf{I}_j + \widehat{\mathbf{D}}_j(\omega), \quad \text{with} \quad \|\widehat{\mathbf{D}}_j\|^{\mathrm{Lip}(\gamma)} \lessdot \varepsilon j^{-1}, \quad \forall j \in \mathbb{N}, \quad (5.48)$$

we get

$$M_L(\Delta_\omega \mathbf{D}_j) - M_R(\Delta_\omega \mathbf{D}_{j'}) = (\Delta_\omega m)\,(j - j')\mathbf{I}_{j,j'} + M_L(\Delta_\omega \widehat{\mathbf{D}}_j) - M_R(\Delta_\omega \widehat{\mathbf{D}}_{j'}).$$

By (4.36), (5.48) and using the property (2.77) one gets

$$\|M_L(\Delta_\omega \mathbf{D}_j) - M_R(\Delta_\omega \mathbf{D}_{j'})\|_{\mathrm{Op}(j,j')} \lessdot \varepsilon \gamma^{-1} \langle j - j'\rangle|\omega_1 - \omega_2|. \qquad (5.49)$$

Recalling (5.47), we get the estimate

$$\|\Delta_\omega \mathbf{A}^-(\ell, j, j')\|_{\mathrm{Op}(j,j')} \lessdot \left(\langle \ell\rangle + \varepsilon \gamma^{-1}\langle j - j'\rangle\right)|\omega_1 - \omega_2|,$$

which implies, by (5.46), for $\varepsilon \gamma^{-1} \leq 1$ that

$$\|\Delta_\omega \mathbf{A}^-(\ell, j, j')^{-1}\|_{\mathrm{Op}(j,j')} \leq N^{2\tau+1}\gamma^{-2}|\omega_1 - \omega_2|.$$

By (5.43), (5.44), we get the estimate

$$\|\Delta_\omega(\mathbf{\Psi}^{(1)})_j^{j'}(\ell)\| \leq N^\tau \gamma^{-1}\|\Delta_\omega(\mathbf{R}^{(1)})_j^{j'}(\ell)\| \qquad (5.50)$$
$$+ N^{2\tau+1}\gamma^{-2}\|(\mathbf{R}^{(1)})_j^{j'}(\ell; \omega_1)\|.$$

Thus (5.42), (5.50) and the definition (2.80) imply

$$|\Psi^{(1)}|D|\|_s^{\mathrm{Lip}(\gamma)} \lessdot N^{2\tau+1}\gamma^{-1}|\mathcal{R}^{(1)}|D|\|_s^{\mathrm{Lip}(\gamma)}.$$

The estimate of $\Psi^{(2)}$ in terms of $\mathcal{R}^{(2)}$ follows by similar arguments and then (5.34) follows.

Now we prove the estimate (5.35). As a notation, we write $\Delta_{12}A := A(u_1) - A(u_2)$, for any operator A depending on u. We prove the estimate (5.35) for the operator $\Psi^{(1)}$. The estimate for $\Psi^{(2)}$ is analogous. By (5.40), for all $j, j' \in \mathbb{N}$, $\ell \in \mathbb{Z}^\nu$, $(\ell, j, j') \neq (0, j, j)$, $|\ell| \leq N$ one has

$$\Delta_{12}(\mathbf{\Psi}^{(1)})_j^{j'}(\ell) = \mathrm{i}\left\{\Delta_{12}\mathbf{A}^-(\ell, j, j')^{-1}\right\}(\mathbf{R}^{(1)})_j^{j'}(\ell; u_1) \qquad (5.51)$$
$$+ \mathrm{i}\mathbf{A}^-(\ell, j, j'; u_2)^{-1}\left\{\Delta_{12}(\mathbf{R}^{(1)})_j^{j'}(\ell)\right\}.$$

Since $\omega \in \Omega_{\nu+1}^{\gamma_1}(\mathbf{u}_1) \cap \Omega_{\nu+1}^{\gamma_2}(\mathbf{u}_2)$ and $\gamma/2 \leq \gamma_1, \gamma_2 \leq 2\gamma$, we have

$$\|\mathbf{A}^-(\ell, j, j'; u_2)^{-1}\left\{\Delta_{12}(\mathbf{R}^{(1)})_j^{j'}(\ell)\right\}\| \leq N^\tau \gamma^{-1}\|\Delta_{12}(\mathbf{R}^{(1)})_j^{j'}(\ell)\|. \qquad (5.52)$$

Moreover, arguing as in (5.45), (5.46) (replacing ω_1 resp. ω_2 by u_1 resp. u_2), one has

$$\|\Delta_{12}\mathbf{A}^-(\ell, j, j')^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{N^{2\tau}}{\gamma^2 \langle j - j'\rangle^2}\|\Delta_{12}\mathbf{A}^-(\ell, j, j')\|_{\mathrm{Op}(j,j')}. \qquad (5.53)$$

By the definition (5.14), we get

$$\Delta_{12}\mathbf{A}^-(\ell, j, j') = M_L(\Delta_{12}\mathbf{D}_j) - M_R(\Delta_{12}\mathbf{D}_{j'})$$

$$\overset{(5.48)}{=} (\Delta_{12}m)(j - j')\,\mathbf{I}_{j,j'} + M_L(\Delta_{12}\widehat{\mathbf{D}}_j) - M_R(\Delta_{12}\widehat{\mathbf{D}}_{j'}), \tag{5.54}$$

therefore by (4.36), (2.77), (5.24)

$$\|\Delta_{12}\mathbf{A}^-(\ell, j, j')\|_{\mathrm{Op}(j,j')} \lessdot \varepsilon\langle j - j'\rangle \|u_1 - u_2\|_{s_0 + \overline{\sigma} + \mathtt{b}}. \tag{5.55}$$

Then (5.53), (5.55), $\varepsilon\gamma^{-1} \leq 1$ imply that

$$\| \left\{ \Delta_{12}\mathbf{A}^-(\ell, j, j')^{-1} \right\} (\mathbf{R}^{(1)})_j^{j'}(\ell; u_1)\|$$

$$\lessdot N^{2\tau}\gamma^{-1}\|(\mathbf{R}^{(1)})_j^{j'}(\ell; u_1)\|\|u_1 - u_2\|_{s_0 + \overline{\sigma} + \mathtt{b}}$$

and recalling (5.51), (5.52), we obtain the estimate

$$\|\Delta_{12}(\mathbf{\Psi}^{(1)})_j^{j'}(\ell)\|$$

$$\lessdot N^{2\tau}\gamma^{-1} \left(\|(\mathbf{R}^{(1)})_j^{j'}(\ell; u_1)\|\|u_1 - u_2\|_{s_0 + \overline{\sigma} + \mathtt{b}} + \|\Delta_{12}(\mathbf{R}^{(1)})_j^{j'}(\ell)\| \right).$$

This last estimate imply the estimate (5.35) for $\Delta_{12}\Psi^{(1)}$, by using the definition of the norm $|\cdot|_s$ in (2.80). The estimate for $\Delta_{12}\Psi^{(2)}$ follows by similar arguments and then the proof is concluded. $\qquad\square$

By (5.31), (5.32), (5.33), we get

$$\mathcal{L}_+ := \Phi^{-1}\mathcal{L}\Phi = \omega \cdot \partial_\varphi \mathbb{I}_2 + \mathcal{D}_+ + \mathcal{R}_+, \tag{5.56}$$

where

$$\mathcal{D}_+ := \mathcal{D} + [\mathcal{R}],$$

$$\mathcal{R}_+ := (\Phi^{-1} - \mathbb{I}_2)\mathcal{R} + \Phi^{-1}\left(\Pi_N^\perp \mathcal{R} + \omega \cdot \partial_\varphi \Phi_{\geq 2} + [\mathcal{D}, \Phi_{\geq 2}] + \mathcal{R}(\Phi - \mathbb{I}_2)\right).$$

Lemma 5.2. *[**The new** 2×2 **block-diagonal part**] The new block-diagonal part is*

$$\mathcal{D}_+ := \mathcal{D} + [\mathcal{R}] = \mathrm{i}\begin{pmatrix} \mathcal{D}_+^{(1)} & 0 \\ 0 & -\overline{\mathcal{D}}_+^{(1)} \end{pmatrix}, \qquad \mathcal{D}_+^{(1)} := \mathcal{D}^{(1)} + [\mathcal{R}^{(1)}] = \mathrm{diag}_{j \in \mathbb{N}}\mathbf{D}_j^+,$$

where

$$\mathbf{D}_j^+ := \mathbf{D}_j + (\mathbf{R}^{(1)})_j^j(0) = m\,j\mathbf{I}_j + \widehat{\mathbf{D}}_j + (\mathbf{R}^{(1)})_j^j(0) = m\,j\mathbf{I}_j + \widehat{\mathbf{D}}_j^+,$$

$$\widehat{\mathbf{D}}_j^+ := \widehat{\mathbf{D}}_j + (\mathbf{R}^{(1)})_j^j(0), \qquad \forall j \in \mathbb{N}, \tag{5.57}$$

and

$$\|\mathbf{D}_j^+ - \mathbf{D}_j\|^{\mathrm{Lip}(\gamma)} \lessdot j^{-1}|\mathcal{R}|D|\|_{s_0}^{\mathrm{Lip}(\gamma)}. \tag{5.58}$$

Moreover, if $\mathbf{u}_i(\omega) = (u_i(\omega), \psi_i(\omega))$, $i = 1, 2$ *are families of Sobolev functions, for all* $\omega \in \Omega_\nu^{\gamma_1}(\mathbf{u}_1) \cap \Omega_\nu^{\gamma_2}(\mathbf{u}_2)$, $\gamma/2 \leq \gamma_1, \gamma_2 \leq 2\gamma$, *for all* $j \in \mathbb{N}$,

$$\| \left(\widehat{\mathbf{D}}_j^+(u_1) - \widehat{\mathbf{D}}_j^+(u_2)\right) - \left(\widehat{\mathbf{D}}_j(u_1) - \widehat{\mathbf{D}}_j(u_2)\right) \|_{L^2} \lessdot |\mathcal{R}(u_1) - \mathcal{R}(u_2)|_{s_0}. \tag{5.59}$$

Proof. Notice that, since $\mathcal{R}^{(1)}(\varphi)$ is selfadjoint, the operators $(\mathbf{R}^{(1)})_j^j(0)$: $\mathbf{E}_j \to \mathbf{E}_j$ are self-adjoint for all $j \in \mathbb{N}$, i.e. $(\mathbf{R}^{(1)})_j^j(0) \in \mathcal{S}(\mathbf{E}_j)$. Since $\mathbf{D}_j, \widehat{\mathbf{D}}_j$ are self-adjoint, we get that $\mathbf{D}_j^+, \widehat{\mathbf{D}}_j^+$ are self-adjoint for all $j \in \mathbb{N}$. Furthermore, by Lemma 2.6

$$\|\mathbf{D}_j^+ - \mathbf{D}_j\|^{\mathrm{Lip}(\gamma)} = \|\widehat{\mathbf{D}}_j^+ - \widehat{\mathbf{D}}_j\|^{\mathrm{Lip}(\gamma)} = \|(\mathbf{R}^{(1)})_j^j(0)\|^{\mathrm{Lip}(\gamma)}$$
$$\leq j^{-1}\sup_{k\in\mathbb{N}}\|(\mathbf{R}^{(1)})_k^k(0)k\|^{\mathrm{Lip}(\gamma)} \leq j^{-1}|\mathcal{R}|D|\|_{s_0}^{\mathrm{Lip}(\gamma)},$$

which is the estimate (5.58).

Since, by (5.57) we have $\left(\widehat{\mathbf{D}}_j^+(u_1) - \widehat{\mathbf{D}}_j^+(u_2)\right) - \left(\widehat{\mathbf{D}}_j(u_1) - \widehat{\mathbf{D}}_j(u_2)\right) = (\mathbf{R}^{(1)})_j^j(0;u_1) - (\mathbf{R}^{(1)})_j^j(0;u_2)$, the estimate (5.59) follows since, applying again Lemma 2.6, for all $j \in \mathbb{N}$

$$\|(\mathbf{R}^{(1)})_j^j(0;u_1) - (\mathbf{R}^{(1)})_j^j(0;u_2)\| \prec |\mathcal{R}^{(1)}(u_1) - \mathcal{R}^{(1)}(u_2)|_{s_0}.$$

$\square$

5.3. The iteration

Let $\nu \geq 0$ and let us suppose that $(\mathbf{Si})_\nu$ are true. We prove $(\mathbf{Si})_{\nu+1}$. To simplify notations, in this proof we write $|\cdot|_s$ for $|\cdot|_s^{\mathrm{Lip}(\gamma)}$.

PROOF OF $(\mathbf{S1})_{\nu+1}$. Since the self-adjoint 2×2 blocks $\mathbf{D}_j^\nu \in \mathcal{S}(\mathbf{E}_j)$ are defined on Ω_ν^γ, the set $\Omega_{\nu+1}^\gamma$ is well-defined and by Lemma 5.1, the following estimates hold on $\Omega_{\nu+1}^\gamma$

$$|\Psi_\nu|D|\|_s \leq_s N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|D|\|_s \overset{(5.17)}{\leq_s} N_\nu^{2\tau+1}N_{\nu-1}^{-\mathtt{a}}\gamma^{-1}|\mathcal{R}_0|D|\|_{s+\mathtt{b}}, \quad (5.60)$$

and in particular, by (5.7), (5.9), (5.6), (5.5), taking δ_0 small enough,

$$|\Psi_\nu|D|\|_{s_0} \leq 1. \quad (5.61)$$

By (5.61), we can apply Lemma 2.10 to the map $\Phi_\nu^{\pm 1} := \exp(\pm\Psi_\nu)$ and using also Lemma 2.6-(ii) we obtain that

$$|\Phi_\nu^{\pm 1} - \mathbb{I}_2|_s \leq |(\Phi_\nu^{\pm 1} - \mathbb{I}_2)|D|\|_s \leq_s |\Psi_\nu|D|\|_s. \quad (5.62)$$

By (5.56) we get $\mathcal{L}_{\nu+1} := \Phi_\nu^{-1}\mathcal{L}_\nu\Phi_\nu = \omega \cdot \partial_\varphi \mathbb{I}_2 + \mathcal{D}_{\nu+1} + \mathcal{R}_{\nu+1}$, where $\mathcal{D}_{\nu+1} := \mathcal{D}_\nu + [\mathcal{R}_\nu]$ and

$$\mathcal{R}_{\nu+1} := (\Phi_\nu^{-1} - \mathbb{I}_2)[\mathcal{R}_\nu]$$
$$+ \Phi_\nu^{-1}\left(\Pi_{N_\nu}^\perp \mathcal{R}_\nu + \omega \cdot \partial_\varphi \Psi_{\nu,\geq 2} + [\mathcal{D}_\nu, \Psi_{\nu,\geq 2}] + \mathcal{R}_\nu(\Phi_\nu - \mathbb{I}_2)\right). \quad (5.63)$$

Note that, since $\mathcal{R}_\nu$ is defined on Ω_ν^γ and Ψ_ν is defined on $\Omega_{\nu+1}^\gamma$, the remainder $\mathcal{R}_{\nu+1}$ is defined on $\Omega_{\nu+1}^\gamma$ too. Since the remainder $\mathcal{R}_\nu : \mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1})$ is Hamiltonian, the map $\Psi_\nu : \mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1})$ is Hamiltonian, then $\Phi_\nu : \mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1})$ is symplectic and the operator $\mathcal{L}_{\nu+1} : \mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^{s-1}(\mathbb{T}^{\nu+1})$ is still Hamiltonian.

Now let us prove the estimates (5.17) for $\mathcal{R}_{\nu+1}$. Applying Lemmata 2.6, 2.7 and the estimates (5.61), (5.60), (5.62), we get

$$|(\Phi_\nu^{-1} - \mathbb{I}_2)[\mathcal{R}_\nu]|D|\|_s, |\Phi_\nu^{-1}\mathcal{R}_\nu(\Phi_\nu - \mathbb{I}_2)|D|\|_s$$
$$\leq_s N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|D|\|_s|\mathcal{R}_\nu|D|\|_{s_0} \quad (5.64)$$

and

$$|\Phi_\nu^{-1}\Pi_{N_\nu}^\perp \mathcal{R}_\nu |D|||_s \leq_s |\Pi_{N_\nu}\mathcal{R}_\nu |D|||_s + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|D|||_s|\mathcal{R}_\nu|D|||_{s_0}. \quad (5.65)$$

Then, it remains to estimate the term $\Phi_\nu^{-1}(\omega\cdot\partial_\varphi\Phi_{\nu,\geq 2} + [\mathcal{D}_\nu,\Phi_{\nu,\geq 2}])|D|$ in (5.63). A direct calculation shows that for all $n\geq 2$

$$\omega\cdot\partial_\varphi(\Psi_\nu^n) + [\mathcal{D}_\nu,\Psi_\nu^n] = \sum_{i+k=n-1}\Psi_\nu^i(\omega\cdot\partial_\varphi\Psi_\nu + [\mathcal{D}_\nu,\Psi_\nu])\Psi_\nu^k$$

$$\overset{(5.32)}{=}\sum_{i+k=n-1}\Psi_\nu^i([\mathcal{R}_\nu] - \mathcal{R}_\nu)\Psi_\nu^k, \quad (5.66)$$

therefore using (5.61), (5.60), (2.89) and that $|\Psi_\nu|_s \overset{Lemma\,2.6-(ii)}{\leq} |\Psi_\nu|D|||_s$ we get that for all $n\geq 2$

$$|(\omega\cdot\partial_\varphi(\Psi_\nu^n) + [\mathcal{D}_\nu,\Psi_\nu^n])|D|||_s$$
$$\leq n^2 C(s)^n \left((|\Psi_\nu|D|||_{s_0})^{n-1}|\mathcal{R}_\nu|D|||_s + (|\Psi_\nu|D|||_{s_0})^{n-2}|\Psi_\nu|D|||_s|\mathcal{R}_\nu|D|||_{s_0}\right)$$
$$\leq 2n^2 C(s)^n N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|D|||_s|\mathcal{R}_\nu|D|||_{s_0}, \quad (5.67)$$

for some constant $C(s) > 0$. Thus

$$|(\omega\cdot\partial_\varphi\Psi_{\nu,\geq 2} + [\mathcal{D}_\nu,\Psi_{\nu,\geq 2}])|D|||_s \leq \sum_{n\geq 2}\frac{1}{n!}|(\omega\cdot\partial_\varphi(\Psi_\nu^n) + [\mathcal{D}_\nu,\Psi_\nu^n])|D|||_s$$

$$\overset{(5.67)}{\leq} N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|D|||_s|\mathcal{R}_\nu|D|||_{s_0}2\sum_{n\geq 2}\frac{C(s)^n n^2}{n!}$$

$$\leq_s N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|D|||_s|\mathcal{R}_\nu|D|||_{s_0}. \quad (5.68)$$

The estimates (5.62), (5.68) and Lemma 2.7 imply that

$$\left|\Phi_\nu^{-1}(\omega\cdot\partial_\varphi\Psi_{\nu,\geq 2} + [\mathcal{D}_\nu,\Psi_{\nu,\geq 2}])|D|\right|_s$$
$$\leq_s N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|D|||_s|\mathcal{R}_\nu|D|||_{s_0}. \quad (5.69)$$

Collecting the estimates (5.64)–(5.69) we obtain the estimate

$$|\mathcal{R}_{\nu+1}|D|||_s \leq_s |\Pi_{N_\nu}\mathcal{R}_\nu|D|||_s + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|D|||_s|\mathcal{R}_\nu|D|||_{s_0}, \quad (5.70)$$

which implies (using the smoothing property (2.100) and (5.9), (5.17))

$$|\mathcal{R}_{\nu+1}|D|||_s \leq_s N_\nu^{-\mathtt{b}}|\mathcal{R}_\nu|D|||_{s+\mathtt{b}} + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|D|||_s|\mathcal{R}_\nu|D|||_{s_0}, \quad (5.71)$$
$$|\mathcal{R}_{\nu+1}|D|||_{s+\mathtt{b}} \leq C(s+\mathtt{b})|\mathcal{R}_\nu|D|||_{s+\mathtt{b}}. \quad (5.72)$$

Hence

$$|\mathcal{R}_{\nu+1}|D|||_{s+\mathtt{b}} \overset{(5.72),(5.17)}{\leq} C(s+\mathtt{b})|\mathcal{R}_0|D|||_{s+\mathtt{b}}N_{\nu-1} \leq |\mathcal{R}_0|D|||_{s+\mathtt{b}}N_\nu,$$

for $N_0 := N_0(s,\mathtt{b}) > 0$ large enough and then the second inequality in (5.17) for $\mathcal{R}_{\nu+1}$ has been proved. Let us prove the first inequality in (5.17) at the step $\nu + 1$. We have

$$|\mathcal{R}_{\nu+1}|D||_s$$

$$\overset{(5.72),(5.17)}{\leq_s} \quad N_\nu^{-\mathtt{b}}N_{\nu-1}|\mathcal{R}_0|D||_{s+\mathtt{b}} + N_\nu^{2\tau+1}N_{\nu-1}^{-2\mathtt{a}}\gamma^{-1}|\mathcal{R}_0|D||_{s_0+\mathtt{b}}|\mathcal{R}_0|D||_{s+\mathtt{b}}$$

$$\leq |\mathcal{R}_0|D||_{s+\mathtt{b}}N_\nu^{-\mathtt{a}},$$

provided

$$N_\nu^{\mathtt{b}-\mathtt{a}}N_{\nu-1}^{-1} \geq 2C(s), \quad \gamma^{-1}|\mathcal{R}_0|D||_{s_0+\mathtt{b}} \leq \frac{N_{\nu-1}^{2\mathtt{a}}N_\nu^{-\mathtt{a}-2\tau-1}}{2C(s)},$$

which are verified by (5.5), (5.6), (5.7) and (5.9), taking N_0 large enough and δ_0 small enough.

The estimate (5.16) for $\widehat{\mathbf{D}}_j^{\nu+1}$ follows by a telescopic argument, since $\widehat{\mathbf{D}}_j^{\nu+1} = \sum_{k=0}^\nu \widehat{\mathbf{D}}_j^{k+1} - \widehat{\mathbf{D}}_j^k$ and applying the estimates (5.58), (5.17), (5.7).

PROOF OF $(\mathbf{S2})_{\nu+1}$ We now construct a Lipschitz extension of the function $\omega \in \Omega_{\nu+1}^\gamma \mapsto \mathbf{D}_j^{\nu+1}(\omega) \in \mathcal{S}(\mathbf{E}_j)$, for all $j \in \mathbb{N}$. We apply Lemma M.5 in [36]. Note that the space $\mathcal{S}(\mathbf{E}_j)$, defined in (2.78), is a Hilbert subspace of $\mathcal{L}(\mathbf{E}_j)$ equipped by the scalar product defined in (2.71). By the inductive hypothesis, there exists a Lipschitz function $\widetilde{\mathbf{D}}_j^\nu : \Omega_o \to \mathcal{S}(\mathbf{E}_j)$, satisfying $\widetilde{\mathbf{D}}_j^\nu(\omega) = \mathbf{D}_j^\nu(\omega)$, for all $\omega \in \Omega_\nu^\gamma$. Now we construct a self-adjoint extension of the self-adjoint operator $\mathbf{D}_j^{\nu+1} = \mathbf{D}_j^\nu + [\mathbf{D}]_j^\nu$, where $[\mathbf{D}]_j^\nu := (\mathbf{R}_\nu^{(1)})_j^j(0)$. By (5.58), for all $j \in \mathbb{N}$, one has that

$$\|[\mathbf{D}]_j^\nu\|^{\mathrm{Lip}(\gamma)}$$

$$= \|\mathbf{D}_j^{\nu+1} - \mathbf{D}_j^\nu\|^{\mathrm{Lip}(\gamma)} \lessdot j^{-1}|\mathcal{R}_\nu|D||_{s_0}^{\mathrm{Lip}(\gamma)} \overset{(5.17)}{\lessdot} N_{\nu-1}^{-\mathtt{a}}|\mathcal{R}_0|D||_{s_0+\mathtt{b}}j^{-1}$$

$$\overset{(5.7)}{\lessdot} N_{\nu-1}^{-\mathtt{a}}\varepsilon j^{-1}$$

and then by Lemma M.5 in [36] there exists a Lipschitz extension $[\widetilde{\mathbf{D}}]_j^\nu : \Omega_o \to \mathcal{S}(\mathbf{E}_j)$ of $[\mathbf{D}]_j^\nu : \Omega_\nu^\gamma \to \mathcal{S}(\mathbf{E}_j)$ still satisfying the above estimate. Therefore we define $\widetilde{\mathbf{D}}_j^{\nu+1} := \widetilde{\mathbf{D}}_j^\nu + [\widetilde{\mathbf{D}}]_j^\nu$.

PROOF OF $(\mathbf{S3})_{\nu+1}$. As a notation we write $\Delta_{12}A := A(u_1) - A(u_2)$, for any operator $A = A(u)$ depending on u. Now we will estimate the operator $\Delta_{12}\mathcal{R}_{\nu+1}$, where $\mathcal{R}_{\nu+1}$ is defined in (5.63). Moreover, we define

$$R_\nu(s) := \max\{|\mathcal{R}_\nu(u_1)|_s, |\mathcal{R}_\nu(u_2)|_s\}, \qquad \forall s \in [s_0, s_0 + \mathtt{b}].$$

Note that, by (5.17) and (5.7) and by Lemma 2.6-(ii), one gets

$$R_\nu(s_0) \lessdot \varepsilon N_{\nu-1}^{-\mathtt{a}}, \qquad R_\nu(s_0 + \mathtt{b}) \lessdot \varepsilon N_{\nu-1}. \tag{5.73}$$

By (5.34), (5.35), (5.73), (5.21), (5.22), Lemma 2.6-(ii), one has

$$|\Delta_{12}\Psi_\nu|_{s_0} \lessdot N_\nu^{2\tau+1}N_{\nu-1}^{-\mathtt{a}}\varepsilon\gamma^{-1}\|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}}, \tag{5.74}$$

$$|\Delta_{12}\Psi_\nu|_{s_0+\mathtt{b}} \lessdot N_\nu^{2\tau+1}N_{\nu-1}\varepsilon\gamma^{-1}\|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}}, \tag{5.75}$$

$$|\Psi_\nu(u_1)|_{s_0}, |\Psi_\nu(u_2)|_{s_0} \lessdot N_\nu^{2\tau+1}N_{\nu-1}^{-\mathtt{a}}\varepsilon\gamma^{-1}, \tag{5.76}$$

$$|\Psi_\nu(u_1)|_{s_0+\mathtt{b}}, |\Psi_\nu(u_2)|_{s_0+\mathtt{b}} \lessdot N_\nu^{2\tau+1}N_{\nu-1}\varepsilon\gamma^{-1}. \tag{5.77}$$

By the estimates (5.62) (applied to $\Phi_\nu = \Phi_\nu(u_i)$, $i = 1, 2$), (5.76)–(5.77) and using also (2.91), one gets

$$|\Phi_\nu^{\pm 1}(u_i) - \mathbb{I}_2|_{s_0} \lessdot N_\nu^{2\tau+1} N_{\nu-1}^{-\mathsf{a}} \varepsilon \gamma^{-1}, \tag{5.78}$$

$$|\Phi_\nu^{\pm 1}(u_i) - \mathbb{I}_2|_{s_0+\mathsf{b}} \lessdot N_\nu^{2\tau+1} N_{\nu-1} \varepsilon \gamma^{-1}, \tag{5.79}$$

$$|\Delta_{12}\Phi_\nu^{\pm 1}|_{s_0} \lessdot N_\nu^{2\tau+1} N_{\nu-1}^{-\mathsf{a}} \varepsilon \gamma^{-1} \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}}, \tag{5.80}$$

$$|\Delta_{12}\Phi_\nu^{\pm 1}|_{s_0+\mathsf{b}} \lessdot N_\nu^{2\tau+1} N_{\nu-1} \varepsilon \gamma^{-1} \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}}. \tag{5.81}$$

We estimate separately the terms of $\Delta_{12}\mathcal{R}_{\nu+1}$, where, by (5.63)

$$\mathcal{R}_{\nu+1} = (\Phi_\nu^{-1} - \mathbb{I}_2)[\mathcal{R}_\nu] + \Phi_\nu^{-1}\mathcal{H}_\nu,$$
$$\mathcal{H}_\nu := \Pi_{N_\nu}^\perp \mathcal{R}_\nu + \omega \cdot \partial_\varphi \Psi_{\nu,\geq 2} + [\mathcal{D}_\nu, \Psi_{\nu,\geq 2}] + \mathcal{R}_\nu(\Phi_\nu - \mathbb{I}_2). \tag{5.82}$$

In the following we will use that by (5.5), (5.6), (5.9) (choosing δ_0 small enough)

$$N_\nu^{2\tau+1} N_{\nu-1}^{-\mathsf{a}} \varepsilon \gamma^{-1} \leq 1, \quad \forall \nu \geq 0. \tag{5.83}$$

Lemma 2.6-(iii), Lemma 2.7 and the estimates (5.78)–(5.81), (5.73), (5.21), (5.22), (5.83) imply that

$$|\Delta_{12}\left\{(\Phi_\nu^{-1} - \mathbb{I}_2)[\mathcal{R}_\nu]\right\}|_{s_0}, \; |\Delta_{12}\left\{\mathcal{R}_\nu(\Phi_\nu - \mathbb{I}_2)\right\}|_{s_0} \tag{5.84}$$
$$\lessdot N_\nu^{2\tau+1} N_{\nu-1}^{-2\mathsf{a}} \varepsilon^2 \gamma^{-1} \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}},$$

$$|\Delta_{12}\left\{(\Phi_\nu^{-1} - \mathbb{I}_2)[\mathcal{R}_\nu]\right\}|_{s_0+\mathsf{b}}, \; |\Delta_{12}\left\{\mathcal{R}_\nu(\Phi_\nu - \mathbb{I}_2)\right\}|_{s_0+\mathsf{b}} \tag{5.85}$$
$$\lessdot N_{\nu-1} \varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}}.$$

Moreover by (2.100), (5.21), (5.22) one gets

$$|\Pi_{N_\nu}^\perp \Delta_{12} \mathcal{R}_\nu|_{s_0} \lessdot N_\nu^{-\mathsf{b}} N_{\nu-1} \varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}}, \tag{5.86}$$

$$|\Pi_{N_\nu}^\perp \Delta_{12} \mathcal{R}_\nu|_{s_0+\mathsf{b}} \lessdot N_{\nu-1} \varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}}. \tag{5.87}$$

It remains to estimate only the term $\Delta_{12}\left\{\omega \cdot \partial_\varphi \Phi_{\nu,\geq 2} + [\mathcal{D}_\nu, \Phi_{\nu,\geq 2}]\right\}$, where we recall that $\Phi_{\nu,\geq 2} = \sum_{n\geq 2} \frac{\Psi_\nu^n}{n!}$. By (5.66), for all $n \geq 2$ we have

$$\Delta_{12}\left\{\omega \cdot \partial_\varphi \Psi_\nu^n + [\mathcal{D}_\nu, \Psi_\nu^n]\right\} = \sum_{i+k=n-1} \Delta_{12}\left\{\Psi_\nu^i([\mathcal{R}_\nu] - \mathcal{R}_\nu)\Psi_\nu^k\right\}. \tag{5.88}$$

Iterating the interpolation estimate of Lemma 2.7 and using (5.21), (5.22), (5.73), (5.74)–(5.77), we have that for all $i + k = n - 1$,

$$|\Delta_{12}\left\{\Psi_\nu^i([\mathcal{R}_\nu] - \mathcal{R}_\nu)\Psi_\nu^k\right\}|_{s_0}$$
$$\leq nC(s_0)^n \left(N_\nu^{2\tau+1} N_{\nu-1}^{-\mathsf{a}} \varepsilon \gamma^{-1}\right)^{n-2} N_\nu^{2\tau+1} N_{\nu-1}^{-2\mathsf{a}} \varepsilon^2 \gamma^{-1} \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}}$$
$$\overset{(5.83)}{\leq} nC(s_0)^n N_\nu^{2\tau+1} N_{\nu-1}^{-2\mathsf{a}} \varepsilon^2 \gamma^{-1} \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}} \tag{5.89}$$

and

$$|\Delta_{12}\left\{\Psi_\nu^i([\mathcal{R}_\nu] - \mathcal{R}_\nu)\Psi_\nu^k\right\}|_{s_0+\mathsf{b}}$$
$$\leq nC(s_0+\mathsf{b})^n \left(N_\nu^{2\tau+1} N_{\nu-1}^{-\mathsf{a}} \varepsilon \gamma^{-1}\right)^{n-1} N_{\nu-1} \varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}}$$
$$\overset{(5.83)}{\leq} nC(s_0+\mathsf{b})^n N_{\nu-1} \varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathsf{b}}. \tag{5.90}$$

Hence

$$
|\omega \cdot \partial_\varphi \Phi_{\nu,\geq 2} + [\mathcal{D}_\nu, \Phi_{\nu,\geq 2}]|_{s_0} \overset{(5.88)}{\leq} \sum_{n\geq 2} \frac{1}{n!} \sum_{i+k=n-1} \left| \Delta_{12} \left\{ \Psi_\nu^i ([\mathcal{R}_\nu] - \mathcal{R}_\nu)\Psi_\nu^k \right\} \right|_{s_0}
$$

$$
\overset{(5.89)}{\leq} N_\nu^{2\tau+1} N_{\nu-1}^{-2\mathtt{a}} \varepsilon^2 \gamma^{-1} \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}} \sum_{n\geq 2} \frac{n^2 C(s_0)^n}{n!}
$$

$$
\lll N_\nu^{2\tau+1} N_{\nu-1}^{-2\mathtt{a}} \varepsilon^2 \gamma^{-1} \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}} \tag{5.91}
$$

and

$$
|\omega \cdot \partial_\varphi \Phi_{\nu,\geq 2} + [\mathcal{D}_\nu, \Phi_{\nu,\geq 2}]|_{s_0+\mathtt{b}}
$$

$$
\overset{(5.88)}{\leq} \sum_{n\geq 2} \frac{1}{n!} \sum_{i+k=n-1} \left| \Delta_{12} \left\{ \Psi_\nu^i ([\mathcal{R}_\nu] - \mathcal{R}_\nu)\Psi_\nu^k \right\} \right|_{s_0+\mathtt{b}}
$$

$$
\overset{(5.90)}{\leq} N_{\nu-1}\varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}} \sum_{n\geq 2} \frac{n^2 C(s_0+\mathtt{b})^n}{n!}
$$

$$
\lll N_{\nu-1}\varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}}. \tag{5.92}
$$

Collecting the estimates (5.86), (5.87), (5.91), (5.92) and recalling the definition of $\mathcal{H}_\nu$ in (5.82), one gets

$$
|\Delta_{12}\mathcal{H}_\nu|_{s_0} \lll \left(N_\nu^{-\mathtt{b}} N_{\nu-1}\varepsilon + N_\nu^{2\tau+1} N_{\nu-1}^{-2\mathtt{a}} \varepsilon^2 \gamma^{-1} \right) \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}} \tag{5.93}
$$

$$
|\Delta_{12}\mathcal{H}_\nu|_{s_0+\mathtt{b}} \lll N_{\nu-1}\varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}}. \tag{5.94}
$$

Arguing as in the proof of (5.17) one can obtain that for $i = 1, 2$

$$
|\mathcal{H}_\nu(u_i)|_{s_0} \lll N_\nu^{-\mathtt{b}} N_{\nu-1}\varepsilon + N_\nu^{2\tau+1} N_{\nu-1}^{-2\mathtt{a}} \varepsilon^2 \gamma^{-1}, \tag{5.95}
$$

$$
|\mathcal{H}_\nu(u_i)|_{s_0+\mathtt{b}} \lll N_{\nu-1}\varepsilon, \tag{5.96}
$$

thus, by (5.82), writing $\Phi_\nu^{-1}\mathcal{H}_\nu = \mathcal{H}_\nu + (\Phi_\nu^{-1} - \mathbb{I}_2)\mathcal{H}_\nu$, using Lemma 2.7, and the estimates (5.78)–(5.81), (5.83)–(5.85), (5.93)–(5.96), one obtains

$$
|\Delta_{12}\mathcal{R}_{\nu+1}|_{s_0} \leq C(\tau,\nu) \left(N_\nu^{-\mathtt{b}} N_{\nu-1}\varepsilon + N_\nu^{2\tau+1} N_{\nu-1}^{-2\mathtt{a}} \varepsilon^2 \gamma^{-1} \right) \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}}
$$

$$
\leq K_0 N_\nu^{-\mathtt{a}} \varepsilon,
$$

and

$$
|\Delta_{12}\mathcal{R}_{\nu+1}|_{s_0+\mathtt{b}} \leq C(\tau,\nu) N_{\nu-1}\varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}} \leq K_0 N_\nu \varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}}
$$

by (5.5), (5.6), taking N_0 large enough and $\varepsilon\gamma^{-1}$ small enough. Then the estimate (5.21), (5.22) has been proved at the step $\nu+1$. The estimates (5.23), (5.24) follow by (5.59), (5.21) and by a telescopic argument.

PROOF OF $(\mathbf{S4})_{\nu+1}$ We have to prove that, if

$$
K_1 N_\nu^\tau \varepsilon \|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}} \leq \rho, \tag{5.97}
$$

for a suitable constant $K_1 = K_1(\tau,\nu) > 0$, then

$$
\omega \in \Omega_{\nu+1}^\gamma(\mathbf{u}_1) \qquad \Longrightarrow \qquad \omega \in \Omega_{\nu+1}^{\gamma-\rho}(\mathbf{u}_2).
$$

By the definiton (5.13), we have $\Omega_{\nu+1}^{\gamma}(\mathbf{u}_1) \subseteq \Omega_{\nu}^{\gamma}(\mathbf{u}_1)$, and by the inductive hypothesis $\Omega_{\nu}^{\gamma}(\mathbf{u}_1) \subseteq \Omega_{\nu}^{\gamma-\rho}(\mathbf{u}_2)$, hence

$$\omega \in \Omega_{\nu+1}^{\gamma}(\mathbf{u}_1) \qquad \Longrightarrow \qquad \omega \in \Omega_{\nu}^{\gamma-\rho}(\mathbf{u}_2) \subseteq \Omega_{\nu}^{\gamma/2}(\mathbf{u}_2). \qquad (5.98)$$

Then for all $j \in \mathbb{N}$, the 2×2 matrices $\mathbf{D}_j^{\nu}(u_2) = \mathbf{D}_j^{\nu}(\omega, u_2(\omega))$ are well defined on $\Omega_{\nu+1}^{\gamma}(\mathbf{u}_1)$. We set for convenience

$$\Delta_{12}\mathbf{A}_{\nu}^{-}(\ell, j, j') := \mathbf{A}_{\nu}^{-}(\ell, j, j'; u_2) - \mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1).$$

By (5.98), on the set $\Omega_{\nu+1}^{\gamma}(\mathbf{u}_1)$, both the operators $\mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1)$ and $\mathbf{A}_{\nu}^{-}(\ell, j, j'; u_2)$ are well defined. By the estimate (5.55),

$$\|\Delta_{12}\mathbf{A}_{\nu}^{-}(\ell, j, j')\|_{\mathrm{Op}(j,j')} \lessdot \varepsilon\langle j - j'\rangle\|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}}. \qquad (5.99)$$

Now we write

$$\mathbf{A}_{\nu}^{-}(\ell, j, j'; u_2) = \mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1) + \Delta_{12}\mathbf{A}_{\nu}^{-}(\ell, j, j')$$
$$= \mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1)\left\{\mathbf{I}_{j,j'} + \mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1)^{-1}\Delta_{12}\mathbf{A}_{\nu}^{-}(\ell, j, j')\right\}. \qquad (5.100)$$

For all $|\ell| \leq N_{\nu}$, we have

$$\|\mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1)^{-1}\Delta_{12}\mathbf{A}_{\nu}^{-}(\ell, j, j')\|_{\mathrm{Op}(j,j')}$$
$$\leq \|\mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1)^{-1}\|_{\mathrm{Op}(j,j')}\|\Delta_{12}\mathbf{A}_{\nu}^{-}(\ell, j, j')\|_{\mathrm{Op}(j,j')}$$
$$\overset{(5.99)}{\lessdot} \frac{\langle\ell\rangle^{\tau}}{\gamma\langle j - j'\rangle}\varepsilon\langle j - j'\rangle\|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}} \lessdot \langle\ell\rangle^{\tau}\varepsilon\gamma^{-1}\|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}}$$
$$\lessdot N_{\nu}^{\tau}\varepsilon\gamma^{-1}\|u_1 - u_2\|_{s_0+\overline{\sigma}+\mathtt{b}} \overset{(5.97)}{\leq} \rho\gamma^{-1}. \qquad (5.101)$$

Since $\rho \leq \gamma/2$, we get that the operator $\mathbf{I}_{j,j'} + \mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1)^{-1}\Delta_{12}\mathbf{A}_{\nu}^{-}(\ell, j, j')$ is invertible and by Neumann series we get

$$\left\|\left(\mathbf{I}_{j,j'} + \mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1)^{-1}\Delta_{12}\mathbf{A}_{\nu}^{-}(\ell, j, j')\right)^{-1}\right\|_{\mathrm{Op}(j,j')} \leq \frac{\gamma}{\gamma - \rho}. \qquad (5.102)$$

By (5.100), (5.102) we get

$$\|\mathbf{A}_{\nu}^{-}(\ell, j, j'; u_2)^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{\gamma}{\gamma - \rho}\|\mathbf{A}_{\nu}^{-}(\ell, j, j'; u_1)^{-1}\|_{\mathrm{Op}(j,j')}$$
$$\leq \frac{\gamma}{\gamma - \rho}\frac{\langle\ell\rangle^{\tau}}{\gamma\langle j - j'\rangle} \leq \frac{\langle\ell\rangle^{\tau}}{(\gamma - \rho)\langle j - j'\rangle}. \qquad (5.103)$$

By similar arguments one can also prove that on the set $\Omega_{\nu+1}^{\gamma}(\mathbf{u}_1)$ the following estimate holds

$$\|\mathbf{A}_{\nu}^{+}(\ell, j, j'; u_2)^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{\langle\ell\rangle^{\tau}}{(\gamma - \rho)\langle j + j'\rangle}. \qquad (5.104)$$

Summarizing we have proved that if $\omega \in \Omega_{\nu+1}^{\gamma}(\mathbf{u}_1)$, then (5.103), (5.104) hold, implying that $\omega \in \Omega_{\nu+1}^{\gamma-\rho}(\mathbf{u}_2)$ (recall the definition (5.13)). This concludes the proof of $(\mathbf{S4})_{\nu+1}$.

5.4. Conjugation to a 2×2-block diagonal operator

In this section we prove that the operator $\mathcal{L}_0$ in (5.1) can be conjugated to the 2×2, time independent, block-diagonal operator $\mathcal{L}_\infty$ in (5.116). This will be proved in Theorem 5.2 and it is a consequence of the KAM reducibility Theorem 5.1. First, we state some auxiliary results.

Corollary 5.1. (KAM transformation) $\forall \omega \in \cap_{\nu \geq 0} \Omega_\nu^\gamma$ the sequence

$$\widetilde{\Phi}_\nu := \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu \tag{5.105}$$

converges in $|\cdot|_s^{\mathrm{Lip}(\gamma)}$ to an operator Φ_∞ and

$$\left|\Phi_\infty^{\pm 1} - \mathbb{I}_2\right|_s^{\mathrm{Lip}(\gamma)} \leq_s |\mathcal{R}_0|D|_{s+\mathtt{b}}^{\mathrm{Lip}(\gamma)} \gamma^{-1}. \tag{5.106}$$

Moreover $\Phi_\infty^{\pm 1}$ is symplectic.

Proof. The proof is similar to the one of Corollary 4.1 in [6] and hence it is omitted. $\qquad\square$

By Theorem 5.1-$(\mathbf{S2})_\nu$, for all $j \in \mathbb{N}$, the sequence of the 2×2 blocks $(\widetilde{\mathbf{D}}_j^\nu)_{\nu \geq 0}$ (defined for $\omega \in \Omega_o$) is a Cauchy sequence in $\mathcal{S}(\mathbf{E}_j)$ (recall (2.78)) with respect to $\|\cdot\|^{\mathrm{Lip}(\gamma)}$, then, it converges to a limit $\mathbf{D}_j^\infty(\omega) \in \mathcal{S}(\mathbf{E}_j)$, for any $\omega \in \Omega_o$. We have

$$\mathbf{D}_j^\infty(\omega) := \lim_{\nu \to +\infty} \widetilde{\mathbf{D}}_j^\nu(\omega) = \widetilde{\mathbf{D}}_j^0(\omega) + \widehat{\mathbf{D}}_j^\infty(\omega). \tag{5.107}$$

$$\widehat{\mathbf{D}}_j^\infty(\omega) := \sum_{\nu \geq 0} \widetilde{\mathbf{D}}_j^{\nu+1}(\omega) - \widetilde{\mathbf{D}}_j^\nu(\omega).$$

It could happen that $\Omega_{\nu_0}^\gamma = \emptyset$ (see (5.13)) for some ν_0. In such a case the iterative process of Theorem 5.1 stops after finitely many steps. However, we can always set $\widetilde{\mathbf{D}}_j^\nu := \widetilde{\mathbf{D}}_j^{\nu_0}$, $\forall \nu \geq \nu_0$, and the functions $\mathbf{D}_j^\infty : \Omega_o \to \mathcal{S}(\mathbf{E}_j)$ are always well defined.

Corollary 5.2. (Final blocks) *For all* $\nu \geq 0, j \in \mathbb{N}$,

$$\|\mathbf{D}_j^\infty - \widetilde{\mathbf{D}}_j^\nu\|^{\mathrm{Lip}(\gamma)} \lessdot N_{\nu-1}^{-\mathtt{a}}\varepsilon j^{-1}, \quad \|\widehat{\mathbf{D}}_j^\infty\|^{\mathrm{Lip}(\gamma)} \lessdot \varepsilon j^{-1}. \tag{5.108}$$

Proof. The bound (5.108) follows by a telescoping argument, applying the estimate (5.20). $\qquad\square$

Now we define the Cantor set

$$\Omega_\infty^{2\gamma} := \Omega_\infty^{2\gamma}(\mathbf{u})$$

$$:= \Bigg\{ \omega \in \Omega_o : \|\mathbf{A}_\infty^-(\ell, j, j'; \omega)^{-1}\| \leq \frac{\langle \ell \rangle^\tau}{2\gamma \langle j - j' \rangle}, \quad \forall (\ell, j, j') \in \mathbb{Z}^\nu \times \mathbb{N} \times \mathbb{N},$$

$$(\ell, j, j') \neq (0, j, j), \qquad \|\mathbf{A}_\infty^+(\ell, j, j'; \omega)^{-1}\| \leq \frac{\langle \ell \rangle^\tau}{2\gamma \langle j + j' \rangle},$$

$$\forall (\ell, j, j') \in \mathbb{Z}^\nu \times \mathbb{N} \times \mathbb{N} \Bigg\}, \tag{5.109}$$

where the operators $\mathbf{A}_\infty^\pm(\ell, j, j') = \mathbf{A}_\infty^\pm(\ell, j, j'; \omega) = \mathbf{A}_\infty^\pm(\ell, j, j'; \omega, u(\omega))$: $\mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j) \to \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ are defined by

$$\mathbf{A}_\infty^-(\ell, j, j') := \omega \cdot \ell \mathbf{I}_{j,j'} + M_L(\mathbf{D}_j^\infty) - M_R(\mathbf{D}_{j'}^\infty), \qquad (5.110)$$

$$\mathbf{A}_\infty^+(\ell, j, j') := \omega \cdot \ell \mathbf{I}_{j,j'} + M_L(\mathbf{D}_j^\infty) + M_R(\overline{\mathbf{D}}_{j'}^\infty). \qquad (5.111)$$

Lemma 5.3. (Cantor set)

$$\Omega_\infty^{2\gamma} \subset \cap_{\nu \geq 0} \Omega_\nu^\gamma. \qquad (5.112)$$

Proof. It suffices to show that for any $\nu \geq 0$, $\Omega_\infty^{2\gamma} \subseteq \Omega_\nu^\gamma$. We argue by induction. For $\nu = 0$, since $\Omega_0^\gamma = \Omega_o$, it follows from the definition (5.109) that $\Omega_\infty^{2\gamma} \subseteq \Omega_0^\gamma$. Assume that $\Omega_\infty^{2\gamma} \subseteq \Omega_\nu^\gamma$ for some $\nu \geq 0$ and let us prove that $\Omega_\infty^{2\gamma} \subseteq \Omega_{\nu+1}^\gamma$. Let $\omega \in \Omega_\infty^{2\gamma}$. By the inductive hypothesis $\omega \in \Omega_\nu^\gamma$, hence by Theorem 5.1, the self-adjoint matrices $\mathbf{D}_j^\nu(\omega) \in \mathcal{S}(\mathbf{E}_j)$ are well defined for all $j \in \mathbb{N}$ and $\mathbf{D}_j^\nu(\omega) = \widetilde{\mathbf{D}}_j^\nu(\omega)$. By the definitions (5.14), (5.15), also the matrices $\mathbf{A}_\nu^\pm(\ell, j, j'; \omega)$ are well defined. Since $\omega \in \Omega_\infty^{2\gamma}$, $\mathbf{A}_\infty^-(\ell, j, j'; \omega)$ is invertible and we may write

$$\begin{aligned} \mathbf{A}_\nu^-(\ell, j, j'; \omega) &= \mathbf{A}_\infty^-(\ell, j, j'; \omega) + \Delta_\infty^-(\ell, j, j'; \omega) \\ &= \mathbf{A}_\infty^-(\ell, j, j'; \omega)\left(\mathbf{I}_{j,j'} + \mathbf{A}_\infty^-(\ell, j, j'; \omega)^{-1}\Delta_\infty^-(\ell, j, j'; \omega)\right) \end{aligned}$$

where

$$\Delta_\infty^-(\ell, j, j'; \omega) := M_L\left(\mathbf{D}_j^\nu(\omega) - \mathbf{D}_j^\infty(\omega)\right) - M_R\left(\mathbf{D}_{j'}^\nu(\omega) - \mathbf{D}_{j'}^\infty(\omega)\right).$$

By the property (2.77) and by the estimate (5.108)

$$\|\Delta_\infty^-(\ell, j, j'; \omega)\|_{\mathrm{Op}(j,j')} \lessdot N_{\nu-1}^{-\mathsf{a}}\varepsilon j^{-1}. \qquad (5.113)$$

Moreover, one has that for all $|\ell| \leq N_\nu$, $j, j' \in \mathbb{N}$, with $(\ell, j, j') \neq (0, j, j)$

$$\|\mathbf{A}_\infty^-(\ell, j, j'; \omega)^{-1}\Delta_\infty^-(\ell, j, j'; \omega)\|_{\mathrm{Op}(j,j')} \lessdot \frac{N_\nu^\tau N_{\nu-1}^{-\mathsf{a}}}{2\gamma\langle j - j'\rangle}\varepsilon \leq \frac{1}{2}, \qquad (5.114)$$

by (5.6), and for δ_0 in (5.9) small enough. Therefore the operator $\mathbf{A}_\nu^-(\ell, j, j'; \omega)$ is invertible, with inverse given by the Neumann series. For all $|\ell| \leq N_\nu$, $j, j' \in \mathbb{N}$ with $(\ell, j, j') \neq (0, j, j')$

$$\begin{aligned} \|\mathbf{A}_\nu^-(\ell, j, j'; \omega)^{-1}\|_{\mathrm{Op}(j,j')} &\leq \frac{\|\mathbf{A}_\infty^-(\ell, j, j'; \omega)^{-1}\|_{\mathrm{Op}(j,j')}}{1 - \|\mathbf{A}_\infty^-(\ell, j, j'; \omega)^{-1}\Delta_\infty^-(\ell, j, j'; \omega)\|_{\mathrm{Op}(j,j')}} \\ &\overset{(5.114)}{\leq} 2\|\mathbf{A}_\infty^-(\ell, j, j'; \omega)^{-1}\|_{\mathrm{Op}(j,j')} \overset{(5.109)}{\leq} \frac{\langle\ell\rangle^\tau}{\gamma\langle j - j'\rangle}. \end{aligned}$$

By similar arguments, one can also obtain that for any $(\ell, j, j') \in \mathbb{Z}^\nu \times \mathbb{N} \times \mathbb{N}$ with $|\ell| \leq N_\nu$

$$\|\mathbf{A}_\nu^+(\ell, j, j'; \omega)^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{\langle\ell\rangle^\tau}{\gamma\langle j + j'\rangle}$$

and then the lemma follows. $\qquad\square$

To state the main result of this section we introduce the operator

$$\mathcal{D}_\infty := \mathrm{i} \begin{pmatrix} \mathcal{D}_\infty^{(1)} & 0 \\ 0 & -\overline{\mathcal{D}}_\infty^{(1)} \end{pmatrix}, \quad \mathcal{D}_\infty^{(1)} := \mathrm{diag}_{j\in\mathbb{N}} \mathbf{D}_j^\infty, \tag{5.115}$$

where the 2×2 self-adjoint blocks $\mathbf{D}_j^\infty \in \mathcal{S}(\mathbf{E}_j)$ are defined in (5.107). Furthermore, we introduce, for $\omega \in \Omega_o$, the operator

$$\mathcal{L}_\infty(\omega) := \omega \cdot \partial_\varphi \mathbb{I}_2 + \mathcal{D}_\infty(\omega). \tag{5.116}$$

Then $\mathcal{L}_\infty(\omega)$ is a φ-independent block-diagonal bounded linear Hamiltonian operator $\mathcal{L}_\infty(\omega) : \mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^{s-1}(\mathbb{T}^{\nu+1})$, for any $s \geq 1$.

Theorem 5.2. *Under the same assumptions of Theorem 5.1, the following holds:*

(i)   *For all $\omega \in \Omega_\infty^{2\gamma}$ and $s \in [s_0, S - \bar\sigma - \mathtt{b}]$, the transformations $\Phi_\infty^{\pm 1}$: $\mathbf{H}_0^s(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1})$ satisfy the estimates*

$$|\Phi_\infty^{\pm 1} - \mathbb{I}_2|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon\gamma^{-1}(1 + \|\mathbf{u}\|_{s+\bar\sigma+\mathtt{b}}^{\mathrm{Lip}(\gamma)}). \tag{5.117}$$

(ii)   *On the set $\Omega_\infty^{2\gamma}$, the Hamiltonian operator $\mathcal{L}_0$ in (5.1) is conjugated to the Hamiltonian operator $\mathcal{L}_\infty$ by Φ_∞, namely for all $\omega \in \Omega_\infty^{2\gamma}$,*

$$\mathcal{L}_\infty(\omega) = \Phi_\infty^{-1}(\omega)\mathcal{L}_0(\omega)\Phi_\infty(\omega). \tag{5.118}$$

Proof. (i)   Since $\Omega_\infty^{2\gamma}(\mathbf{u}) \overset{(5.112)}{\subseteq} \cap_{\nu\geq 0}\Omega_\nu^\gamma(\mathbf{u})$, the estimate (5.106) holds on the set $\Omega_\infty^{2\gamma}$, and $\forall s_0 \leq s \leq S - \bar\sigma - \mathtt{b}$,

$$|\Phi_\infty^{\pm 1} - \mathbb{I}_2|_s^{\mathrm{Lip}(\gamma)} \leq_s \gamma^{-1}|\mathcal{R}_0|D|\,|_{s+\mathtt{b}}^{\mathrm{Lip}(\gamma)} \overset{(5.4)}{\leq_s} \varepsilon\gamma^{-1}(1 + \|\mathbf{u}\|_{s+\bar\sigma+\mathtt{b}}^{\mathrm{Lip}(\gamma)}),$$

which is the claimed estimate (5.117).

(ii)   By (5.18), (5.105) we get

$$\mathcal{L}_\nu = \widetilde{\Phi}_{\nu-1}^{-1}\mathcal{L}_0\widetilde{\Phi}_{\nu-1} = \omega \cdot \partial_\varphi \mathbb{I}_2 + \mathcal{D}_\nu + \mathcal{R}_\nu, \qquad \widetilde{\Phi}_\nu = \Phi_0 \circ \ldots \circ \Phi_\nu. \tag{5.119}$$

Note that, for all $\nu \geq 0$,

$$|\mathcal{D}_\infty^{(1)} - \mathcal{D}_\nu^{(1)}|_s^{\mathrm{Lip}(\gamma)} \overset{Lemma\,2.6-(ii)}{\leq} |(\mathcal{D}_\infty^{(1)} - \mathcal{D}_\nu^{(1)})|D|\,|_s^{\mathrm{Lip}(\gamma)}$$

$$= \sup_{j\in\mathbb{N}} j\|\mathbf{D}_j^\infty - \mathbf{D}_j^\nu\|^{\mathrm{Lip}(\gamma)} \overset{(5.108)}{\lessdot} \varepsilon N_{\nu-1}^{-\mathtt{a}} \overset{\nu\to+\infty}{\to} 0 \tag{5.120}$$

and for any $s \in [s_0, S - \bar\sigma - \mathtt{b}]$

$$|\mathcal{R}_\nu|_s^{\mathrm{Lip}(\gamma)} \overset{Lemma\,2.6-(ii)}{\leq} |\mathcal{R}_\nu|D|\,|_s^{\mathrm{Lip}(\gamma)}$$

$$\overset{(5.17),(5.4)}{\lessdot} \varepsilon N_{\nu-1}^{-\mathtt{a}}\left(1 + \|\mathbf{u}\|_{s+\bar\sigma+\mathtt{b}}^{\mathrm{Lip}(\gamma)}\right) \overset{\nu\to+\infty}{\to} 0.$$

Hence, $|\mathcal{L}_\nu - \mathcal{L}_\infty|_s^{\mathrm{Lip}(\gamma)} \overset{\nu\to+\infty}{\to} 0$ for all $s_0 \leq s \leq S - \bar\sigma - \mathtt{b}$. Since, by Lemma 5.1, $\widetilde{\Phi}_\nu^{\pm 1} \overset{\nu\to+\infty}{\to} \Phi_\infty^{\pm 1}$ in the norm $|\cdot|_s^{\mathrm{Lip}(\gamma)}$, formula (5.118) follows by passing to the limit in (5.119). $\qquad\square$

Corollary 5.3. *Assume* (4.3) *with* $\mu = \bar{\sigma} + \mathtt{b} + s_0$, *then for any* $s_0 \leq s \leq S - \bar{\sigma} - \mathtt{b} - s_0$, *for any* $\varphi \in \mathbb{T}^\nu$, *the maps* $\Phi_\infty^{\pm 1}(\varphi) : \mathbf{H}_0^s(\mathbb{T}_x) \to \mathbf{H}_0^s(\mathbb{T}_x)$ *and they satisfy the estimates*

$$|\Phi_\infty^{\pm 1}(\varphi)|_{s,x} \lesssim_s 1 + \|u\|_{s+\bar{\sigma}+\mathtt{b}+s_0}.$$

Proof. The claimed estimate follows by Lemma 2.13-(ii) and by the estimate (5.117). $\qquad \square$

6. Inversion of the operator $\mathcal{L}$

We define

$$\mathcal{W}_1 := \mathcal{S}\mathcal{B}(\mathcal{A}\mathbb{I}_2)\mathcal{V}\Phi_\infty, \quad \mathcal{W}_2 := \mathcal{S}\mathcal{B}(\mathcal{A}\mathbb{I}_2)\rho\mathcal{V}\Phi_\infty \qquad (6.1)$$

(recall the Definitions (4.9), (4.28), (4.49), (4.30) and Corollary 5.1). By Sects. 4 and 5, the operator $\mathcal{L}$ in (4.1) may be written as

$$\mathcal{L} = \mathcal{W}_2 \mathcal{L}_\infty \mathcal{W}_1^{-1}, \qquad (6.2)$$

where the operator $\mathcal{L}_\infty$ is given in (5.116).

Lemma 6.1. *There exists* $\mu_0 := \mu_0(\tau, \nu) > 0$, $\mu_0 \geq \bar{\sigma} + \mathtt{b}$ *such that if* (4.3) *holds with* $\mu = \mu_0$, *then the operators* $\mathcal{W}_1$ *and* $\mathcal{W}_2$ *defined in* (6.1) *satisfies for any* $s_0 \leq s \leq S - \mu_0$, $m = 1, 2$

$$\mathcal{W}_m : \mathbf{H}_0^{s+\frac{1}{2}}(\mathbb{T}^{\nu+1}) \to H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2), \quad \mathcal{W}_m^{-1} : H_0^{s+\frac{1}{2}}(\mathbb{T}^{\nu+1}, \mathbb{R}^2) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1})$$

$$\|\mathcal{W}_m^{\pm 1}\mathbf{h}_\pm\|_s^{\mathrm{Lip}(\gamma)} \lesssim_s \|\mathbf{h}_\pm\|_{s+\mu_0}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\mu_0}^{\mathrm{Lip}(\gamma)}\|\mathbf{h}_\pm\|_{s_0+\mu_0}^{\mathrm{Lip}(\gamma)}, \quad m = 1, 2$$

for any Lipschitz family of Sobolev functions $\mathbf{h}_+(\cdot; \omega) \in \mathbf{H}_0^{s+\mu_0}(\mathbb{T}^{\nu+1})$, $\mathbf{h}_-(\cdot; \omega) \in H^{s+\mu_0}(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$, $\omega \in \Omega_\infty^{2\gamma}(\mathbf{u})$.

Proof. The lemma follows by the estimates (4.16), (4.37), (4.38), (4.55), (4.40), (5.117) and applying also Lemma 2.8. $\qquad \square$

For all $\ell \in \mathbb{Z}^\nu$, for all $j \in \mathbb{N}$, for all $\omega \in \Omega_o = \Omega_o(\mathbf{u})$, we define the 2×2 matrix $\mathbf{B}_\infty(\ell, j; \omega) = \mathbf{B}_\infty(\ell, j; \omega, u(\omega))$ as

$$\mathbf{B}_\infty(\ell, j; \omega) := \omega \cdot \ell \, \mathbf{I}_j + \mathbf{D}_j^\infty(\omega) \qquad (6.3)$$

(recall (2.63), (5.107)). Then we define the set

$$\Lambda_\infty^{2\gamma}(\mathbf{u}) := \left\{ \omega \in \Omega_o(\mathbf{u}) : \|\mathbf{B}_\infty(\ell, j; \omega)^{-1}\| \leq \frac{\langle \ell \rangle^\tau}{2\gamma j}, \quad \forall (\ell, j) \in \mathbb{Z}^\nu \times \mathbb{N} \right\}. \quad (6.4)$$

We prove the following

Lemma 6.2. *[Invertibility of $\mathcal{L}_\infty$] For all* $\omega \in \Lambda_\infty^{2\gamma}(\mathbf{u})$, *the operator* $\mathcal{L}_\infty$ *is invertible and its inverse* $\mathcal{L}_\infty^{-1} : \mathbf{H}_0^{s+2\tau+1}(\mathbb{T}^{\nu+1}) \to \mathbf{H}_0^s(\mathbb{T}^{\nu+1})$ *satisfies the tame estimate*

$$\|\mathcal{L}_\infty^{-1}\mathbf{h}\|_s^{\mathrm{Lip}(\gamma)} \lessdot \gamma^{-1}\|\mathbf{h}\|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)}$$

for any Lipschitz family $\mathbf{h}(\cdot; \omega) \in \mathbf{H}_0^{s+2\tau+1}(\mathbb{T}^{\nu+1})$, $\omega \in \Lambda_\infty^{2\gamma}(\mathbf{u})$

Proof. By (5.115), (5.116) the operator $\mathcal{L}_\infty$ has the form

$$\mathcal{L}_\infty = \begin{pmatrix} \mathcal{L}_\infty^{(1)} & 0 \\ 0 & \overline{\mathcal{L}}_\infty^{(1)} \end{pmatrix}, \qquad \mathcal{L}_\infty^{(1)} := \omega \cdot \partial_\varphi + \mathrm{i}\mathcal{D}_\infty^{(1)},$$

then it suffices to prove that $\mathcal{L}_\infty^{(1)}$ is invertible and its inverse satisfies the claimed estimate. Let $\omega \in \Lambda_\infty^{2\gamma}(\mathbf{u})$, and let $h \in H_0^{s+\tau}(\mathbb{T}^{\nu+d})$. Using the block-representation (2.64), we have

$$[\mathcal{L}_\infty^{(1)}]^{-1}h(\varphi,x) = \sum_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{N}} \mathbf{B}_\infty(\ell,j;\omega)^{-1}[\widehat{\mathbf{h}}_j(\ell)]e^{\mathrm{i}\ell\cdot\varphi}, \tag{6.5}$$

where the 2×2 self-adjoint matrices $\mathbf{B}_\infty(\ell,j;\omega)$ are defined in (6.3). Using (2.11), we get

$$\|[\mathcal{L}_\infty^{(1)}]^{-1}h\|_s^2 \leq \sum_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{N}} \langle\ell,j\rangle^{2s}\|\mathbf{B}_\infty(\ell,j;\omega)^{-1}\widehat{\mathbf{h}}_j(\ell)\|_{L^2}^2$$

$$\overset{(6.4)}{\lesssim} \sum_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{N}} \langle\ell,j\rangle^{2s}\langle\ell\rangle^{2\tau}\gamma^{-2}\|\widehat{\mathbf{h}}_j(\ell)\|_{L^2}^2 \lesssim \gamma^{-2}\|h\|_{s+\tau}^2. \tag{6.6}$$

Now let us consider a Lipschitz family of Sobolev functions $\omega \in \Lambda_\infty^{2\gamma}(\mathbf{u}) \to h(\cdot\,;\omega) \in H_0^{s+2\tau+1}$. Then, for any $\omega_1,\omega_2 \in \Lambda_\infty^{2\gamma}(\mathbf{u})$ one has that

$$[\mathcal{L}_\infty^{(1)}(\omega_1)]^{-1}h(\omega_1) - [\mathcal{L}_\infty^{(1)}(\omega_2)]^{-1}h(\omega_2)$$
$$= [\mathcal{L}_\infty^{(1)}(\omega_1)]^{-1}\left(h(\omega_1) - h(\omega_2)\right) + \left(\mathcal{L}_\infty^{(1)}(\omega_1) - \mathcal{L}_\infty^{(1)}(\omega_2)\right)h(\omega_2). \tag{6.7}$$

Arguing as in (6.6), the first term in (6.7) satisfies

$$\left\|[\mathcal{L}_\infty^{(1)}(\omega_1)]^{-1}\left(h(\omega_1) - h(\omega_2)\right)\right\|_s \lesssim \gamma^{-1}\|h(\omega_1) - h(\omega_2)\|_{s+\tau}$$
$$\lesssim \gamma^{-2}\|h\|_{s+\tau}^{\mathrm{Lip}(\gamma)}|\omega_1 - \omega_2|. \tag{6.8}$$

Now we estimate the second term in (6.7). One has that

$$\left(\mathcal{L}_\infty^{(1)}(\omega_1)^{-1} - \mathcal{L}_\infty^{(1)}(\omega_2)^{-1}\right)h(\omega_2)$$
$$= \sum_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{N}} \left(\mathbf{B}_\infty(\ell,j;\omega_1)^{-1} - \mathbf{B}_\infty(\ell,j;\omega_2)^{-1}\right)\widehat{\mathbf{h}}_j(\ell;\omega_2)e^{\mathrm{i}\ell\cdot\varphi}.$$

Since

$$\mathbf{B}_\infty(\ell,j;\omega_1)^{-1} - \mathbf{B}_\infty(\ell,j;\omega_2)^{-1}$$
$$= \mathbf{B}_\infty(\ell,j;\omega_1)^{-1}\left(\mathbf{B}_\infty(\ell,j;\omega_2) - \mathbf{B}_\infty(\ell,j;\omega_1)\right)\mathbf{B}_\infty(\ell,j;\omega_2)^{-1},$$

using that $\omega_1,\omega_2 \in \Lambda_\infty^{2\gamma}(\mathbf{u})$, one has that

$$\left\|\mathbf{B}_\infty(\ell,j;\omega_1)^{-1} - \mathbf{B}_\infty(\ell,j;\omega_2)^{-1}\right\|$$
$$\leq \frac{\langle\ell\rangle^{2\tau}}{\gamma^2 j^2}\left\|\mathbf{B}_\infty(\ell,j;\omega_2) - \mathbf{B}_\infty(\ell,j;\omega_1)\right\|. \tag{6.9}$$

Furthermore, by (5.107), (5.12) one gets

$$\|\mathbf{B}_\infty(\ell, j; \omega_2) - \mathbf{B}_\infty(\ell, j; \omega_1)\|$$
$$\leq |\omega_1 - \omega_2||\ell| + |m(\omega_1) - m(\omega_2)|j + \|\widehat{\mathbf{D}}_j^\infty(\omega_2) - \widehat{\mathbf{D}}_j^\infty(\omega_1)\|$$
$$\overset{(4.36),(5.108)}{\lessdot} \left(|\ell| + \varepsilon\gamma^{-1}j\right)|\omega_1 - \omega_2|. \tag{6.10}$$

The estimates (6.9), (6.10) (using that $\varepsilon\gamma^{-1} \leq 1$) imply that

$$\left\|\mathbf{B}_\infty(\ell, j; \omega_1)^{-1} - \mathbf{B}_\infty(\ell, j; \omega_2)^{-1}\right\| \lessdot \langle\ell\rangle^{2\tau+1}\gamma^{-2}|\omega_1 - \omega_2|. \tag{6.11}$$

Therefore

$$\left\|\left(\mathcal{L}_\infty^{(1)}(\omega_1)^{-1} - \mathcal{L}_\infty^{(1)}(\omega_2)^{-1}\right) h(\omega_2)\right\|_s^2$$
$$\overset{(2.11)}{\leq} \sum_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{N}} \langle\ell, j\rangle^{2s} \left\|\left(\mathbf{B}_\infty(\ell, j; \omega_1)^{-1} - \mathbf{B}_\infty(\ell, j; \omega_2)^{-1}\right)\widehat{\mathbf{h}}_j(\ell; \omega_2)\right\|_{L^2}^2$$
$$\overset{(6.9)}{\lessdot} \gamma^{-4} \sum_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{N}} \langle\ell, j\rangle^{2s}\langle\ell\rangle^{4\tau+2}\|\widehat{\mathbf{h}}_j(\ell; \omega_2)\|_{L^2}^2|\omega_1 - \omega_2|^2 \tag{6.12}$$

which implies (recalling (2.11)) that

$$\left\|\left(\mathcal{L}_\infty^{(1)}(\omega_1) - \mathcal{L}_\infty^{(1)}(\omega_2)\right) h(\omega_2)\right\|_s \lessdot \gamma^{-2}\|h\|_{s+2\tau+1}^{\sup}|\omega_1 - \omega_2|. \tag{6.13}$$

Hence, by (6.7), (6.8), (6.13), one gets

$$\gamma\|[\mathcal{L}_\infty^{(1)}]^{-1}h\|_s^{\mathrm{lip}} \lessdot \gamma^{-1}\|h\|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)}. \tag{6.14}$$

Recalling (6.6) and the definition of the norm $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$ in (2.5), the lemma follows. $\qquad\square$

Theorem 6.1. *[**Invertibility of** $\mathcal{L}$] Let $\gamma \in (0,1)$ and $\tau > 0$. There exists a constant*

$$\mu_1 = \mu_1(\tau, \nu) \geq \mu_0 \geq \overline{\sigma} + \mathtt{b} \tag{6.15}$$

where μ_0 is given in Lemma 6.1, such that for any Lipschitz family $\mathbf{u}(\cdot; \omega) \in H_0^S(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$, $S \geq s_0 + \mu_1$, satisfying

$$\|\mathbf{u}\|_{s_0+\mu_1}^{\mathrm{Lip}(\gamma)} \leq 1 \tag{6.16}$$

there exists a constant $\delta_1 = \delta_1(S, \tau, \nu) > 0$ (possibly smaller than the constant δ_0 given in Theorem 5.1) such that if

$$\varepsilon\gamma^{-1} \leq \delta_1, \tag{6.17}$$

then for all $\omega \in \mathbf{\Omega}_\infty^{2\gamma}(\mathbf{u}) := \Omega_\infty^{2\gamma}(\mathbf{u}) \cap \Lambda_\infty^{2\gamma}(\mathbf{u})$ (see (6.4), (5.109)), the operator $\mathcal{L}$ defined in (4.1) is invertible and its inverse $\mathcal{L}^{-1} : H_0^{s+\mu_1}(\mathbb{T}^{\nu+1}, \mathbb{R}^2) \to H_0^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$ satisfies $\forall s_0 \leq s \leq S - \mu_1$ the tame estimate

$$\|\mathcal{L}^{-1}\mathbf{h}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \gamma^{-1}\left(\|\mathbf{h}\|_{s+\mu_1}^{\mathrm{Lip}(\gamma)} + \|\mathbf{u}\|_{s+\mu_1}^{\mathrm{Lip}(\gamma)}\|\mathbf{h}\|_{s_0+\mu_1}^{\mathrm{Lip}(\gamma)}\right), \tag{6.18}$$

for any Lipschitz family $\mathbf{h}(\cdot; \omega) \in H_0^{s+\mu_1}(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$, $\omega \in \mathbf{\Omega}_\infty^{2\gamma}(\mathbf{u})$.

Proof. The estimate (6.18) follows by (6.2) and by Lemmata 6.1, 6.2. $\qquad\square$

7. The Nash–Moser iteration

Our next goal is to prove Theorem 3.1. It will be a consequence of Theorem 7.1 below where we construct iteratively a sequence of better and better approximate solutions of the operator $\mathcal{F}(\mathbf{u}) = \mathcal{F}(\varepsilon, \omega, \mathbf{u})$, defined in (3.7), and of the Sects. 8 and 9.

We consider the finite-dimensional subspaces

$$\mathcal{H}_n := \left\{\mathbf{u} \in L_0^2(\mathbb{T}^{\nu+1}, \mathbb{R}^2) : \quad \mathbf{u} = \Pi_n \mathbf{u}\right\}$$

where Π_n is the projector

$$\Pi_n \mathbf{u} := (\Pi_n u, \Pi_n \psi), \qquad \Pi_n h(\varphi, x) := \sum_{|(\ell, j)| \leq N_n} \widehat{h}_j(\ell) e^{i(\ell \cdot \varphi + j \cdot x)} \qquad (7.1)$$

with $N_n = N_0^{\chi^n}$ (see (5.5)). We also define $\Pi_n^\perp := \mathrm{Id} - \Pi_n$. The projectors Π_n, $\Pi_n^\perp$ satisfy the following classical smoothing properties for the weighted norm $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$, namely $\forall s, b \geq 0$,

$$\|\Pi_n \mathbf{u}\|_{s+b}^{\mathrm{Lip}(\gamma)} \leq K_n^b \|\mathbf{u}\|_s^{\mathrm{Lip}(\gamma)}, \qquad \|\Pi_n^\perp \mathbf{u}\|_s^{\mathrm{Lip}(\gamma)} \leq K_n^{-b} \|\mathbf{u}\|_{s+b}^{\mathrm{Lip}(\gamma)}. \qquad (7.2)$$

In view of the Nash–Moser Theorem 7.1 we introduce the constants

$$\kappa := 6\mu_1 + 19, \quad \mathtt{b}_1 := 2\mu_1 + 4 + \kappa(1 + \chi^{-1}) + 1, \qquad (7.3)$$

$$\mathtt{a}_1 := \kappa \chi^{-1} - 2\mu_1, \quad \chi := \frac{3}{2} \qquad (7.4)$$

where $\mu_1 := \mu_1(\tau, \nu) > 0$ is given in Theorem 6.1.

Theorem 7.1. (Nash–Moser) *Let* $\gamma \in (0,1)$ *and* $\tau > 0$. *Assume that* $f \in \mathcal{C}^q(\mathbb{T}^\nu \times \mathbb{T}, \mathbb{R})$, *with* $q \geq s_0 + \mathtt{b}_1$. *There exist* $\delta \in (0,1)$ *small enough and* $N_0 > 0$, $C_* > 0$ *large enough such that if*

$$\varepsilon \gamma^{-1} \leq \delta \qquad (7.5)$$

then:

$(\mathcal{P}1)_n$ *For all* $n \geq 0$, *there exists a function* $\mathbf{u}_n := (u_n, \psi_n) : \mathcal{G}_n \subseteq \Omega \to \mathcal{H}_n$, $\omega \mapsto \mathbf{u}_n(\omega) = (u_n(\omega), \psi_n(\omega))$, *with*

$$\|\mathbf{u}_n\|_{s_0+\mu_1}^{\mathrm{Lip}(\gamma)} \leq 1, \qquad (7.6)$$

$\mathbf{u}_0 := 0$, *where* $\mathcal{G}_n$ *are Cantor like subsets of* Ω *defined inductively by:*

$$\mathcal{G}_0 := \Omega_{\gamma, \tau}, \quad \mathcal{G}_{n+1} := \mathbf{\Omega}_\infty^{2\gamma_n}(\mathbf{u}_n), \quad \forall n \geq 0, \qquad (7.7)$$

(recall (2.7)) where $\gamma_n := \gamma(1 + 2^{-n})$ *and the set* $\mathbf{\Omega}_\infty^{2\gamma_n}(\mathbf{u}_n)$ *is given in Theorem 6.1, with* $\Omega_o(\mathbf{u}_n) = \mathcal{G}_n$. *There exists a constant* $C_*' > C_*$ *such that for all* $n \geq 1$, *the difference* $\mathbf{h}_n := \mathbf{u}_n - \mathbf{u}_{n-1}$ *satisfies*

$$\|\mathbf{h}_n\|_{s_0+\mu_1}^{\mathrm{Lip}(\gamma)} \leq C_*' \varepsilon \gamma^{-1} N_n^{-\mathtt{a}_1}. \qquad (7.8)$$

$(\mathcal{P}2)_n$ *For all* $n \geq 0$, $\|\mathcal{F}(\mathbf{u}_n)\|_{s_0}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon N_n^{-\kappa}$.
$(\mathcal{P}3)_n$ *For all* $n \geq 0$, $\|\mathbf{u}_n\|_{s_0+\mathtt{b}_1}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon \gamma^{-1} N_n^\kappa$, $\qquad \|\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mathtt{b}_1}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon N_n^\kappa$.

All the Lip norms are defined on $\mathcal{G}_n$.

Proof. To simplify notations, in this proof we write $\| \cdot \|_s$ instead of $\| \cdot \|_s^{\mathrm{Lip}(\gamma)}$.

STEP 1: *Proof of* $(\mathcal{P}1,2,3)_0$. They follow, since by (3.7),

$$\mathcal{F}(0) = \begin{pmatrix} 0 \\ \varepsilon f_\perp(\varphi, x) \end{pmatrix}, \quad \|\mathcal{F}(0)\|_s \leq \varepsilon \|f\|_s$$

and taking $C_* > \max\{\|f\|_{s_0} N_0^\kappa, \|f\|_{s_0 + \mathsf{b}_1}\}$.

STEP 2: *Assume that* $(\mathcal{P}1,2,3)_n$ *hold for some* $n \geq 0$, *and prove* $(\mathcal{P}1,2,3)_{n+1}$. We are going to define the successive approximation $\mathbf{u}_{n+1}$. By $(\mathcal{P}1)_n$, one has $\|\mathbf{u}_n\|_{s_0 + \mu_1} \leq 1$, moreover the smallness condition (7.5) implies the smallness condition (6.17) of Theorem 6.1, by taking $\delta < \delta_1 = \delta_1(S, \tau, \nu)$ with $S = s_0 + \mu_1 + \mathsf{b}_1$. Then Theorem 6.1 can be applied to the linearized operator

$$\mathcal{L}_n = \mathcal{L}(\mathbf{u}_n) = \partial_\mathbf{u} \mathcal{F}(\mathbf{u}_n), \tag{7.9}$$

implying that for all $\omega \in \mathcal{G}_{n+1} = \mathbf{\Omega}_\infty^{2\gamma_n}(\mathbf{u}_n)$, the operator $\mathcal{L}_n$ is invertible and its inverse satisfies $\forall s_0 \leq s \leq s_0 + \mathsf{b}_1$, $\forall \mathbf{h} \in H_0^{s+\mu_1}(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$, the tame estimate

$$\|\mathcal{L}_n^{-1} \mathbf{h}\|_s \leq_s \gamma^{-1} \left(\|\mathbf{h}\|_{s+\mu_1} + \|\mathbf{u}_n\|_{s+\mu_1} \|\mathbf{h}\|_{s_0 + \mu_1} \right). \tag{7.10}$$

Specializing the above estimate for $s = s_0$, using (7.6), one gets

$$\|\mathcal{L}_n^{-1} \mathbf{h}\|_{s_0} \leq_{s_0} \gamma^{-1} \|\mathbf{h}\|_{s_0 + \mu_1}. \tag{7.11}$$

We define the successive approximation

$$\mathbf{u}_{n+1} := \mathbf{u}_n + \mathbf{h}_{n+1}, \quad \mathbf{h}_{n+1} := -\Pi_{n+1} \mathcal{L}_n^{-1} \Pi_{n+1} \mathcal{F}(\mathbf{u}_n) \in \mathcal{H}_{n+1} \tag{7.12}$$

where Π_n is defined in (7.1). We now show that the iterative scheme in (7.12) is rapidly converging. We write

$$\mathcal{F}(\mathbf{u}_{n+1}) = \mathcal{F}(\mathbf{u}_n) + \mathcal{L}_n \mathbf{h}_{n+1} + Q_n$$

where $\mathcal{L}_n := \partial_\mathbf{u} \mathcal{F}(\mathbf{u}_n)$ and

$$Q_n := Q(\mathbf{u}_n, \mathbf{h}_{n+1}), \quad Q(\mathbf{u}_n, \mathbf{h}) := \mathcal{F}(\mathbf{u}_n + \mathbf{h}) - \mathcal{F}(\mathbf{u}_n) - \mathcal{L}_n \mathbf{h}, \tag{7.13}$$

$\mathbf{h} \in \mathcal{H}_{n+1}$. Then, by the the definition of $\mathbf{h}_{n+1}$ in (7.12), writing $\Pi_{n+1} = \mathrm{Id} - \Pi_{n+1}^\perp$, we have

$$\mathcal{F}(\mathbf{u}_{n+1}) = \mathcal{F}(\mathbf{u}_n) - \mathcal{L}_n \Pi_{n+1} \mathcal{L}_n^{-1} \Pi_{n+1} \mathcal{F}(\mathbf{u}_n) + Q_n$$
$$= \Pi_{n+1}^\perp \mathcal{F}(\mathbf{u}_n) + R_n + Q_n, \tag{7.14}$$

where

$$R_n := [\mathcal{L}_n, \Pi_{n+1}^\perp] \mathcal{L}_n^{-1} \Pi_{n+1} \mathcal{F}(\mathbf{u}_n) = [\Pi_{n+1}, \mathcal{L}_n] \mathcal{L}_n^{-1} \Pi_{n+1} \mathcal{F}(\mathbf{u}_n). \tag{7.15}$$

We first note that, for all $\omega \in \mathcal{G}_n$, by $(\mathcal{P}2)_n$ and by the smallness condition (7.5), one has that

$$\|\mathcal{F}(\mathbf{u}_n)\|_{s_0} \gamma^{-1} \leq 1. \tag{7.16}$$

Lemma 7.1. *On the set $\mathcal{G}_{n+1}$, defining*

$$B_n := \|\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mathtt{b}_1} + \varepsilon\|\mathbf{u}_n\|_{s_0+\mathtt{b}_1}, \tag{7.17}$$

we have

$$B_{n+1} \leq_{s_0+\mathtt{b}_1} N_{n+1}^{2\mu_1+6} B_n, \tag{7.18}$$

$$\|\mathcal{F}(\mathbf{u}_{n+1})\|_{s_0} \leq_{s_0+\mathtt{b}_1} N_{n+1}^{2\mu_1+4-\mathtt{b}_1} B_n + N_{n+1}^{2\mu_1+6}\varepsilon\gamma^{-2}\|\mathcal{F}(\mathbf{u}_n)\|_{s_0}^2. \tag{7.19}$$

Proof. We first estimate $\mathbf{h}_{n+1}$ defined in (7.12).

Estimates of $\mathbf{h}_{n+1}$. By (7.12) and (7.2), (7.10) (applied for $s = s_0 + \mathtt{b}_1$), (7.11), (7.6), we get

$$\begin{aligned}
\|\mathbf{h}_{n+1}\|_{s_0+\mathtt{b}_1} &\leq_{s_0+\mathtt{b}_1} \gamma^{-1}\left(\|\Pi_{n+1}\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mu_1+\mathtt{b}_1}\right.\\
&\qquad\qquad\left.+\|\mathbf{u}_n\|_{s_0+\mu_1+\mathtt{b}_1}\|\Pi_{n+1}\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mu_1}\right)\\
&\overset{(7.2)}{\leq_{s_0+\mathtt{b}_1}} N_{n+1}^{2\mu_1}\gamma^{-1}\left(\|\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mathtt{b}_1} + \|\mathbf{u}_n\|_{s_0+\mathtt{b}_1}\|\mathcal{F}(\mathbf{u}_n)\|_{s_0}\right),\\
&\overset{(7.16)}{\leq_{s_0+\mathtt{b}_1}} N_{n+1}^{2\mu_1}\left(\gamma^{-1}\|\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mathtt{b}_1} + \|\mathbf{u}_n\|_{s_0+\mathtt{b}_1}\right) \tag{7.20}
\end{aligned}$$

$$\|\mathbf{h}_{n+1}\|_{s_0} \leq_{s_0} \gamma^{-1}N_{n+1}^{\mu_1}\|\mathcal{F}(\mathbf{u}_n)\|_{s_0}. \tag{7.21}$$

Now we estimate the terms Q_n in (7.13) and R_n in (7.15).

Estimate of Q_n. By (7.13), (3.7), (2.1), (7.6), (7.2), we have the quadratic estimate

$$\|Q(\mathbf{u}_n, \mathbf{h})\|_s \leq_s \varepsilon N_{n+1}^6\left(\|\mathbf{h}\|_s\|\mathbf{h}\|_{s_0} + \|\mathbf{u}_n\|_s\|\mathbf{h}\|_{s_0}^2\right), \tag{7.22}$$

$\forall \mathbf{h} \in \mathcal{H}_{n+1}$, $\forall s \geq s_0$. Then the term Q_n in (7.13) satisfies, by (7.22), (7.20), (7.21), (7.16)

$$\|Q_n\|_{s_0+\mathtt{b}_1} \leq_{s_0+\mathtt{b}_1} N_{n+1}^{2\mu_1+6}\varepsilon\left(\|\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mathtt{b}_1}\gamma^{-1} + \|\mathbf{u}_n\|_{s_0+\mathtt{b}_1}\right), \tag{7.23}$$

$$\|Q_n\|_{s_0} \leq_{s_0} N_{n+1}^{2\mu_1+6}\varepsilon\gamma^{-2}\|\mathcal{F}(\mathbf{u}_n)\|_{s_0}^2. \tag{7.24}$$

Estimate of R_n. Now we estimate the term R_n defined in (7.15). By (7.9), (4.1), Lemma 2.1, (7.2), (7.6) the operator $\mathcal{L}_n$ satisfies the estimates

$$\|[\mathcal{L}_n, \Pi_{n+1}^{\perp}]\mathbf{h}\|_{s_0} \leq_{s_0+\mathtt{b}_1} N_{n+1}^{-\mathtt{b}_1+2}\varepsilon\left(\|\mathbf{h}\|_{s_0+\mathtt{b}_1} + \|\mathbf{u}_n\|_{s_0+\mathtt{b}_1}\|\mathbf{h}\|_{s_0+2}\right),$$

$$\|[\mathcal{L}_n, \Pi_{n+1}^{\perp}]\mathbf{h}\|_{s_0+\mathtt{b}_1} = \|[\Pi_{n+1}, \mathcal{L}_n]\mathbf{h}\|_{s_0+\mathtt{b}_1}$$

$$\leq_{s_0+\mathtt{b}_1} N_{n+1}^2\varepsilon\left(\|\mathbf{h}\|_{s_0+\mathtt{b}_1} + \|\mathbf{u}_n\|_{s_0+\mathtt{b}_1}\|\mathbf{h}\|_{s_0+2}\right).$$

The above estimates, together with the estimates (7.10), (7.11), (7.2) and using also (7.16) imply

$$\|R_n\|_{s_0} \leq_{s_0+\mathtt{b}_1} N_{n+1}^{2\mu_1+4-\mathtt{b}_1}\varepsilon\left(\|\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mathtt{b}_1}\gamma^{-1} + \|\mathbf{u}_n\|_{s_0+\mathtt{b}_1}\right), \tag{7.25}$$

$$\|R_n\|_{s_0+\mathtt{b}_1} \leq_{s_0+\mathtt{b}_1} N_{n+1}^{2\mu_1+4}\varepsilon\left(\|\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mathtt{b}_1}\gamma^{-1} + \|\mathbf{u}_n\|_{s_0+\mathtt{b}_1}\right). \tag{7.26}$$

Estimates of $\mathbf{u}_{n+1}$. By (7.12) and by the estimates (7.20) one gets

$$\|\mathbf{u}_{n+1}\|_{s_0+\mathtt{b}_1} \leq_{s_0+\mathtt{b}_1} N_{n+1}^{2\mu_1}\left(\|\mathbf{u}_n\|_{s_0+\mathtt{b}_1} + \|\mathcal{F}(\mathbf{u}_n)\|_{s_0+\mathtt{b}_1}\gamma^{-1}\right). \tag{7.27}$$

Finally, by (7.14), (7.23), (7.24), (7.25), (7.26), (7.27), (7.2), $\varepsilon\gamma^{-1} \leq 1$ and recalling the definition (7.17), we deduce the estimates (7.19), (7.18). $\qquad\square$

The estimates (7.19), (7.18), together with (7.3), (7.4), $(\mathcal{P}2)_n, (\mathcal{P}3)_n$, (7.5), taking δ_1 small enough and N_0 large enough, imply $(\mathcal{P}2)_{n+1}, (\mathcal{P}3)_{n+1}$.

The estimate (7.8) at the step $n+1$ follows since

$$\|\mathbf{h}_{n+1}\|_{s_0+\mu_1} \overset{(7.2)}{\leq} N_{n+1}^{\mu_1}\|\mathbf{h}_{n+1}\|_{s_0} \overset{(7.21)}{\leq} C(s_0)N_{n+1}^{2\mu_1}\gamma^{-1}\|\mathcal{F}(\mathbf{u}_n)\|_{s_0}$$

$$\overset{(\mathcal{P}2)_n}{\leq} C(s_0)C_* N_{n+1}^{2\mu_1}N_n^{-\kappa}\varepsilon\gamma^{-1}$$

which implies the claimed estimate, by (7.4) and taking $C_*' = C(s_0)C_*$.

The estimate (7.6) at the step $n+1$ follows since

$$\|\mathbf{u}_{n+1}\|_{s_0+\mu_1} \leq \sum_{k=0}^{n+1}\|\mathbf{h}_k\|_{s_0+\mu_1} \overset{(7.8)}{\leq} C_*'\varepsilon\gamma^{-1}\sum_{k\geq 0}N_k^{-\mathbf{a}_1} \leq C'\varepsilon\gamma^{-1} \leq 1$$

by taking δ in (7.5) small enough. Then $(\mathcal{P}1)_{n+1}$ follows and the proof is concluded. $\qquad\square$

8. Measure estimates

In this Section we estimate the measure of the set

$$\mathcal{G}_\infty := \cap_{n\geq 0}\mathcal{G}_n \tag{8.1}$$

where the sets $\cdots \subseteq \mathcal{G}_{n+1} \subseteq \mathcal{G}_n \subseteq \cdots \subseteq \mathcal{G}_0$ are given in Theorem 7.1-$(\mathcal{P}1)_n$. First, let us define the constants

$$\tau^* := \nu + 2, \qquad \tau := 2\tau^* + \nu + 1, \tag{8.2}$$

$$\gamma^* := 5\gamma, \qquad \gamma_n^* := \gamma^*(1 + 2^{-n}), \qquad \forall n \geq 0 \tag{8.3}$$

and recall also that the constants

$$\gamma_n = \gamma(1 + 2^{-n}), \qquad \forall n \geq 0, \tag{8.4}$$

are given in Theorem 7.1-$(\mathcal{P}1)_n$. We prove the following

Theorem 8.1. *One has*

$$|\Omega\backslash\mathcal{G}_\infty| \lessdot \gamma.$$

The rest of this section is devoted to the proof of Theorem 8.1. First, we write

$$\Omega\backslash\mathcal{G}_\infty = (\Omega\backslash\mathcal{G}_0)\bigcup_{n\geq 0}(\mathcal{G}_n\backslash\mathcal{G}_{n+1}). \tag{8.5}$$

Since $\mathcal{G}_0 = \Omega_{\gamma,\tau}$, see (7.7), it follows by standard volume estimates

$$\Omega\backslash\mathcal{G}_0 = O(\gamma). \tag{8.6}$$

For any $n \geq 0$, we define the set

$$\Omega_{\gamma^*,\tau^*}^{(I)}(\mathbf{u}_n) := \left\{\omega \in \mathcal{G}_n : |\omega\cdot\ell + m(u_n)j| \geq \frac{\gamma_n^*\langle j\rangle}{\langle \ell\rangle^{\tau^*}}, \right. \tag{8.7}$$

$$\left. \forall(\ell,j) \in (\mathbb{Z}^\nu \times \mathbb{N}_0)\backslash\{(0,0)\}\right\}$$

where we recall that the constant $m(u_n) = m(\omega, u_n(\omega))$ satisfies the estimates (4.36) (recall also that $\mathbb{N}_0 = \{0\} \cup \mathbb{N} = \{0, 1, 2, \ldots\}$). For all $n \geq 0$, we make the splitting

$$\mathcal{G}_n \backslash \mathcal{G}_{n+1} = \mathcal{A}_n^{(1)} \cup \mathcal{A}_n^{(2)}, \tag{8.8}$$

where

$$\mathcal{A}_n^{(1)} := (\mathcal{G}_n \backslash \mathcal{G}_{n+1}) \cap \Omega_{\gamma^*, \tau^*}^{(I)}(\mathbf{u}_n), \tag{8.9}$$

$$\mathcal{A}_n^{(2)} := (\mathcal{G}_n \backslash \mathcal{G}_{n+1}) \cap (\mathcal{G}_n \backslash \Omega_{\gamma^*, \tau^*}^{(I)}(\mathbf{u}_n)). \tag{8.10}$$

Estimate of $\mathcal{A}_n^{(1)}$.

By (8.9), using the inductive definition of the sets $\mathcal{G}_n$ given in (7.7) and recalling (6.4), (5.109), one has that for all $n \geq 0$,

$$\mathcal{A}_n^{(1)} = \bigcup_{(\ell, j) \in \mathbb{Z}^\nu \times \mathbb{N}} Q_{\ell j}(\mathbf{u}_n) \bigcup_{\substack{\ell \in \mathbb{Z}^\nu \\ j, j' \in \mathbb{N} \\ (\ell, j, j') \neq (0, j, j)}} R_{\ell j j'}^-(\mathbf{u}_n) \bigcup_{\substack{\ell \in \mathbb{Z}^\nu \\ j, j' \in \mathbb{N}}} R_{\ell j j'}^+(\mathbf{u}_n) \tag{8.11}$$

where

$$Q_{\ell j}(\mathbf{u}_n) := \Big\{ \omega \in \mathcal{G}_n \cap \Omega_{\gamma^*, \tau^*}^{(I)}(\mathbf{u}_n) : \text{ the \ operator } \mathbf{B}_\infty(\ell, j; \omega, u_n(\omega))$$

$$\text{is \ not \ invertible \ or \ it \ is \ invertible \ and}$$

$$\|\mathbf{B}_\infty(\ell, j; \omega, u_n(\omega))^{-1}\| > \frac{\langle \ell \rangle^\tau}{2\gamma_n j} \Big\}, \tag{8.12}$$

$$R_{\ell j j'}^-(\mathbf{u}_n) := \Big\{ \omega \in \mathcal{G}_n \cap \Omega_{\gamma^*, \tau^*}^{(I)}(\mathbf{u}_n) : \text{ the \ operator } \mathbf{A}_\infty^-(\ell, j, j'; \omega, u_n(\omega))$$

$$\text{is \ not \ invertible \ or \ it \ is \ invertible \ and}$$

$$\|\mathbf{A}_\infty^-(\ell, j, j'; \omega, u_n(\omega))^{-1}\|_{\mathrm{Op}(j, j')} > \frac{\langle \ell \rangle^\tau}{2\gamma_n \langle j - j' \rangle} \Big\}, \tag{8.13}$$

$$R_{\ell j j'}^+(\mathbf{u}_n) := \Big\{ \omega \in \mathcal{G}_n \cap \Omega_{\gamma^*, \tau^*}^{(I)}(\mathbf{u}_n) : \text{ the \ operator } \mathbf{A}_\infty^+(\ell, j, j'; \omega, u_n(\omega))$$

$$\text{is \ not \ invertible \ or \ it \ is \ invertible \ and}$$

$$\|\mathbf{A}_\infty^+(\ell, j, j'; \omega, u_n(\omega))^{-1}\|_{\mathrm{Op}(j, j')} > \frac{\langle \ell \rangle^\tau}{2\gamma_n \langle j + j' \rangle} \Big\}, \tag{8.14}$$

where we recall that by (5.110), (5.111), (6.3)

$$\mathbf{B}_\infty(\ell, j; \omega, u_n(\omega)) := \omega \cdot \ell \mathbf{I}_j + \mathbf{D}_j^\infty(\omega, u_n(\omega)), \quad \ell \in \mathbb{Z}^\nu, \quad j \in \mathbb{N}, \tag{8.15}$$

$$\mathbf{A}_\infty^-(\ell, j, j'; \omega, u_n(\omega)) := \omega \cdot \ell \, \mathbf{I}_{j, j'} + M_L(\mathbf{D}_j^\infty(\omega, u_n(\omega))) \tag{8.16}$$

$$- M_R(\mathbf{D}_{j'}^\infty(\omega, u_n(\omega))),$$

for all $\ell \in \mathbb{Z}^\nu$, $j, j' \in \mathbb{N}$, $(\ell, j, j') \neq (0, j, j)$ and

$$\mathbf{A}_\infty^+(\ell, j, j'; \omega, u_n(\omega)) := \omega \cdot \ell \, \mathbf{I}_{j, j'} + M_L(\mathbf{D}_j^\infty(\omega, u_n(\omega))) \tag{8.17}$$

$$+ M_R(\overline{\mathbf{D}_{j'}^\infty(\omega, u_n(\omega))})$$

for $\ell \in \mathbb{Z}^\nu$, $j, j' \in \mathbb{N}$.

First we need to establish several auxiliary Lemmas.

Lemma 8.1. *For all $n \geq 1$,*

$$\sup_{j \in \mathbb{N}} \left\| \widehat{\mathbf{D}}_j^\infty(u_n) - \widehat{\mathbf{D}}_j^\infty(u_{n-1}) \right\| \lessdot \varepsilon N_{n-1}^{-\mathtt{a}}, \quad \forall \omega \in \mathcal{G}_n, \tag{8.18}$$

where $\widehat{\mathbf{D}}_j^\infty(u_n) = \widehat{\mathbf{D}}_j^\infty(\omega, u_n(\omega))$ is given in (5.107) and $\mathtt{a}$ is defined in (5.6).

Proof. We first apply Theorem 5.1-(**S4**)$_\nu$ with $\nu = n$, $\gamma = \gamma_{n-1}$, $\gamma - \rho = \gamma_n$, and $\mathbf{u}_1$, $\mathbf{u}_2$, replaced, respectively, by $\mathbf{u}_{n-1}$, $\mathbf{u}_n$, in order to conclude that

$$\Omega_n^{\gamma_{n-1}}(\mathbf{u}_{n-1}) \subseteq \Omega_n^{\gamma_n}(\mathbf{u}_n). \tag{8.19}$$

The smallness condition in (5.25) is satisfied because $\overline{\sigma} + \mathtt{b} < \mu_1$ (see (6.15)) and so

$$K_1 N_{n-1}^\tau \varepsilon \|u_n - u_{n-1}\|_{s_0+\overline{\sigma}+\mathtt{b}} \leq K_1 N_{n-1}^\tau \varepsilon \|u_n - u_{n-1}\|_{s_0+\mu_1}$$

$$\leq K_1 N_{n-1}^\tau \varepsilon \|\mathbf{u}_n - \mathbf{u}_{n-1}\|_{s_0+\mu_1} \overset{(7.8)}{\leq} C_*' K_1 \varepsilon^2 \gamma^{-1} N_{n-1}^\tau N_n^{-\mathtt{a}_1}$$

$$\leq \gamma_{n-1} - \gamma_n =: \rho = \gamma 2^{-n}$$

for $\varepsilon \gamma^{-1}$ small enough, N_0 large enough and using that $\mathtt{a}_1 > \tau$ (see (7.3), (7.4), (6.15), (5.6)). Then, by the definitions (7.7), (6.4), (5.109), we have

$$\mathcal{G}_n \subseteq \mathcal{G}_{n-1} \cap \Omega_\infty^{2\gamma_{n-1}}(\mathbf{u}_{n-1})$$

$$\overset{(5.112)}{\subseteq} \bigcap_{\nu \geq 0} \Omega_\nu^{\gamma_{n-1}}(\mathbf{u}_{n-1}) \subset \Omega_n^{\gamma_{n-1}}(\mathbf{u}_{n-1}) \overset{(8.19)}{\subseteq} \Omega_n^{\gamma_n}(\mathbf{u}_n).$$

Next, for all $\omega \in \mathcal{G}_n \subset \Omega_n^{\gamma_{n-1}}(\mathbf{u}_{n-1}) \cap \Omega_n^{\gamma_n}(\mathbf{u}_n)$ both the operators $\widehat{\mathbf{D}}_j^n(u_{n-1})$ and $\widehat{\mathbf{D}}_j^n(u_n)$ are well defined and applying the estimate (5.24) with $\nu = n$, we deduce that

$$\sup_{j \in \mathbb{N}} \left\| \widehat{\mathbf{D}}_j^n(u_n) - \widehat{\mathbf{D}}_j^n(u_{n-1}) \right\| \overset{(5.24)}{\lessdot} \varepsilon \|u_n - u_{n-1}\|_{s_0+\overline{\sigma}+\mathtt{b}}$$

$$\overset{(6.15)}{\lessdot} \varepsilon \|\mathbf{u}_n - \mathbf{u}_{n-1}\|_{s_0+\mu_1} \overset{(7.8)}{\lessdot} \varepsilon N_n^{-\mathtt{a}_1}. \tag{8.20}$$

Moreover by (5.107), (5.108) (with $\nu = n$), for all $j \in \mathbb{N}$, we get

$$\left\| \widehat{\mathbf{D}}_j^\infty(u_{n-1}) - \widehat{\mathbf{D}}_j^n(u_{n-1}) \right\| + \left\| \widehat{\mathbf{D}}_j^\infty(u_n) - \widehat{\mathbf{D}}_j^n(u_n) \right\| \lessdot \varepsilon N_{n-1}^{-\mathtt{a}}. \tag{8.21}$$

Therefore, for all $\omega \in \mathcal{G}_n$, $\forall j \in \mathbb{N}$,

$$\left\| \widehat{\mathbf{D}}_j^\infty(u_n) - \widehat{\mathbf{D}}_j^\infty(u_{n-1}) \right\|$$

$$\leq \left\| \widehat{\mathbf{D}}_j^n(u_n) - \widehat{\mathbf{D}}_j^n(u_{n-1}) \right\| + \left\| \widehat{\mathbf{D}}_j^\infty(u_{n-1}) - \widehat{\mathbf{D}}_j^n(u_{n-1}) \right\|$$

$$+ \left\| \widehat{\mathbf{D}}_j^\infty(u_n) - \widehat{\mathbf{D}}_j^n(u_n) \right\|$$

$$\overset{(8.20),(8.21)}{\lessdot} \varepsilon(N_{n-1}^{-\mathtt{a}} + N_n^{-\mathtt{a}_1}) \lessdot \varepsilon N_{n-1}^{-\mathtt{a}}, \tag{8.22}$$

since by (7.3), (7.4) one has $\mathtt{a}_1 > \mu_1$, and by (6.15) $\mu_1 \geq \overline{\sigma} + \mathtt{b} \overset{(5.6)}{\geq} \mathtt{a}$. Then the claimed estimate is proved. $\qquad \square$

Lemma 8.2. *For $\varepsilon\gamma^{-1}$ small enough, for all $n \geq 1$, $|\ell| \leq N_{n-1}$,*

$$Q_{\ell j}(\mathbf{u}_n) \subseteq Q_{\ell j}(\mathbf{u}_{n-1}), \qquad R_{\ell jj'}^{\pm}(\mathbf{u}_n) \subseteq R_{\ell jj'}^{\pm}(\mathbf{u}_{n-1}). \tag{8.23}$$

Proof. We prove that $R_{\ell jj'}^-(\mathbf{u}_n) \subseteq R_{\ell jj'}^-(\mathbf{u}_{n-1})$. The proof of the other inclusions is analogous. For all $j, j' \in \mathbb{N}$, $|\ell| \leq N_{n-1}$, $(\ell, j, j') \neq (0, j, j)$, $\omega \in \mathcal{G}_n$, we write

$$\mathbf{A}_\infty^-(\ell, j, j'; u_n) = \mathbf{A}_\infty^-(\ell, j, j'; u_{n-1}) \left(\mathbf{I}_{j,j'} + \mathbf{A}_\infty^-(\ell, j, j'; u_{n-1})^{-1} \Delta_\infty(j, j', n)\right),$$

where

$$\Delta_\infty(j, j', n) := M_L \left(\mathbf{D}_j^\infty(u_n) - \mathbf{D}_j^\infty(u_{n-1})\right) \tag{8.24}$$
$$- M_R \left(\mathbf{D}_{j'}^\infty(u_n) - \mathbf{D}_{j'}^\infty(u_{n-1})\right).$$

Note that

$$\Delta_\infty(j, j', n) \overset{(5.12),(5.107)}{=} (m(u_n) - m(u_{n-1}))(j - j')\mathbf{I}_{j,j'}$$
$$+ M_L \left(\widehat{\mathbf{D}}_j^\infty(u_n) - \widehat{\mathbf{D}}_j^\infty(u_{n-1})\right) - M_R \left(\widehat{\mathbf{D}}_{j'}^\infty(u_n) - \widehat{\mathbf{D}}_{j'}^\infty(u_{n-1})\right). \tag{8.25}$$

By (8.25), (4.36), (2.77), (8.18) one gets

$$\|\Delta_\infty(j, j', n)\|_{\mathrm{Op}(j,j')} \lessdot \varepsilon \langle j - j' \rangle N_{n-1}^{-\mathtt{a}}. \tag{8.26}$$

Hence for $|\ell| \leq N_{n-1}$ we get

$$\left\|\mathbf{A}_\infty^-(\ell, j, j'; u_{n-1})^{-1} \Delta_\infty(j, j', n)\right\|_{\mathrm{Op}(j,j')} \leq \frac{\langle \ell \rangle^\tau}{2\gamma_{n-1}\langle j - j' \rangle} \|\Delta_\infty(j, j', n)\|_{\mathrm{Op}(j,j')}$$

$$\overset{(8.26)}{\lessdot} \varepsilon\gamma^{-1} N_{n-1}^{\tau-\mathtt{a}} \leq \frac{1}{2}$$

for $\varepsilon\gamma^{-1}$ small enough and since $\mathtt{a} > \tau$ (see (5.6)). Then for all $|\ell| \leq N_{n-1}$, for all $j, j' \in \mathbb{N}$, the operator $\mathbf{A}_\infty^-(\ell, j, j'; u_n)$ is invertible by Neumann series and

$$\|\mathbf{A}_\infty^-(\ell, j, j'; u_n)^{-1}\|_{\mathrm{Op}(j,j')} \leq \|\mathbf{A}_\infty^-(\ell, j, j'; u_{n-1})^{-1}\|_{\mathrm{Op}(j,j')} \left(1 + C\varepsilon\gamma^{-1} N_{n-1}^{\tau-\mathtt{a}}\right)$$

$$\leq \frac{\langle \ell \rangle^\tau}{2\gamma_{n-1}\langle j - j' \rangle} \left(1 + C\varepsilon\gamma^{-1} N_{n-1}^{\tau-\mathtt{a}}\right).$$

Since by the definition of γ_n,

$$\frac{\gamma_{n-1} - \gamma_n}{\gamma_n} = \frac{1}{1 + 2^n},$$

it follows that for $\varepsilon\gamma^{-1}$ sufficiently small

$$C\varepsilon\gamma^{-1} N_{n-1}^{\tau-\mathtt{a}} \leq \frac{\gamma_{n-1} - \gamma_n}{\gamma_n}.$$

Hence

$$\|\mathbf{A}_\infty^-(\ell, j, j'; u_n)^{-1}\|_{\mathrm{Op}(j,j')} \leq \frac{\langle \ell \rangle^\tau}{2\gamma_n \langle j - j' \rangle}$$

which implies the claimed inclusion $R_{\ell jj'}^-(\mathbf{u}_n) \subseteq R_{\ell jj'}^-(\mathbf{u}_{n-1})$ in (8.23). $\qquad\square$

Corollary 8.1. *For any $n \geq 1$*
(i) $Q_{\ell j}(\mathbf{u}_n) = \emptyset$, *for all* $|\ell| \leq N_{n-1}$,

(ii) $R^-_{\ell j j'}(\mathbf{u}_n) = \emptyset$, for all $|\ell| \leq N_{n-1}$, $(\ell, j, j') \neq (0, j, j)$,

(iii) $R^+_{\ell j j'}(\mathbf{u}_n) = \emptyset$, for all $|\ell| \leq N_{n-1}$.

Hence for any $n \geq 1$,

$$\mathcal{A}^{(1)}_n \overset{(8.11)}{=} \bigcup_{\substack{|\ell| > N_{n-1} \\ j \in \mathbb{N}}} Q_{\ell j}(\mathbf{u}_n) \bigcup \bigcup_{\substack{|\ell| > N_{n-1} \\ j, j' \in \mathbb{N} \\ (\ell, j, j') \neq (0, j, j)}} R^-_{\ell j j'}(\mathbf{u}_n) \bigcup \bigcup_{\substack{|\ell| > N_{n-1} \\ j, j' \in \mathbb{N}}} R^+_{\ell j j'}(\mathbf{u}_n). \tag{8.27}$$

Proof. By definition, $R^\pm_{\ell j j'}(\mathbf{u}_n), Q_{\ell j}(\mathbf{u}_n) \subset \mathcal{G}_n$ and, by (8.23), for all $|\ell| \leq N_{n-1}$, we have $R^\pm_{\ell j j'}(\mathbf{u}_n) \subseteq R^\pm_{\ell j j'}(\mathbf{u}_{n-1})$ and $Q_{\ell j}(\mathbf{u}_n) \subseteq Q_{\ell j}(\mathbf{u}_{n-1})$. On the other hand again by definition $R^\pm_{\ell j j'}(\mathbf{u}_{n-1}) \cap \mathcal{G}_n, Q_{\ell j}(\mathbf{u}_{n-1}) \cap \mathcal{G}_n = \emptyset$. As a consequence, $\forall |\ell| \leq N_{n-1}$, $R^\pm_{\ell j j'}(\mathbf{u}_n)$, $Q_{\ell j}(\mathbf{u}_n) = \emptyset$. $\square$

Lemma 8.3. *For all $n \geq 0$ the following statements hold:*

(i)   *If $Q_{\ell j}(\mathbf{u}_n) \neq \emptyset$, then $\ell \neq 0$ and $j \prec |\ell|$.*

(ii)   *If $R^-_{\ell j j'}(\mathbf{u}_n) \neq \emptyset$, then $\ell \neq 0$ and $|j - j'| \prec |\ell|$, $j, j' \prec |\ell|^{\tau^*}$.*

(iii)   *If $R^+_{\ell j j'}(\mathbf{u}_n) \neq \emptyset$, then $\ell \neq 0$ and $j, j' \prec |\ell|$.*

Proof. We prove item (ii). The proofs of items (i) and (iii) are similar. The statement follows by the following claim:

CLAIM If $\varepsilon \gamma^{-1}$ is small enough and

$$\langle \ell \rangle^{\tau^*} \leq \langle j - j' \rangle \min\{j, j'\} \tag{8.28}$$

then for all $\omega \in \mathcal{G}_n \cap \Omega^{(I)}_{\gamma^*, \tau^*}(\mathbf{u}_n)$ (recall (8.7)), the matrix $\mathbf{A}^-_\infty(\ell, j, j') = \mathbf{A}^-_\infty(\ell, j, j'; \omega, u_n(\omega))$ is invertible and

$$\|\mathbf{A}^-_\infty(\ell, j, j')^{-1}\|_{\mathrm{Op}(j, j')} \leq \frac{\langle \ell \rangle^\tau}{2\gamma_n \langle j - j' \rangle}.$$

PROOF OF THE CLAIM. By (5.110), (5.107), (5.12), (5.8), we can write

$$\mathbf{A}^-_\infty(\ell, j, j') = \mathbf{I}_\infty(\ell, j, j') + \Delta_\infty(j, j'), \tag{8.29}$$

where

$$\mathbf{I}_\infty(\ell, j, j') := (\omega \cdot \ell + m(j - j')) \, \mathbf{I}_{j,j'}, \quad \Delta_\infty(j, j') := M_L(\widehat{\mathbf{D}}^\infty_j) - M_R(\widehat{\mathbf{D}}^\infty_{j'}).$$

Since $\omega \in \Omega^{(I)}_{\gamma^*, \tau^*}(\mathbf{u}_n)$, for any $(\ell, j, j') \neq (0, j, j)$, the operator $\mathbf{I}_\infty(\ell, j, j')$ is invertible and its inverse satisfies the bound

$$\|\mathbf{I}_\infty(\ell, j, j')^{-1}\|_{\mathrm{Op}(j, j')} \leq \frac{\langle \ell \rangle^{\tau^*}}{\gamma^*_n \langle j - j' \rangle}. \tag{8.30}$$

Moreover the operatorial norm of the operator $\Delta_\infty(j, j')$ satisfies

$$\|\Delta_\infty(j, j')\|_{\mathrm{Op}(j, j')} \overset{(2.77),(5.108)}{\prec} \varepsilon \left(\frac{1}{j} + \frac{1}{j'} \right) \prec \frac{\varepsilon}{\min\{j, j'\}}. \tag{8.31}$$

The estimates (8.30), (8.31) imply that

$$\|\mathbf{I}_\infty(\ell, j, j')^{-1} \Delta_\infty(j, j')\|_{\mathrm{Op}(j, j')} \prec \frac{\varepsilon \langle \ell \rangle^{\tau^*}}{\gamma^*_n \langle j - j' \rangle \min\{j, j'\}} \leq \frac{1}{2},$$

by (8.28) and for $\varepsilon\gamma^{-1}$ small enough. Hence by (8.29), the matrix $\mathbf{A}_\infty^-(\ell,j,j')$ is invertible by Neumann series and

$$\|\mathbf{A}_\infty^-(\ell,j,j')^{-1}\|_{\mathrm{Op}(j,j')}$$
$$\leq 2\|\mathbf{I}_\infty(\ell,j,j')^{-1}\|_{\mathrm{Op}(j,j')} \overset{(8.30)}{\leq} \frac{2\langle\ell\rangle^{\tau^*}}{\gamma_n^*\langle j-j'\rangle} \leq \frac{\langle\ell\rangle^\tau}{2\gamma_n\langle j-j'\rangle},$$

since by (8.3), (8.4) we have $\tau < \tau^*$ and $\gamma_n^* > 4\gamma_n$.

By the definition (8.13), the claim implies that if $\omega \in \mathcal{G}_n \cap \Omega_{\gamma^*,\tau^*}^{(I)}(\mathbf{u}_n)$ and if the condition (8.28) holds, then $\omega \notin R_{\ell j j'}^-(\mathbf{u}_n)$, hence if $R_{\ell j j'}^-(\mathbf{u}_n) \neq \emptyset$, then

$$\langle j-j'\rangle \min\{j,j'\} < \langle\ell\rangle^{\tau^*}. \tag{8.32}$$

For $\ell = 0$, since $\langle\ell\rangle = \max\{|\ell|,1\} = 1$, the above condition becomes $\langle j - j'\rangle \min\{j,j'\} < 1$ which is violated since $\langle j-j'\rangle \min\{j,j'\} = \max\{|j-j'|,1\} \min\{j,j'\} \geq 1$, therefore $R_{0jj'}(\mathbf{u}_n) = \emptyset$ for any $j \neq j'$. Finally, by (8.32), we may easily deduce that

$$j,j' \lessdot \langle\ell\rangle^{\tau^*}.$$

By similar arguments, it can be proved that if $R_{\ell j j'}^-(\mathbf{u}_n) \neq \emptyset$, then $|j-j'| \lessdot \langle\ell\rangle$ and the proof is concluded. $\qquad\square$

Combining Corollary 8.1 and Lemma 8.3, recalling the formulae (8.11), (8.27), we get

$$\mathcal{A}_0^{(1)} = \bigcup_{\substack{\ell\neq 0\\ j\in\mathbb{N}\\ j\lessdot|\ell|}} Q_{\ell j}(\mathbf{u}_0) \bigcup \bigcup_{\substack{\ell\neq 0\\ j,j'\in\mathbb{N}\\ (\ell,j,j')\neq(0,j,j)\\ |j-j'|\lessdot|\ell|\\ j,j'\lessdot\langle\ell\rangle^{\tau^*}}} R_{\ell j j'}^-(\mathbf{u}_0) \bigcup \bigcup_{\substack{\ell\neq 0\\ j,j'\in\mathbb{N}\\ j,j'\lessdot|\ell|}} R_{\ell j j'}^+(\mathbf{u}_0), \tag{8.33}$$

$$\mathcal{A}_n^{(1)} = \bigcup_{\substack{|\ell|>N_{n-1}\\ j\in\mathbb{N}\\ j\lessdot\langle\ell\rangle}} Q_{\ell j}(\mathbf{u}_n) \bigcup \bigcup_{\substack{|\ell|>N_{n-1}\\ \ell\neq 0\\ j,j'\in\mathbb{N}\\ (\ell,j,j')\neq(0,j,j)\\ j,j'\lessdot\langle\ell\rangle^{\tau^*}}} R_{\ell j j'}^-(\mathbf{u}_n) \bigcup \bigcup_{\substack{|\ell|>N_{n-1}\\ \ell\neq 0\\ j,j'\in\mathbb{N}\\ j,j'\lessdot\langle\ell\rangle}} R_{\ell j j'}^+(\mathbf{u}_n)$$

$$\tag{8.34}$$

$\forall n \geq 1$. The measure of the resonant sets on the right hand side of the latter identities now are estimated separately:

Lemma 8.4. *For $\varepsilon\gamma^{-1}$ small enough, if $Q_{\ell j}(\mathbf{u}_n)$, $R_{\ell j j'}^\pm(\mathbf{u}_n) \neq \emptyset$, then*

$$|Q_{\ell j}(\mathbf{u}_n)|, |R_{\ell j j'}^\pm(\mathbf{u}_n)| \lessdot \gamma\langle\ell\rangle^{-\tau}.$$

Proof. We prove the estimate of the set $R_{\ell j j'}^-(\mathbf{u}_n)$. The other estimates can be proven arguing similarly. Recall that for all $j \in \mathbb{N}$, the 2×2 blocks $\mathbf{D}_j^\infty = m\,j\,\mathbf{I}_j + \widehat{\mathbf{D}}_j^\infty \in \mathcal{S}(\mathbf{E}_j)$, are self-adjoint and Lipschitz continuous with respect to the parameter ω. We set

$$\mathrm{spec}(\widehat{\mathbf{D}}_j^\infty(\omega)) := \left\{r_1^{(j)}(\omega), r_2^{(j)}(\omega)\right\} \quad \text{with} \quad r_1^{(j)}(\omega) \leq r_2^{(j)}(\omega), \tag{8.35}$$

By Lemma 2.5-(i) the functions $\omega \mapsto r_k^{(j)}(\omega)$ are Lipschitz with respect to ω, since

$$|r_k^{(j)}(\omega_1) - r_k^{(j)}(\omega_2)| \leq \|\widehat{\mathbf{D}}_j^\infty(\omega_1) - \widehat{\mathbf{D}}_j^\infty(\omega_2)\| \leq \|\widehat{\mathbf{D}}_j^\infty\|^{\mathrm{lip}}|\omega_1 - \omega_2|$$
$$\overset{(5.108)}{\lessdot} \varepsilon\gamma^{-1}j^{-1}|\omega_1 - \omega_2|. \tag{8.36}$$

Setting also

$$\mathrm{spec}(\mathbf{D}_j^\infty(\omega)) := \left\{\lambda_1^{(j)}(\omega), \lambda_2^{(j)}(\omega)\right\} \quad \text{with} \quad \lambda_1^{(j)}(\omega) \leq \lambda_2^{(j)}(\omega),$$

by Lemma 2.5-(ii) we have that

$$\lambda_k^{(j)}(\omega) = m(\omega)\,j + r_k^{(j)}(\omega), \qquad k = 1, 2. \tag{8.37}$$

By the definition (5.110) and by Lemmata 2.4, 2.5-(ii) the operator $\mathbf{A}_\infty^-(\ell, j, j') : \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j) \to \mathcal{L}(\mathbf{E}_{j'}, \mathbf{E}_j)$ is self-adjoint with respect to the scalar product (2.71) and

$$\mathrm{spec}\left(\mathbf{A}_\infty^-(\ell, j, j'; \omega)\right) = \left\{\omega \cdot \ell + \lambda_k^{(j)}(\omega) - \lambda_{k'}^{(j')}(\omega), \quad k, k' = 1, 2\right\}.$$

Therefore, recalling the definition (8.13) and by Lemma 2.5-(iii) we get

$$R_{\ell j j'}^-(\mathbf{u}_n) \subseteq \bigcup_{k,k'=1}^{2} \widetilde{R}_{\ell j j'}(k, k'), \tag{8.38}$$

where

$$\widetilde{R}_{\ell j,j'}(k, k') := \left\{\omega : |\omega \cdot \ell + \lambda_k^{(j)}(\omega) - \lambda_{k'}^{(j')}(\omega)| < \frac{2\gamma_n\langle j - j'\rangle}{\langle\ell\rangle^\tau}\right\}.$$

We estimate the measure of the set $\widetilde{R}_{\ell j j'}(k, k')$ defined above for all $k, k' = 1, 2$. Since, by Lemma 8.3-(ii), $\ell \neq 0$, we can write

$$\omega = \frac{\ell}{|\ell|}s + v, \quad \text{with} \quad v \cdot \ell = 0.$$

and we define

$$\phi(s) := |\ell|s + \lambda_k^{(j)}(s) - \lambda_{k'}^{(j')}(s), \tag{8.39}$$

where for any $j \in \mathbb{N}$, for all $k = 1, 2$

$$\lambda_k^{(j)}(s) := \lambda_k^{(j)}\left(\frac{\ell}{|\ell|}s + v\right).$$

According to (8.37), (8.36)

$$\lambda_k^{(j)}(s) = m(s)\,j + r_k^{(j)}(s), \qquad |r_k^{(j)}|^{\mathrm{lip}} \lessdot \varepsilon\gamma^{-1}j^{-1}. \tag{8.40}$$

One gets

$$|\phi(s_1) - \phi(s_2)| \geq |\ell||s_1 - s_2| - |m(s_1) - m(s_2)||j - j'|$$
$$- \left(|r_k^{(j)}|^{\mathrm{lip}} + |r_{k'}^{(j')}|^{\mathrm{lip}}\right)|s_1 - s_2|$$
$$\overset{(4.36),(8.40)}{\geq} \left(|\ell| - \varepsilon\gamma^{-1}\langle j - j'\rangle\right)|s_1 - s_2|$$
$$\overset{Lemma\,8.3-(ii)}{\geq} \frac{|\ell|}{2}|s_1 - s_2|$$

for $\varepsilon\gamma^{-1}$ small enough. The above estimate implies that

$$\left|\left\{s: \frac{\ell}{|\ell|}s + v \in \widetilde{R}_{\ell j j'}(k, k')\right\}\right| \lessdot \frac{4\gamma\langle j - j'\rangle}{\langle \ell\rangle^{\tau+1}}$$

and by Fubini Theorem we get

$$|\widetilde{R}_{\ell j j'}(k, k')| \lessdot \frac{4\gamma\langle j - j'\rangle}{\langle \ell\rangle^{\tau+1}}.$$

The claimed estimate follows by recalling (8.38) and using that by Lemma 8.3-(ii), $|j - j'| \lessdot |\ell|$. $\qquad\square$

Lemma 8.5. *For all $n \geq 0$, we get*

$$|\mathcal{A}_n^{(1)}| \lessdot \gamma N_{n-1}^{-1}.$$

Proof. We prove the estimate for $\mathcal{A}_n^{(1)}$ in (8.34) with $n \geq 1$. The estimate for $\mathcal{A}_0^{(1)}$ in (8.33) follows similarly. One has

$$|\mathcal{A}_n^{(1)}| \lessdot \sum_{\substack{|\ell|>N_{n-1}\\ j\in\mathbb{N}\\ j\lessdot\langle\ell\rangle}} |Q_{\ell j}(\mathbf{u}_n)| + \sum_{\substack{|\ell|>N_{n-1}\\ j,j'\in\mathbb{N}\\ (\ell,j,j')\neq(0,j,j)\\ |j-j'|\lessdot\langle\ell\rangle\\ j,j'\lessdot\langle\ell\rangle^{\tau^*}}} |R_{\ell j j'}^{-}(\mathbf{u}_n)| + \sum_{\substack{|\ell|>N_{n-1}\\ j,j'\in\mathbb{N}\\ j,j'\lessdot\langle\ell\rangle}} |R_{\ell j j'}^{+}(\mathbf{u}_n)|$$

$$\overset{Lemma\,8.4}{\lessdot} \gamma\left(\sum_{\substack{|\ell|>N_{n-1}\\ j\in\mathbb{N}\\ j\lessdot\langle\ell\rangle}} \frac{1}{\langle\ell\rangle^\tau} + \sum_{\substack{|\ell|>N_{n-1}\\ j,j'\in\mathbb{N}\\ j,j'\lessdot\langle\ell\rangle^{\tau^*}}} \frac{1}{\langle\ell\rangle^\tau} + \sum_{\substack{|\ell|>N_{n-1}\\ j,j'\in\mathbb{N}\\ j,j'\lessdot\langle\ell\rangle}} \frac{1}{\langle\ell\rangle^\tau}\right)$$

$$\lessdot \gamma \sum_{|\ell|>N_{n-1}} \left(\frac{1}{\langle\ell\rangle^{\tau-1}} + \frac{1}{\langle\ell\rangle^{\tau-2\tau^*}} + \frac{1}{\langle\ell\rangle^{\tau-2}}\right) \overset{(8.2)}{\lessdot} \gamma N_{n-1}^{-1} \qquad (8.41)$$

which is the claimed estimate. $\qquad\square$

Estimate of $\mathcal{A}_n^{(2)}$. By the same arguments used to prove Lemma 8.5, one may deduce that the sets $\mathcal{A}_n^{(2)}$ defined in (8.10) satisfy the lemma below.

Lemma 8.6.

$$|\mathcal{A}_n^{(2)}| \lessdot \gamma N_{n-1}^{-1}, \qquad \forall n \geq 0.$$

Proof of Theorem 8.1. The theorem follows by (8.5), (8.6), (8.8), by Lemmata 8.6, 8.5, using that the series $\sum_{n \geq 0} N_{n-1}^{-1} < +\infty$, since $N_n = N_0^{\chi^n}$, $N_{-1} := 1$ with $N_0 > 1$.

9. Proof of the main theorems concluded

PROOF OF THEOREM 3.1. Set $\gamma = \varepsilon^a$, with $0 < a < 1$. Then $\varepsilon \gamma^{-1} = \varepsilon^{1-a}$ and hence the smallness condition (7.5) is fullfilled by taking ε small enough. By the estimate (7.8) we deduce that the sequence $(\mathbf{u}_n)_{n \geq 1}$ is a Cauchy sequence with respect to the norm $\| \cdot \|_{s_0+\mu_1}^{\mathrm{Lip}(\gamma)}$ on the set $\mathcal{G}_\infty$ defined in (8.1), then it converges to a limit $\mathbf{u}_\infty$, which satisfies the estimate

$$\|\mathbf{u}_\infty\|_{s_0+\mu_1}^{\mathrm{Lip}(\gamma)} \lessdot \varepsilon \gamma^{-1} \lessdot \varepsilon^{1-a} \overset{\varepsilon \to 0}{\to} 0. \tag{9.1}$$

Moreover, by Theorem 7.1-$(\mathcal{P}2)_n$, we deduce that for all $\omega \in \mathcal{G}_\infty$, $\mathcal{F}(\mathbf{u}_\infty) = 0$ and by Theorem 8.1, since $\gamma = \varepsilon^a$, we get

$$|\Omega \backslash \mathcal{G}_\infty| \to 0 \qquad \text{as} \qquad \varepsilon \to 0.$$

PROOF OF THEOREM 1.1. Recalling the splitting (3.4), (3.5), if $(v_0(\varphi), p_0(\varphi))$ are the solution of the Eq. (3.5) found in Lemma 3.1, and by applying Theorem 3.1, we get that for any $\omega \in \mathcal{G}_\infty$ the function $(v_\infty, p_\infty) = (v_0, p_0) + \mathbf{u}_\infty$ satisfies $F(v_\infty, p_\infty) = 0$. Furthermore, choosing $\gamma = \varepsilon^a$, with $0 < a < 1/2$, by (3.6)

$$\|v_0\|_s \lessdot \varepsilon^{1-2a} \overset{\varepsilon \to 0}{\to} 0, \quad \|p_0\|_s \lessdot \varepsilon^{1-a} \overset{\varepsilon \to 0}{\to} 0,$$

hence, (1.10) follows by recalling (9.1). Finally (1.9) follows since

$$\int_{\mathbb{T}^{\nu+1}} v_\infty(\varphi, x) \, d\varphi \, dx = \int_{\mathbb{T}^\nu} v_0(\varphi) \, d\varphi, \quad \int_{\mathbb{T}^{\nu+1}} p_\infty(\varphi, x) \, d\varphi \, dx = \int_{\mathbb{T}^\nu} p_0(\varphi) \, d\varphi$$

and by Lemma 3.1, v_0 and p_0 have zero average in φ, this concludes the proof of Theorem 1.1.

9.1. Linear stability

In this section we prove Theorem 1.2. The linearized equation on a quasi-periodic Sobolev function $\mathbf{v}(\omega t, x) = (v(\omega t, x), p(\omega t, x))$, with $v, p \in H^S(\mathbb{T}^{\nu+1}, \mathbb{R})$ has the form (1.11). Since the linearized vector field

$$L(\omega t) := \begin{pmatrix} 0 & 1 \\ a(\omega t)\partial_{xx} + \mathcal{R}(\omega t) & 0 \end{pmatrix}$$

(recall (1.12), (1.13)) preserves the space of the functions with zero average in x, the equation (1.11) can be splitted into the two systems

$$\begin{cases} \partial_t \widehat{v}_0 = \widehat{p}_0 \\ \partial_t \widehat{p}_0 = 0 \end{cases} \tag{9.2}$$

$$\begin{cases} \partial_t \widehat{u} = \widehat{\psi} \\ \partial_t \widehat{\psi} = a(\omega t)\partial_{xx}\widehat{u} + \mathcal{R}(\omega t)[\widehat{u}], \end{cases} \tag{9.3}$$

where, recalling (3.1), $\widehat{v}_0 = \pi_0 \widehat{v}$, $\widehat{p}_0 = \pi_0 \widehat{p}$, $\widehat{u} = \pi_0^\perp \widehat{v}$, $\widehat{\psi} = \pi_0^\perp \widehat{p}$. By assumption, the initial datum $\widehat{p}^{(0)}$ has zero average in x, hence the solution of the system (9.2) is given by $\widehat{v}_0(t) = const$, $\widehat{p}_0(t) = 0$, for all $t \in \mathbb{R}$, implying that the system projected on the zero Fourier mode in x is stable. It remains to establish the stability for the system projected on the x-zero average functions (9.3). By Lemma 4.3, by (4.27), by Lemmata 4.6, 5.3, using also Lemma 2.13, there exists $\overline{\mu} > 0$ such that for $S > s_0 + \overline{\mu}$, for any $s_0 \leq s \leq S - \overline{\mu}$, for any $\omega \in \Omega_\infty^{2\gamma}$ (see (5.109)), the linear and continuous maps $\mathcal{T}_1(\omega t) := \mathcal{S}(\omega t) \circ \mathcal{B}$ and $\mathcal{T}_2(\omega t) := \mathcal{V}(\omega t) \circ \Phi_\infty(\omega t)$ satisfy

$$\mathcal{T}_1(\omega t) : \mathbf{H}_0^{s-\frac{1}{2}}(\mathbb{T}_x) \to H_0^s(\mathbb{T}_x, \mathbb{R}) \times H_0^{s-1}(\mathbb{T}_x, \mathbb{R}),$$
$$\mathcal{T}_1(\omega t)^{-1} : H_0^s(\mathbb{T}_x, \mathbb{R}) \times H_0^{s-1}(\mathbb{T}_x, \mathbb{R}) \to \mathbf{H}_0^{s-\frac{1}{2}}(\mathbb{T}_x),$$
$$\mathcal{T}_2(\omega t)^{\pm 1} : \mathbf{H}_0^{s-\frac{1}{2}}(\mathbb{T}_x) \to \mathbf{H}_0^{s-\frac{1}{2}}(\mathbb{T}_x).$$

Setting $A\mathbf{h}(t,x) = \mathbf{h}(t + \alpha(\omega t), x)$, $A^{-1}\mathbf{h}(\tau + \tilde{\alpha}(\omega \tau), x)$, where α and $\tilde{\alpha}$ are given in Sect. 4.3, by the results of Sects. 4.1–4.4, by Theorem 5.2 and using the arguments of Sect. 2.2, we get that a curve $\widehat{\mathbf{u}}(t) = (\widehat{u}(t), \widehat{\psi}(t)) \in H_0^s(\mathbb{T}_x, \mathbb{R}) \times H_0^{s-1}(\mathbb{T}_x, \mathbb{R})$ is a solution of the PDE (9.3) if and only if

$$\mathbf{h}(t) = \begin{pmatrix} h(t) \\ \overline{h}(t) \end{pmatrix} = \mathcal{T}_2(\omega t)^{-1} \circ A^{-1} \circ \mathcal{T}_1(\omega t)^{-1}\widehat{\mathbf{u}}(t) \in \mathbf{H}_0^{s-\frac{1}{2}}(\mathbb{T}_x)$$

is a solution of the PDE

$$\begin{cases} \partial_t h = -\mathrm{i}\mathcal{D}_\infty^{(1)} h \\ \partial_t \overline{h} = \mathrm{i}\overline{\mathcal{D}}_\infty^{(1)} \overline{h} \end{cases} \tag{9.4}$$

where $\mathcal{D}_\infty^{(1)}$ is defined by (5.115), (5.107). Using that $\mathcal{D}_\infty^{(1)}$ is a 2×2-block diagonal operator, it is straightforward to verify that the commutator $[\mathcal{D}_\infty^{(1)}, |D|^s] = 0$. Furthermore, using the self-adjointness of $\mathcal{D}_\infty^{(1)}$ one sees by a standard energy estimate that $\partial_t \|h(t, \cdot)\|_{H_x^s}^2 = 0$, implying that

$$\|h(t, \cdot)\|_{H_x^s} = const, \qquad \forall t \in \mathbb{R}.$$

Arguing as in the proof of Theorem 1.5 in [6] one concludes that $\|\widehat{\mathbf{u}}(t, \cdot)\|_{H_x^s \times H_x^{s-1}} \leq_s \|\widehat{\mathbf{u}}(0, \cdot)\|_{H_x^s \times H_x^{s-1}}$ for all $t \in \mathbb{R}$, which proves the linear stability of (9.3) and the proof of Theorem 1.2 is concluded.

Acknowledgements

References

[1] Alazard, T., Baldi, P.: Gravity capillary standing water waves. Arch. Rat. Mech. Anal. **217**(3), 741–830 (2015)

[2] Arosio, A., Spagnolo, S.: Global solutions of the Cauchy problem for a nonlinear hyperbolic equation. In: Brezis, H., Lions, J.L. (eds.) Nonlinear PDEs and Their Applications, Collége de France Seminar, vol. VI, 126, Research Notes on Mathematics 109. Pitman, Boston (1984)

[3] Arosio, A., Panizzi, S.: On the well-posedness of the Kirchhoff string. Trans. Am. Math. Soc. **348**(1), 305–330 (1996)

[4] Baldi, P.: Periodic solutions of forced Kirchhoff equations. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5) **8**, 117–141 (2009)

[5] Baldi, P.: Periodic solutions of fully nonlinear autonomous equations of Benjamin–Ono type. Ann. Inst. H. Poincaré (C) Anal. Non Linéaire **30**(1), 33–77 (2013)

[6] Baldi, P., Berti, M., Montalto, R.: KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation. Math. Ann. **359**(1–2), 471–536 (2014)

[7] Baldi, P., Berti, M., Montalto, R.: KAM for autonomous quasi-linear perturbations of KdV. Ann. Inst. H. Poincaré (C) Anal. Non Linéaire **33**, 1589–1638 (2016)

[8] Baldi, P., Berti, M., Montalto, R.: KAM for autonomous quasi-linear perturbations of mKdV. Boll. Unione Mat. Ital. **9**, 143–188 (2016)

[9] Bernstein, S.N.: Sur une classe d' équations fonctionelles aux dérivées partielles. Izv. Akad. Nauk SSSR Ser. Mat. **4**, 17–26 (1940)

[10] Berti, M., Biasco, P., Procesi, M.: KAM theory for the Hamiltonian DNLW. Ann. Sci. Éc. Norm. Supér. (4) **46**(fascicule 2), 301–373 (2013)

[11] Berti, M., Biasco, P., Procesi, M.: KAM theory for the reversible derivative wave equation. Arch. Ration. Mech. Anal. **212**, 905–955 (2014)

[12] Berti, M., Bolle, P.: Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential. Nonlinearity **25**, 2579–2613 (2012)

[13] Berti, M., Bolle, P.: Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential. Eur. J. Math. **15**, 229–286 (2013)

[14] Berti, M., Corsi, L., Procesi, M.: An abstract Nash–Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds. Commun. Math. Phys. **334**(3), 1413–1454 (2015)

[15] Berti, M., Montalto, R.: Quasi-periodic Standing Wave Solutions of Gravity Capillary Standing Water Waves. arXiv:1602.02411 (2016)

[16] Bourgain J.: Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE. Int. Math. Res. Not. (11) (1994)

[17] Bourgain, J.: Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. Ann. Math. **148**, 363–439 (1998)

[18] Bourgain, J.: Periodic solutions of nonlinear wave equations. In: Harmonic Analysis and Partial Differential Equations. Chicago Lectures in Mathematics, pp. 69–97. University of Chicago Press (1999)

[19] Bourgain, J.: Green's Function Estimates for Lattice Schrödinger Operators and Applications, Annals of Mathematics Studies 158. Princeton University Press, Princeton (2005)

[20] Craig, W.: Problèmes de petits diviseurs dans les équations aux dérivées partielles, Panoramas et Synthèses, 9. Société Mathématique de France, Paris (2000)

[21] Craig, W., Wayne, C.E.: Newton's method and periodic solutions of nonlinear wave equation. Commun. Pure Appl. Math. **46**, 1409–1498 (1993)

[22] Chierchia, L., You, J.: KAM Tori for 1D nonlinear wave equations with periodic boundary conditions. Commun. Math. Phys **211**, 497–525 (2000)

[23] D'Ancona, P., Spagnolo, S.: Global solvability for the degenerate Kirchhoff equation with real analytic data. Invent. Math. **108**, 247–262 (1992)

[24] Dickey, R.W.: Infinite systems of nonlinear oscillation equations related to the string. Proc. Am. Math. Soc. **23**(3), 459–468 (1969)

[25] Eliasson, L.H., Kuksin, S.: On reducibility of Schrödinger equations with quasiperiodic in time potentials. Commun. Math. Phys. **286**, 125–135 (2009)

[26] Eliasson, L.H., Kuksin, S.: KAM for non-linear Schrödinger equation. Ann. Math. **172**, 371–435 (2010)

[27] Feola, R.: KAM for Quasi-linear Forced Hamiltonian NLS. arXiv:1602.01341 (2016)

[28] Feola, R., Procesi, M.: Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations. J. Differ. Equ. **259**(7), 3389–3447 (2015)

[29] Iooss, G., Plotnikov, P.I., Toland, J.F.: Standing waves on an infinitely deep perfect fluid under gravity. Arch. Ration. Mech. Anal. **177**(3), 367–478 (2005)

[30] Iooss, G., Plotnikov, P.I.: Small divisor problem in the theory of three-dimensional water gravity waves. Mem. Am. Math. Soc. 200(940) (2009)

[31] Iooss, G., Plotnikov, P.I.: Asymmetrical three-dimensional travelling gravity waves. Arch. Ration. Mech. Anal. **200**(3), 789–880 (2011)

[32] Lions, J.L.: On some questions in boundary value problems of mathematical physics. In: de la Penha, G.M., Medeiros, L.A. (eds.) Contemporary developments in continuum mechanics and PDEs. North-Holland, Amsterdam (1978)

[33] Kappeler, T., Pöschel, J.: KAM and KdV. Springer, Berlin (2003)

[34] Klainermann, S., Majda, A.: Formation of singularities for wave equations including the nonlinear vibrating string. Commun. Pure Appl. Math. **33**, 241–263 (1980)

[35] Kuksin, S.: A KAM theorem for equations of the Korteweg–de Vries type. Rev. Math. Math. Phys. **10**(3), 1–64 (1998)

[36] Kuksin, S., Pöschel, J.: Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. Ann. Math. **2**(143), 149–179 (1996)

[37] Lax, P.: Development of singularities of solutions of nonlinear hyperbolic partial differential equations. J. Math. Phys. **5**, 611–613 (1964)

[38] Liu, J., Yuan, X.: Spectrum for quantum Duffing oscillator and small-divisor equation with large-variable coefficient. Commun. Pure Appl. Math. **63**(9), 1145–1172 (2010)

[39] Liu, J., Yuan, X.: A KAM theorem for Hamiltonian partial differential equations with unbounded perturbations. Commun. Math. Phys. **307**, 629–673 (2011)

[40] Manfrin, R.: On the global solvability of Kirchhoff equation for non-analytic initial data. J. Differ. Equ. **211**, 38–60 (2005)

[41] Matsuyama, T., Ruzhansky, M.: Global well-posedness of the Kirchhoff equation and Kirchhoff systems. Springer proceedings in mathematics and statistics. Anal. Methods Interdiscip. Appl. **116**, 81–96 (2014)

[42] Pokhozhaev, S.I.: On a class of quasilinear hyperbolic equations, Mat. Sbornik **96**, 152–166 (1975) (English transl.: Mat. USSR Sbornik 25, 145–158 (1975))

[43] Procesi, M., Xu, X.: Quasi-Töplitz functions in KAM theorem. SIAM J. Math. Anal. **45**(4), 2148–2181 (2013)

[44] Rabinowitz, P.H.: Periodic solutions of nonlinear hyperbolic partial differential equations, part I and II. Commun. Pure Appl. Math. **20**, 145–205 (1967)

[45] Rabinowitz, P.H.: Periodic solutions of nonlinear hyperbolic partial differential equations, part I and II. Commun. Pure Appl. Math. **22**, 15–39 (1969)

Riccardo Montalto
University of Zürich
Wintherthurerstrasse 190
8057 Zurich
Switzerland
e-mail: riccardo.montalto@math.uzh.ch