Long Time Dynamics of Quasi-linear Hamiltonian Klein–Gordon Equations on the Circle

Roberto Feola[1] · Filippo Giuliani[2]

Abstract

We consider a class of Hamiltonian Klein–Gordon equations with a quasilinear, quadratic nonlinearity under periodic boundary conditions. For a large set of masses, we provide a precise description of the dynamics for an open set of small initial data of size ε showing that the corresponding solutions remain close to oscillatory motions over a time scale $\varepsilon^{-\frac{9}{4}+\delta}$ for any $\delta > 0$. The key ingredients of the proof are normal form methods, para-differential calculus and a modified energy approach.

Keywords Long time dynamics · Normal forms · Quasilinear Klein–Gordon · Long time approximation

Contents

✉ Filippo Giuliani
  filippo.giuliani@polimi.it

  Roberto Feola
  roberto.feola@uniroma3.it

1   Dipartimento di Matematica e Fisica, Universita' degli studi di Roma Tre, Largo San Leonardo Murialdo 1, 00144 Rome, Italy

2   Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy

🖄 Springer

1 Introduction

In this paper we study the long-time dynamics of the quasi-linear Klein–Gordon equation on the compact interval $[0, 2\pi]$, under periodic boundary conditions

$$u_{tt} - u_{xx} + mu + N(u) = 0 \quad u(t, x) \in \mathbb{R}, \quad t \in \mathbb{R}, \quad x \in \mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z}), \qquad (1.1)$$

where

$$m \in [1, 2], \qquad N(u) := f_2(u, u_x, u_{xx}) \qquad (1.2)$$

and $f_2 \in C^\infty(\mathbb{R}^3; \mathbb{R})$ is a homogeneous polynomial of degree 2.
We consider Klein–Gordon equations possessing a Hamiltonian structure, namely we require that the nonlinearity N in (1.2) has the form

$$N(u) = \partial_u G(u, u_x) - \frac{d}{dx}\left(\partial_{u_x} G(u, u_x)\right), \qquad (1.3)$$

where $G \in C^\infty(\mathbb{R}^2; \mathbb{R})$ is a homogeneous polynomial of degree 3.

The origin $u = 0$ is an elliptic equilibrium for the equation (1.1). The long-time stability of this fixed point has been proved in the seminal works [23–25] by Delort for a positive measure set of masses m. For long-time stability here we mean the following property: given $N \geq 1$ there exists $\varepsilon_0 = \varepsilon_0(N)$ such that for all solutions $u(t)$ with initial datum u_0 such that $\|u_0\|_{H^s} < \varepsilon$ there exist constants $c = c(N), C = C(N) > 0$, independent of ε, such that $\|u(t)\|_{H^s} \leq C\varepsilon$ for all $t \in [0, c\,\varepsilon^{-N}]$. This is a result of orbital stability for the elliptic equilibrium $u = 0$ and it does not provide information about the dynamics of small data solutions. While the theory of long-time stability is quite well-established for semilinear PDEs on some compact manifolds (without trying to be exhaustive [2, 3, 21, 22, 30], and the more recent results [5, 15]), there are few results concerning quasi-linear PDEs using normal form techniques. We mention for instance [9, 10, 34], and the partial results (namely, where the stability time scale is ε^{-m} for some *fixed* $m \in \mathbb{Q}, m \geq 1$) [6, 14, 26, 33, 37, 42, 44]. We point out that results of long / global in time existence for Klein–Gordon equations as (1.1) have been proved also in the Euclidean setting, and we cite for instance [43, 46, 52].

The aim of the present paper is to provide an accurate description of the long-time dynamics for solutions arising from an open set of small initial data. With long-time behavior (or dynamics) we mean the evolution of the system (1.1) beyond the local time scale. Since the equation (1.1) is quadratic, the local time scale for solutions with initial data of size ε here corresponds to ε^{-1}. If the equation does not present 3-wave resonant interactions, as for instance in the pure gravity water waves system and KdV equation, then one can expect to obtain a quadratic lifespan ε^{-2}, over which the solutions remains of size ε. Our aim is to provide dynamical information beyond this time scale, where we have to deal with the effect of higher order resonances.

We prove that the long-time evolution of initial data, within the above mentioned open set, is almost recurrent, in the sense that these solutions are very close to nonlinear oscillations, over the range of time $[0, c\,\varepsilon^{-9/4+\delta}]$ for any $\delta > 0$ and some $c > 0$ independent of ε.

The oscillatory motions that we consider are supported on finite dimensional tori and they can be minimal or not, in the sense that they may fill densely the underlying torus or not. It is possible that certain motions (of the former kind) may be continued to periodic or quasi-periodic solutions by using KAM techniques. However, these would provide the description of the dynamics of a zero measure set of data in phase space. In any case, at the best of our knowledge, such KAM results are not still available in literature for the quasi-linear Klein–Gordon equation on the circle. Concerning the existence of periodic and quasi-periodic solutions for semilinear Klein–Gordon equations on the circle we refer to [7, 12, 16, 19, 20, 39, 41, 49, 55] and to [8, 11, 13, 17, 54] (and reference therein) for the higher dimensional case.

An interesting research direction in the study of Hamiltonian PDEs concerns the orbital stability of invariant objects rather than fixed points, such as plane waves and quasi-periodic tori. About that, we mention [31, 45, 47, 56]. We also quote the stability result [18] for traveling waves of the Burger-Hilbert equation.

We point out that our result is not an orbital stability result, in the sense that we do not only prove the existence of solutions that stay close to certain embedded tori, but also that the solutions of equation (1.1) follow (in the sense of Sobolev norms) periodic or quasi-periodic orbits that can be explicitly computed.

To explain the main difficulties in proving our result, we can think to set action-angle variables on the finite dimensional embedded torus complemented by cartesian coordinates in the normal infinite dimensional directions. In the following we refer to tangential and normal directions with respect to the torus.

To prove that solutions $u(t)$ of the Klein–Gordon equation starting close to the tours follow closely some of the orbits $\varphi(t)$ supported on it we shall study the evolution of the error function $R := u - \varphi$.

One of the main issue is that the tangential and normal dynamics of R are coupled. We will use normal form techniques to almost decouple them at linear level.

Since the finite dimensional torus is embedded in an infinite dimensional phase space, we certainly need to control the deviation between $u(t)$ and $\varphi(t)$ along infinitely many normal directions. In this analysis the quasi-linear nature of the equation provides the main issue. We deal with this by combining para-differential calculus techniques and using a modified energy method.

A further issue comes from the analysis of tangential directions. Indeed, we need to control the deviation of the trajectories both in the actions and angles directions.

As it is well known, even for integrable Hamiltonian systems, the linear dynamics along the angle directions is unstable. To overcome this problem we need to impose certain non-

resonance conditions on the linear frequencies of oscillations. This accounts to exclude just a zero measure set of masses.

We will continue this discussion, giving more details, in Sect. 1.2.

1.1 Main Result

To state our main result is useful to look at equation (1.1) as a first order system, which, in appropriate complex coordinates $Z := (z^+, z^-)$ (see (2.10)), reads as

$$
\begin{cases}
\dot{z}^+ = i\Lambda z^+ + \frac{i}{\sqrt{2}}\Lambda^{-1/2} N\left(\Lambda^{-1/2}\left(\frac{z^+ + z^-}{\sqrt{2}}\right)\right), & \Lambda := \sqrt{-\partial_{xx} + m} \\
\dot{z}^- = i\Lambda z^- - \frac{i}{\sqrt{2}}\Lambda^{-1/2} N\left(\Lambda^{-1/2}\left(\frac{z^+ + z^-}{\sqrt{2}}\right)\right).
\end{cases}
\tag{1.4}
$$

We look for solutions in the Sobolev real subspaces

$$
H^s := \left(H^s(\mathbb{T};\mathbb{C}) \times H^s(\mathbb{T};\mathbb{C})\right) \cap \mathcal{U}, \qquad \mathcal{U} := \{(z^+, z^-) \in L^2(\mathbb{T};\mathbb{C}^2) \ : \ \overline{z^+} = z^-\},
\tag{1.5}
$$

with $s \geq 0$. With abuse of notation we denote by $\|\cdot\|_{H^s}$ (see (2.3)) both the norm on $H^s(\mathbb{T};\mathbb{C})$ and on the product space H^s.

Fix $\varepsilon > 0$ small, $N \in \mathbb{N}$, a *symmetric* subset $S := \{j_1, \ldots, j_N\} \subset \mathbb{Z}$ (meaning that $j \in S$ implies $-j \in S$) and let $(\xi, \theta) \in \mathcal{O}^N \times \mathbb{T}^N$ for some compact subset $\mathcal{O} \subset \mathbb{R}_+ := (0, \infty]$ containing the origin. We consider small amplitude, oscillating functions of the form

$$
\varepsilon \varphi(t) = \varepsilon \varphi_{\xi,\theta}(t) := \sum_{j \in S} \varepsilon \sqrt{\xi_j}\, e^{i(\theta_j + jx + \omega_j t)}, \qquad \xi_{j_i} := \xi_i, \quad \theta_{j_i} := \theta_i, \quad i = 1, \ldots, N,
\tag{1.6}
$$

where the frequencies ω_j, $j \in S$ satisfy

$$
\sup_{j \in S} |\omega_j - \Lambda(j)| \lesssim \varepsilon^2, \qquad \Lambda(j) := \sqrt{j^2 + m}.
$$

These oscillating functions will be obtained as nonlinear corrections of small amplitude linear solutions[1] through normal form methods.

Our goal is to prove that there is plenty of solutions of (1.4) which are "well-approximated" by functions as in (1.6) for a very *long time scale* provided that the ω_j are chosen as *suitable* corrections of the linear frequencies of oscillation. More precisely, the main result is the following.

Theorem 1.1 *Let* $n \in \mathbb{N}$ *and set* $N = 2n + 1$. *Consider the equation* (1.4)–(1.3) *and the subset* $S = \{-n, -(n-1), \ldots, n-1, n\} \subset \mathbb{Z}$. *There exist a symmetric matrix* $\mathsf{C} \in \mathbb{R}^{N \times N}$ *such that for almost all* $m \in [1, 2]$ *there is* $s_0 > 1/2$ *such that for any* $s \geq s_0$ *the following holds. There is* $\varepsilon_0 = \varepsilon_0(s, S)$ *such that for all* $0 < \varepsilon < \varepsilon_0$ *and for any* $\xi \in \mathbb{R}_+^N, \theta \in \mathbb{T}^N$ *there exists an open set* $\mathcal{U}_{\xi,\theta}^N \subseteq H^s(\mathbb{T};\mathbb{C})$ *such that for any* $z_0 \in \mathcal{U}_{\xi,\theta}^N$ *the solution* $z(t)$ *of* (1.4) *with* $z(0) = z_0$ *belongs to* $C([0, T]; H^s)$, *with* $T = \varepsilon^{-\left(\frac{9}{4}\right)^-}$ *and satisfies*

$$
\sup_{t \in [0,T]} \|z(t) - \varepsilon\varphi(t)\|_{H^s} \lesssim_s \varepsilon^2,
\tag{1.7}
$$

where $\varepsilon\varphi$ *is defined as in* (1.6) *with frequencies*

$$
\omega_j = \omega_j(\xi) := \sqrt{j^2 + m} + \varepsilon^2 \sum_{k \in S} \mathsf{C}_{jk}\xi_k, \qquad j \in S.
$$

---

[1] Indeed, for $\omega_j = \Lambda(j)$ functions of the form (1.6) are solutions of the linearized problem of (1.4) at zero.

Remark 1.2 We remark that we mostly exploit the symmetry of the set S in order to avoid certain resonant interactions. We need the additional property

$$j \notin S \quad \Leftrightarrow \quad |j| > \max_{k \in S}\{|k|\}$$

only for the construction of a modified energy in Sect. 5.2.2.

The matrix $\mathbb{C}$, which provides the corrections to the linear frequencies of oscillation $\Lambda(j)$, is completely determined by the choice of the set S and by the coefficients of the nonlinearity N.

When the matrix $\mathbb{C}$ is invertible one can modulate the frequencies $(\omega_j)_{j \in S}$ through the choice of the amplitudes $(\xi_j)_{j \in S}$. When the PDE presents no external/physical parameters this is the approach implemented to find small amplitude quasi-periodic solutions, which bifurcate from quasi-periodic functions of the form (1.6) for which diophantine non-resonance conditions are imposed (we mention for instance [1, 32, 38, 40]).

To prove our result these conditions are not required. In principle we could consider resonant oscillating motions as

$$\varepsilon\varphi(t, x) = \varepsilon \left(\sqrt{\xi_j}e^{\mathrm{i}jx} + \sqrt{\xi_{-j}}e^{-\mathrm{i}jx} \right) e^{\mathrm{i}(\theta_j + \omega_j t)},$$

where we assumed that $\theta_j = \theta_{-j}$ and $\omega_j = \omega_{-j}$. The existence of such functions depend on the choice of the nonlinearity. Therefore, our result allow to control the dynamics of more initial data with respect to the ones considered by KAM theory, but only for finite (even if long) time. On the other hand, Birkhoff normal form methods allow to control the evolution of any small (enough) initial data over any polynomial time scale, but they do not provide a description of the orbits as the one provided by Theorem 1.1.

Description of the Long-time Dynamics and Time Scales Theorem 1.1 provides an accurate description of the long-time dynamics for general Hamiltonian Klein–Gordon equations with quadratic nonlinearities. The strength of this result relies on the fact that we are able to approximate solutions of the Klein–Gordon equation over long time with explicit oscillating functions $\varphi(t)$ which are solutions of *integrable* finite dimensional systems. Such systems are obtained as the restriction of the truncated Hamiltonian, after some steps of normalization, to some finite dimensional subspace. Since these ones are integrable in the sense of Liouville, all the trajectories lie on (finite dimensional) tori and they are conjugated to a linear flow $\theta \mapsto \omega t + \theta$. Hence, once the nonlinearity has been fixed, the evolution of the functions $\varphi(t)$ is completely determined, even for infinite time. We remark again that the orbits $\varphi(t)$ may lie also on *resonant* tori.

In the framework of Hamiltonian PDEs, a similar result has been provided in the work by Bernier–Grébert [4] for the generalized KdV and Benjamin–Ono equations on the circle. A description of the dynamics is provided for any polynomial time scale ε^{-r} and for many small initial data by using the rational normal form method [5]. However, apart from the fact that the above mentioned PDEs are not quasi linear equations, the information provided about the long time behavior of small data solutions differ substantially from the ones given by Theorem 1.1, indeed:

1. The solutions of the generalized KdV and Benjamin–Ono equations are approximated by oscillating functions whose expression is not explicit. For instance, it is not clear whether these are periodic, quasi-periodic or none of the previous.
2. The vicinity among the solutions of the PDEs and their approximating oscillating functions is not ensured in the phase space (where the initial data belong). Indeed, the

approximation (as (1.7)) is provided in a Sobolev space which is less regular than the space of initial data (with a loss of one derivative).

With respect to [4] here we pay these extra information on the dynamics by considering "less" initial data (even if they still form open sets in the phase space) and shorter time scales. Concerning the latter, the fact that we are not able to provide a result as Theorem 1.1 for any polynomial time scale is not merely a technical issue.

We observe that just to obtain an orbital stability result, one would require the integrability of the normalized Hamiltonian at any order, and this requirement is not satisfied by the Klein–Gordon system (and it is not expected in general from non integrable equations, even on one dimensional spatial domains). The normalized Hamiltonian of Klein–Gordon at degree six is not integrable, it has only half of the required constants of motion, usually called *super-actions*. This is due to the dispersion relation of Klein–Gordon, and more precisely on the multiplicity of the linear eigenvalues $i\Lambda(j)$.

Therefore the study of the dynamics of the normal form at higher orders can be very complicated to study and we cannot expect to give an accurate description of it, especially when the number of degrees of freedom (that is N in Theorem 1.1) is large.

For this reason, in this situation, we have an expected optimal time scale for which the approximation (1.7) holds. Certainly the approximation (1.7) provided by Theorem 1.1 cannot hold true beyond the time scale ε^{-3^-}. One can see it from the proof of Lemma 6.5, which provides energy estimates for the linearized part of the equation (1.12) for low frequencies. We now explain this issue and how we obtain the time scale $\varepsilon^{-\frac{9}{4}^-}$ in the main theorem. It is useful to denote by $\varepsilon^{-2-\sigma}$ the time scale of the approximation and highlight the restriction on the parameter $\sigma > 0$.

We start by noticing that a sub problem of our analysis consists in controlling the deviation among the trajectories of a finite dimensional integrable Hamiltonian which is quadratic in the action variables I (up to linear terms in actions which play no role in this analysis); we refer to Hamiltonian (6.10) where $I_j = |w_j|^2$.

The key problem is to control the linearization of the Hamiltonian vector field at one of the trajectories considered. This linearized problem approximately describes the time evolution of the first variation of the action-angle variables.

The main issue is that the variation in the angles direction grows linearly in time. The first bound in (6.19), which can be satisfied by taking $\sigma < 1$, shows how this fact reflects on the estimate of the energy for low frequencies. We remark that this is a finite dimensional issue (independent of the PDE's context that we are considering) and it cannot be overcome neither by the presence (if any) of integrable Hamiltonian terms of higher order. Indeed, in a small amplitude regime, the quadratic term in actions provides the main contribution. Hence, this gives the upper bound ε^{-3^-} on the time of validity of the approximation (1.7).

A further restriction on the parameter σ comes from the fact that the normalized Hamiltonian of the Klein–Gordon is integrable only up to degree six. Since we are interested in approximating solutions of the Klein–Gordon equation with trajectories of an *integrable* system (hence, which are completely known), we need to consider $\varepsilon\varphi$ in (1.7) as a solution of the Klein–Gordon equation up to remainders $O(\varepsilon^5)$, see (4.7). This means that the residual term $\text{Res}_{\mathcal{H}}(\varepsilon\varphi)$ in the equation (1.12) cannot be made too small. This affects the last bound in (6.19), where we remark that the parameter β has to be chosen $> 2 + 2\sigma$ in order to control the nonlinear terms of the equation (1.12). If we relax the constraints on β we need to enforce the limitations on σ, and viceversa. Then, in order to satisfy all the constraints on β and σ, it turns out that σ has to be taken strictly less than $1/4$. This gives the time scale $\varepsilon^{-\frac{9}{4}^-}$

provided in Theorem 1.1. Clearly this loss, with respect to the time scale ε^{-3^-}, depends on our method and we reserve to improve it as matter of future investigations.

The Set of Initial Data The open set $\mathcal{U}_{\xi,\theta}^N$ is the neighborhood of an embedded N-dimensional torus in the phase space. Then the set of initial data for which we can describe the long-time dynamics forms an open covering of a continuous family of embedded N-dimensional tori.

The following proposition provides a characterization of this set.

Proposition 1.3 *Under the assumptions of Theorem 1.1 there exists a diffeomorphism $\mathcal{F}: B_\varepsilon(H^s) \to H^s$, where $B_\varepsilon(H^s)$ is the ball of radius $\varepsilon > 0$ centered at the origin, such that the following holds. Defining*

$$\mathcal{B}_S^N := \left\{ \varepsilon \varphi_{\xi,\theta}(0) := \sum_{j \in S} \varepsilon \sqrt{\xi_j} e^{i(\theta_j + jx)} \ : \ (\xi, \theta) \in \mathcal{O}^N \times \mathbb{T}^N \right\}, \tag{1.8}$$

one has that the union of the open sets $\mathcal{U}_{\xi,\theta}^N$ is a covering of an $\varepsilon^{7/2}$-neighbourhood of the embedded manifold $\mathcal{F}(\mathcal{B}_S^N) \subseteq H^s$, namely

$$\mathcal{A}_{\varepsilon^{7/2}}^N := \left\{ u := u_1 + u_2 \in H^s \ : \ u_1 \in \mathcal{F}(\mathcal{B}_S^N), \|u_2\|_{H^s} \leq \frac{\varepsilon^{7/2}}{2} \right\} \subseteq \bigcup_{(\xi,\theta) \in \mathcal{O}^N \times \mathbb{T}^N} \mathcal{U}_{\xi,\theta}^N. \tag{1.9}$$

Remark 1.4 We note that the union of the sets $\mathcal{B}_S^N$, as N and S vary respectively in $\mathbb{N}$ and in the symmetric subsets of $\mathbb{Z}$, contains the set of trigonometric polynomials.

The result above implies that solutions evolving from a *large* family of initial data can be approximated, in the Sobolev topology, by oscillatory motions thanks to Theorem 1.1. We remark that, since trigonometric polynomials are *dense* in Sobolev spaces, in principle taking N larger one can consider more and more functions $u_1 \in \mathcal{F}(\mathcal{S}^N)$ in (1.9), being the map $\mathcal{F}$ a diffeomorphism. The price to pay in order to apply Theorem 1.1 is to shrink to zero the ball of radius ε in which the approximation holds.

Relation with Modulation Theory The functions $\varepsilon \varphi$ shall be chosen as good approximate solutions of the quasi-linear Klein–Gordon equation (1.1). An efficient way to construct approximate solutions for a Hamiltonian PDE is to perform a normalization of the Hamiltonian and consider orbits of its resonant part.

A different method to study the long time evolution and stability of oscillatory solutions to partial differential equations is provided by the modulation theory. In this framework we mention the work by Düll [27], where the NLS approximation is provided for a quadratic quasi-linear Klein–Gordon equation on $\mathbb{R}$ (we also mention [28, 29, 50, 51], and reference therein, for similar results on other PDEs). The description of the dynamics in these papers is given over the cubic time scale ε^{-2}, where the only resonant effects are given by 3-wave resonances. The crucial difference with respect to our analysis is that here we deal with higher order resonances taking advantage of considering the PDE on the torus and of its Hamiltonian structure.

1.2 Strategy of the Proof

First of all we remark that equation (1.4), in view of the assumption (1.3), can be written as

$$\partial_t Z = X_H(Z), \quad Z = \begin{bmatrix} z \\ \bar{z} \end{bmatrix}, \tag{1.10}$$

where X_H is the Hamiltonian vector field of the KG equation in complex coordinates and H admits the expansion (see (2.12))

$$H(Z) = H^{(2)}(Z) + H^{(3)}(Z), \qquad H^{(2)}(Z) = \int_{\mathbb{T}} \bar{z} \Lambda z \, dx,$$

where $H^{(3)}$ is homogeneous of degree 3 in the variables Z, Z_x. Precise details on the Hamiltonian structure will be given in Sect. 2.1. We now discuss the strategy of the proof of our main result.

(1) Approximate solution.
The first step is to find a suitable approximate solution of the KG equation, of the form (1.6), supported on some symmetric set of Fourier modes $S \subset \mathbb{Z}$. It turns out that, for our aims, the choice $\omega_j \equiv \Lambda(j)$ (namely to consider *linear* solutions of KG) provides a rough approximation. In order to determine a better approximation through oscillatory motion we perform some steps of "weak" Birkhoff normal form. In other words we construct an invertible, bounded, symplectic map Φ_B with the following properties:

• The transformed Hamiltonian has the form

$$\mathcal{H} := H \circ \Phi_B^{-1} = H^{(2)} + \mathcal{H}_{\mathrm{res}}^{(4,0)} + \mathcal{H}^{(>)} + \mathfrak{R}^{(\geq 6)}, \qquad (1.11)$$

where

$$\mathcal{H}_{\mathrm{res}}^{(4,0)}(W) = \tfrac{1}{2} \mathtt{C} I \cdot I \qquad I := (|w_j|^2)_{j \in S}, \qquad \Phi_B(Z) := W = \left[\begin{smallmatrix} w \\ \bar{w} \end{smallmatrix} \right],$$

with $\mathtt{C}$ some symmetric $N \times N$ matrix, $\mathfrak{R}^{(\geq 6)}$ has a zero of order 6 at the origin and $\mathcal{H}^{(>)}$ is an Hamiltonian function which vanishes on the finite dimensional subspace $\mathcal{U}_S = \{w_n = 0 \ \ n \notin S\}$. We remark that the matrix $\mathtt{C}$ is completely determined in terms of the coefficients of the original Hamiltonian H and the set of modes S. This is the content of Proposition 3.9

• Functions $\varepsilon \varphi$ of the form (1.6) are approximate solutions for the transformed equation

$$\partial_t W = X_{\mathcal{H}}(W), \qquad \Phi_B(Z) = W,$$

in the sense that

$$\mathrm{Res}_{\mathcal{H}}(\varepsilon \varphi) := -\varepsilon \partial_t \varphi + X_{\mathcal{H}}(\varepsilon \varphi) \sim \varepsilon^5.$$

The precise estimates are given in Lemma 4.1.

(2) The error function. Once we constructed the approximate oscillatory solution we introduce the "error" function

$$\varepsilon^\beta V = W - \varepsilon \varphi, \qquad \beta > 2.$$

Our aim is to control the norms of the function V, over a long time interval, to measure the deviation between the true solution W of the KG equation in the Birkhoff coordinates and the approximate solution $\varepsilon \varphi$. Thanks to estimates on the map Φ_B we are able to show that (see Proposition 4.2) $Z - \varepsilon \varphi$ remains *small* over the time scale $T \sim \varepsilon^{-(9/4)^-}$, i.e. satisfies (1.7), provided that

$$\sup_{t \in [0,T]} \varepsilon^\beta \|V\|_{H^s} \leq 2\varepsilon^\beta, \qquad T \sim \varepsilon^{-(9/4)^-}.$$

The core of the paper is then to show that the above estimate holds true. We have that V solves the problem (see Sect. 4.2)

$$\dot{V} = dX_{\mathcal{H}}(\varepsilon \varphi)[V] + \varepsilon^\beta \mathcal{Q}(\varepsilon \varphi)[V, V] + \varepsilon^{-\beta} \mathrm{Res}_{\mathcal{H}}(\varepsilon \varphi), \qquad (1.12)$$

where

$$\mathcal{Q}(\varepsilon\varphi)[V, V] := \int_0^1 (1 - t) \, d^2 X_{\mathcal{H}}(\varepsilon\varphi + t\varepsilon^\beta V)[V, V] \, dt \,.$$

In order to study the evolution of the Sobolev norms of solutions of (1.12) we provide a priori energy estimates for such equation, recalling that the term $\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)$ is perturbative thanks to the choice of $\varepsilon\varphi$. Due to the quasi-linear nature of the nonlinearity we first need to study the pseudo-differential structure of the linear operator $dX_{\mathcal{H}}(\varepsilon\varphi)[\cdot]$ and provide a para-differential formulation of the nonlinear term $\mathcal{Q}(\varepsilon\varphi)[V, V]$. This is the content of Sect. 4.3.

(3) High/low frequencies analysis and normal forms. We first note that given V, defined for all $t \in [0, T]$, $\partial_t \|\varepsilon^\beta V\|_s^2 = \partial_t (D^{2s} V, V)_{L^2}$ where D is the Fourier multiplier $D = \sqrt{-\partial_{xx} + 1}$ and $(\cdot, \cdot)_{L^2}$ the L^2-scalar product. Using equation (1.12) we have that $\partial_t \|\varepsilon^\beta V\|_s^2$ is controlled from above by the sum of three terms

$$\partial_t \|\varepsilon^\beta V\|_s^2 \lesssim A_{dX_{\mathcal{H}}} + A_{\mathcal{Q}} + A_{Res},$$

with the following properties:

- A_{Res} is the contribution coming from the residual term. In particular we show that

$$A_{Res} \lesssim \varepsilon^\beta \sup_{[0,T]} \|V\|_{H^s} \sup_{[0,T]} \|\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)\|_{H^s} \lesssim \varepsilon^{\beta+5},$$

The smallness of the residual $\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)$ is obtained by construction of $\varepsilon\varphi$.

- The term $A_{\mathcal{Q}}$ is the contribution coming from the quadratic part in V. The energy estimates for the quadratic terms $\mathcal{Q}(\varepsilon\varphi)$ are obtained thanks to the use of para-differential techniques. So we get

$$A_{\mathcal{Q}} \lesssim \varepsilon^{3\beta} \sup_{[0,T]} \|V\|_{H^s}^3.$$

Its smallness is guaranteed by taking β large enough.

- The term $A_{dX_{\mathcal{H}}}$ is the most delicate one and comes from the contribution of the linearized operator $dX_{\mathcal{H}}(\varepsilon\varphi)$ at $\varepsilon\varphi$. The main issue is to control the effect of this linear part. We remark that a direct estimate of this term provides a bound like

$$A_{dX_{\mathcal{H}}} \lesssim \varepsilon^{2\beta+1} \sup_{[0,T]} \|V\|_{H^s}^2.$$

An estimate like this allows a control of the norm of V only over the trivial time scale of size ε^{-1}. This is due to the fact that our equation has a quadratic nonlinearity. Of course this estimate is not sufficient to our aims.

In Sects. 5–6 we improve such estimate to show the stability of V over a longer time scale.

The analysis of the linearized operator is done by performing normal form methods. Since we want to reach stability beyond the nonlinear time scale we shall deal with resonances of order (three and) four, i.e. solutions of

$$\sigma_1 \Lambda(j_1) + \sigma_2 \Lambda(j_2) + \sigma_3 \Lambda(j_3) + \sigma_4 \Lambda(j_4) = 0, \qquad \sigma_i = \pm, \ j_i \in \mathbb{Z}.$$

By picking the mass $m \in [1, 2]$ in a positive Lebesgue measure subset, we can prove that (see Lemma 3.8) these resonances are of the form (up to permutations)

$$j_1 = j_2, \quad j_3 = j_4, \qquad \sigma_1 + \sigma_2 = 0, \qquad \sigma_3 + \sigma_4 = 0.$$

Then the Hamiltonian terms Fourier supported on these resonances (*resonant terms*) cannot be eliminated and we have to guarantee that they do not let growth much the Sobolev norms

of R. In principle it is not clear how to deal with the resonant terms Fourier supported on

$$j_1, j_3 \in \text{supp}(\widehat{V}) \quad \text{and} \quad j_2, j_4 \in \text{supp}(\widehat{\varphi}).$$

Indeed such resonant monomials are not dependent just by the actions $|\widehat{\varphi}(j)|^2$, $|\widehat{V}(j)|^2$. For instance they may have the form

$$\widehat{\varphi}(j_1) \, \overline{\widehat{V}(j_1)} \, \widehat{\varphi}(j_3) \, \overline{\widehat{V}(j_3)}, \tag{1.13}$$

which are clearly angle-dependent and they could generate hyperbolicity and unstable phenomena.

To overcome this problem we exploit the fact that the Fourier support S of $\varepsilon\varphi$ is compact. The idea is to split the study of the time evolution of low and high modes of the error function V. Thanks to the weak normal form of step (1) the projection on the low frequencies of the linearized vector field at $\varepsilon\varphi$ of the Hamiltonian terms $\mathcal{H}^{(>)}$ and $\mathfrak{R}^{(\geq 6)}$ vanish. Then the linear part of the equation for the low frequency arise from the linearization of an integrable finite dimensional Hamiltonian $H^{(2)} + \mathcal{H}^{(4,0)}_{\text{res}}$.

Despite the integrability, the linearized problem at an approximate solution of an integrable system presents instability in some directions (the angles directions). Thanks to the fact that we pushed the non-normalized part of the Hamiltonian at higher orders and to the choice of the initial data (6.2), we are able to control the evolution of the low modes for long time.

In normal form coordinates the linear dynamics of low and high frequencies are decoupled, and terms like (1.13) do not appear in the linear problem for high frequencies. In Sect. 5, we provide energy estimates for the projection of V onto high frequencies where we deal with the quasi-linear nature of the equation by using normal form techniques and energy methods.

2 Functional Setting

2.1 Hamiltonian Formalism

We denote by $H^s(\mathbb{T}) := H^s(\mathbb{T}; \mathbb{C})$ the usual Sobolev space of functions $\mathbb{T} \ni x \mapsto u(x) \in \mathbb{C}$. We consider the following Fourier expansion of a function $u(x)$, $x \in \mathbb{T}$,

$$u(x) = \frac{1}{\sqrt{2\pi}} \sum_{n \in \mathbb{Z}} \widehat{u}(n) e^{inx}, \qquad \widehat{u}(n) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} u(x) e^{-inx} \, dx. \tag{2.1}$$

We shall also use the notation

$$u_n^{+1} := u_n := \widehat{u}(n) \quad \text{and} \quad u_n^{-1} := \overline{u_n} := \overline{\widehat{u}(n)}. \tag{2.2}$$

For $\xi \in \mathbb{R}$ we define $\langle \xi \rangle := \sqrt{1 + |\xi|^2}$ and we denote by $\langle D \rangle$ the Fourier multiplier with symbol $\langle \xi \rangle$, i.e.

$$\langle D \rangle e^{inx} = \langle n \rangle e^{inx}, \quad n \in \mathbb{Z}.$$

We endow $H^s(\mathbb{T}; \mathbb{C})$ with the norm

$$\|u(\cdot)\|_{H^s}^2 := \sum_{j \in \mathbb{Z}} \langle j \rangle^{2s} |u_j|^2. \tag{2.3}$$

Moreover, for $r \in \mathbb{R}^+$, we denote by $B_r(X)$ the ball of the Banach space X with radius r centred at the origin.

Notation. We shall use the notation $A \lesssim B$ to denote $A \leq CB$ where C is a positive constant depending on parameters fixed once for all, for instance d and s. We use the notation $\lesssim_q$ to highlight the dependence of the constant C on some parameter q. We use the notation $A \sim B$ to denote that $C_1 A \leq B \leq C_2 A$ for some constants $C_1, C_2 > 0$.

The Klein–Gordon equation in (1.1) reads

$$\partial_{tt} u + \Lambda^2 u + N(u) = 0 \tag{2.4}$$

where Λ is the Fourier multiplier defined by linearity as

$$\Lambda \, e^{in \cdot x} = \Lambda(n) e^{in \cdot x}, \qquad \Lambda(n) := \sqrt{|n|^2 + m}, \qquad \forall n \in \mathbb{Z}. \tag{2.5}$$

Introducing the variable $v = \dot{u} = \partial_t u$ we can rewrite equation (2.4) as

$$\dot{u} = -v, \qquad \dot{v} = \Lambda^2 u + N(u). \tag{2.6}$$

By (1.3) we note that (2.6) can be written in the Hamiltonian form

$$\partial_t \begin{bmatrix} u \\ v \end{bmatrix} = X_{H_{\mathbb{R}}}(u, v) = J \begin{pmatrix} \partial_u H_{\mathbb{R}}(u, v) \\ \partial_v H_{\mathbb{R}}(u, v) \end{pmatrix}, \qquad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

where ∂ denotes the L^2-gradient of the Hamiltonian function

$$H_{\mathbb{R}}(u, v) = \int_{\mathbb{T}} \left(\frac{1}{2} v^2 + \frac{1}{2} (\Lambda^2 u) u + G(u, u_x) \right) dx, \tag{2.7}$$

on the phase space $H^1(\mathbb{T}; \mathbb{R}) \times L^2(\mathbb{T}; \mathbb{R})$. Indeed we have

$$dH_{\mathbb{R}}(u, v) \begin{bmatrix} \widehat{u} \\ \widehat{v} \end{bmatrix} = -\omega_{\mathbb{R}} \left(X_{H_{\mathbb{R}}}(u, v), \begin{bmatrix} \widehat{u} \\ \widehat{v} \end{bmatrix} \right) \tag{2.8}$$

for any $(u, v), (\widehat{u}, \widehat{v})$ in $H^1(\mathbb{T}; \mathbb{R}) \times L^2(\mathbb{T}; \mathbb{R})$, where $\omega_{\mathbb{R}}$ is the non-degenerate symplectic form

$$\omega_{\mathbb{R}}(W_1, W_2) := \int_{\mathbb{T}} (u_1 v_2 - v_1 u_2) dx, \qquad W_1 := \begin{bmatrix} u_1 \\ v_1 \end{bmatrix}, W_2 := \begin{bmatrix} u_2 \\ v_2 \end{bmatrix}.$$

The Poisson bracket between two Hamiltonians $H_{\mathbb{R}}, G_{\mathbb{R}} : H^1(\mathbb{T}; \mathbb{R}) \times L^2(\mathbb{T}; \mathbb{R}) \to \mathbb{R}$ is defined as

$$\{H_{\mathbb{R}}, G_{\mathbb{R}}\} = \omega_{\mathbb{R}}(X_{H_{\mathbb{R}}}, X_{G_{\mathbb{R}}}). \tag{2.9}$$

We introduce the complex symplectic variables

$$\begin{pmatrix} z \\ \overline{z} \end{pmatrix} = \mathcal{C} \begin{pmatrix} u \\ v \end{pmatrix} := \frac{1}{\sqrt{2}} \begin{pmatrix} \Lambda^{\frac{1}{2}} u + i\Lambda^{-\frac{1}{2}} v \\ \Lambda^{\frac{1}{2}} u - i\Lambda^{-\frac{1}{2}} v \end{pmatrix}, \qquad \begin{pmatrix} u \\ v \end{pmatrix} = \mathcal{C}^{-1} \begin{pmatrix} z \\ \overline{z} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Lambda^{-\frac{1}{2}} (z + \overline{z}) \\ -i\Lambda^{\frac{1}{2}} (z - \overline{z}) \end{pmatrix}, \tag{2.10}$$

where Λ is the Fourier multiplier defined in (2.5). Then the system (2.6) reads as in (1.4). Notice that (1.4) can be written in the Hamiltonian form

$$\partial_t Z = X_H(Z) = iJ \begin{pmatrix} \partial_u H(Z) \\ \partial_{\overline{z}} H(Z) \end{pmatrix} = \begin{pmatrix} i\partial_{\overline{z}} H(Z) \\ -i\partial_z H(Z) \end{pmatrix}, \qquad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad Z = \begin{bmatrix} z \\ \overline{z} \end{bmatrix}, \tag{2.11}$$

with Hamiltonian function (see (2.7))

$$H(Z) = H_{\mathbb{R}}(\mathcal{C}^{-1} Z) = H^{(2)}(Z) + H^{(3)}(Z),$$

$$H^{(2)}(Z) = \int_{\mathbb{T}} \overline{z} \Lambda z \, dx, \qquad H^{(3)}(Z) := \int_{\mathbb{T}} G\Big(\frac{\Lambda^{-1/2}(z + \overline{z})}{\sqrt{2}}, \frac{\Lambda^{-1/2}(z_x + \overline{z}_x)}{\sqrt{2}} \Big) dx, \tag{2.12}$$

and where $\partial_{\bar{z}} = (\partial_{\text{Re}(z)} + i\partial_{\text{Im}(z)})/2$, $\partial_z = (\partial_{\text{Re}(z)} - i\partial_{\text{Im}(z)})/2$. Notice that

$$X_H = \mathcal{C} \circ X_{H_{\mathbb{R}}} \circ \mathcal{C}^{-1} \tag{2.13}$$

and that (using (2.10))

$$dH(z, \bar{z})\begin{bmatrix} h \\ h \end{bmatrix} = (dH_{\mathbb{R}})(u, v)[\mathcal{C}^{-1}\begin{bmatrix} h \\ h \end{bmatrix}] \overset{(2.8),(2.13)}{=} -\omega\left(X_H(Z), \begin{bmatrix} h \\ h \end{bmatrix}\right) \tag{2.14}$$

for any $h \in H^2(\mathbb{T}; \mathbb{C})$ and where the two form ω is given by the pullback of $\omega_{\mathbb{R}}$ through $\mathcal{C}^{-1}$. In complex variables the Poisson bracket in (2.9) reads as

$$\{H, G\} := \omega(X_H, X_G) = i\left(\int_{\mathbb{T}} \partial_z G \partial_{\bar{z}} H - \partial_{\bar{z}} G \partial_z H\right) dx, \tag{2.15}$$

where we set $H = H_{\mathbb{R}} \circ \mathcal{C}^{-1}$, $G = G_{\mathbb{R}} \circ \mathcal{C}^{-1}$.

2.2 Preliminaries

In this section we introduce some classes of operators that we shall consider along the paper.

2.2.1 Basic Para-Differential Calculus

We follow the notation of [36]. We denote by $\mathcal{N}_s^m$, $s \geq 0$, $m \in \mathbb{R}$, the spaces of functions $\mathbb{T} \times \mathbb{R} \ni (x, \xi) \to a(x, \xi)$ of symbols defined by the norms

$$|a|_{\mathcal{N}_s^m} := \sup_{0 \leq \alpha+\beta \leq s} \sup_{\xi \in \mathbb{R}} \langle \xi \rangle^{-m+\beta} \|\partial_\xi^\beta \partial_x^\alpha a(x, \xi)\|_{L^2}. \tag{2.16}$$

The constant $m \in \mathbb{R}$ indicates the *order* of the symbols, while s denotes its regularity. The following result is a consequence of item (i) of Lemma 2.3 in [36].

Lemma 2.1 *Let* $m_1, m_2 \in \mathbb{R}$, $s > 1/2$ *and* $a \in \mathcal{N}_s^{m_1}$, $b \in \mathcal{N}_s^{m_2}$. *One has*

$$|ab|_{\mathcal{N}_s^{m_1+m_2}} + |\{a, b\}|_{\mathcal{N}_{s-1}^{m_1+m_2-1}} + |\sigma(a, b)|_{\mathcal{N}_{s-2}^{m_1+m_2-2}} \lesssim |a|_{\mathcal{N}_s^{m_1}} |b|_{\mathcal{N}_s^{m_2}} \tag{2.17}$$

where

$$\{a, b\} := (\partial_\xi a)(\partial_x b) - (\partial_x a)(\partial_\xi b), \tag{2.18}$$

$$\sigma(a, b) := (\partial_{\xi\xi} a)(\partial_{xx} b) - 2(\partial_{x\xi} a)(\partial_{\xi x} b) + (\partial_{xx} a)(\partial_{\xi\xi} b). \tag{2.19}$$

The Weyl quantization. For a symbol $a(x, \xi)$ in $\mathcal{N}_s^m$ we define its (Weyl) quantization as

$$Op^{\text{W}}(a(x, \xi))h := \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z}} e^{ijx} \sum_{k \in \mathbb{Z}} \hat{a}\left(j - k, \frac{j+k}{2}\right) \hat{h}(k) \tag{2.20}$$

where $\hat{a}(\eta, \xi)$ denotes the Fourier transform of $a(x, \xi)$ in the variable $x \in \mathbb{T}$.
Smoothed symbols. Let $0 < \epsilon < 1/2$ and consider a smooth function $g : \mathbb{R} \to [0, 1]$

$$g(\xi) = \begin{cases} 1 & \text{if } |\xi| \leq 5/4 \\ 0 & \text{if } |\xi| \geq 8/5 \end{cases} \quad \text{and define} \quad \chi(x) = \chi(\epsilon; \xi) := g(|\xi|/\epsilon). \tag{2.21}$$

Given $a(x, \xi) \in \mathcal{N}_s^m$ we define

$$a_\chi(x, \xi) = \mathcal{F}^{-1}\left(\hat{a}(\eta, \xi) \chi(|\eta|/\langle \xi \rangle)\right) \tag{2.22}$$

where $\mathcal{F}$ denotes the Fourier transform in x.

We denote by

$$Op^{BW}(a(x,\xi)) := Op^W(a_\chi(x,\xi)). \tag{2.23}$$

Lemma 2.2 *Let $s_0 > 1$. Then the following holds.*

(i) Let $m \in \mathbb{R}$, $a \in \mathcal{N}_{s_0}^m$, then for any $s \in \mathbb{R}$

$$\|Op^{BW}(a)[h]\|_{H^{s-m}} \lesssim_s |a|_{\mathcal{N}_{s_0}^m}\|h\|_{H^s}, \qquad \forall h \in H^s(\mathbb{T};\mathbb{C}). \tag{2.24}$$

(ii) Let $m_1, m_2 \in \mathbb{R}$, $a \in \mathcal{N}_{s_0+4}^{m_1}$, $b \in \mathcal{N}_{s_0+4}^{m_2}$, then

$$Op^{BW}(a) \circ Op^{BW}(b) = Op^{BW}(ab + \tfrac{1}{2i}\{a,b\} - \tfrac{1}{8}\sigma(a,b)) + R(a,b), \tag{2.25}$$

where $R(a,b)$ is a remainder satisfying, for any $s \in \mathbb{R}$,

$$\|R(a,b)h\|_{H^{s-m_1-m_2+3}} \lesssim \|h\|_{H^s}|a|_{\mathcal{N}_{s_0+4}^{m_1}}|b|_{\mathcal{N}_{s_0+4}^{m_2}}. \tag{2.26}$$

In particular

$$\left[Op^{BW}(a), Op^W(b)\right] = Op^{BW}(\tfrac{1}{i}\{a,b\}) + R(a,b), \tag{2.27}$$

with $R(a,b)$ as in (2.26).

(iii) Let $\rho \geq 0$ and $a = a(x)$, $b = b(x) \in H^{\rho+s_0}(\mathbb{T};\mathbb{C})$ (independent of $\xi \in \mathbb{R}$) then, for any $s \in \mathbb{R}$,

$$\|(Op^{BW}(a) \circ Op^{BW}(b) - Op^{BW}(ab))h\|_{H^{s+\rho}} \lesssim \|h\|_{H^s}\|a\|_{H^{\rho+s_0}}\|b\|_{H^{\rho+s_0}}. \tag{2.28}$$

Proof By the definition of the semi-norm (2.16) we have

$$|\widehat{a}(j,\xi)| \lesssim_s \langle\xi\rangle^m\langle j\rangle^{-s_0}|a|_{\mathcal{N}_{s_0}^m}, \quad \forall j,\xi \in \mathbb{Z}. \tag{2.29}$$

We have (recall (2.23), (2.20))

$$\|Op^{BW}(a)[h]\|_{H^{s-m}}^2 \leq \sum_{j\in\mathbb{Z}}\left(\sum_{|j-k|\leq 2^{-1}|j+k|}\left|a\left(j-k,\frac{j+k}{2}\right)\right||\widehat{h}(k)|\langle j\rangle^{s-m}\right)^2.$$

We note that $|j-k| \leq 2^{-1}|j+k|$ implies that $\langle j\rangle \sim \langle k\rangle$. Then by using (2.29) and Young's inequality we have

$$\|Op^{BW}(a)[h]\|_{H^{s-m}}^2 \lesssim_s |a|_{\mathcal{N}_{s_0}^m}^2 \sum_{j\in\mathbb{Z}}\left(\sum_{|j-k|\leq 2^{-1}|j+k|}\frac{|j+k|^m}{|j-k|^{s_0}\langle j\rangle^m}|\widehat{h}(k)|\langle k\rangle^s\right)^2$$

$$\lesssim_s |a|_{\mathcal{N}_{s_0}^m}^2 \sum_{j\in\mathbb{Z}}\left(\sum_{|j-k|\leq 2^{-1}|j+k|}|j-k|^{-s_0}|\widehat{h}(k)|\langle k\rangle^s\right)^2$$

$$\lesssim_s |a|_{\mathcal{N}_{s_0}^m}^2\|h\|_{H^s}^2.$$

The items (ii)–(iii) follows by Proposition 2.5 in [36]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

2.2.2 Multilinear Operators

We are interested in studying properties of multilinear operators of the form

$$R_p[z_{p+1}] = R_p(z_1, \ldots, z_p)[z_{p+1}] : (C^\infty(\mathbb{T}; \mathbb{C}))^{p+1} \to C^\infty(\mathbb{T}; \mathbb{C}), \quad p = 1, 2,$$

$$R_p[z_{p+1}] = \frac{1}{\sqrt{2\pi}} \sum_{j_{p+1}, j \in \mathbb{Z}} (R_p)_j^{j_{p+1}} \widehat{z}(j_{p+1}) e^{\mathrm{i} j x}, \tag{2.30}$$

with

$$(R_1)_j^{j_2} = \frac{1}{\sqrt{2\pi}} \sum_{\substack{j_1, j_2 \in \mathbb{Z}, \sigma_1 \in \{\pm\} \\ \sigma_1 j_1 + j_2 = j}} r_1^{\sigma_1}(j, j_1, j_2) \widehat{z}_1^{\sigma_1}(j_1), \tag{2.31}$$

$$(R_2)_j^{j_3} = \frac{1}{2\pi} \sum_{\substack{j_1, j_2, j_3 \in \mathbb{Z}, \sigma_1, \sigma_2 \in \{\pm\} \\ \sigma_1 j_1 + \sigma_2 j_2 + j_3 = j}} r_2^{\sigma_1, \sigma_2}(j, j_1, j_2, j_3) \widehat{z}_1^{\sigma_1}(j_1) \widehat{z}_2^{\sigma_2}(j_2), \tag{2.32}$$

where the coefficients $r_p^{\sigma_1 \ldots \sigma_p}(j, j_1, \ldots, j_{p+1}) \in \mathbb{C}$ for any $j, j_1, \ldots, j_{p+1} \in \mathbb{Z}$, $p = 1, 2$. We introduce the following notation: given $j_1, \ldots, j_q \geq 0$, $q \geq 2$ we define

$$\mathrm{max}_i\{j_1, \ldots, j_q\} = i-\text{th largest among } j_1, \ldots, j_q. \tag{2.33}$$

We need the following definition.

Definition 2.3 Let $\rho \in \mathbb{R}$, $p = 1, 2$, $q \in \mathbb{N}$ and $0 < r \leq 1$.
(i) **(Multilinear operators).** We say that a $(p+1)$-linear map R_p as in (2.30)–(2.32) belongs to the class $\mathbf{M}_p^{-\rho}$ if there is $\mu \geq 0$ such that for any $j_1, \ldots, j_{p+1} \in \mathbb{Z}$, $\sigma_1, \ldots, \sigma_p \in \{\pm\}$ one has

$$|r_p^{\sigma_1 \ldots \sigma_p}(\zeta, j_1, \ldots, j_{p+1})| \lesssim \frac{\mathrm{max}_2\{\langle j_1 \rangle, \ldots, \langle j_{p+1} \rangle\}^{\mu + |\rho|}}{\mathrm{max}_1\{\langle j_1 \rangle, \ldots, \langle j_{p+1} \rangle\}^\rho}, \tag{2.34}$$

for $\zeta := \sigma_1 j_1 + \cdots + \sigma_p j_p + j_{p+1}$.
(ii) **(Non-homogenenous operators).** We denote $\mathbf{NH}_q^{-\rho}$ the class of maps $(\varphi, V, u) \mapsto \mathcal{R}(\varphi, V)[u]$ defined on $B_r(H^\mu(\mathbb{T}))^2 \times H^\mu(\mathbb{T})$ for some $\mu > 1/2$, linear in u, such that the following holds. For any $s \geq \mu$, and $(\varphi, V) \in B_r(H^\mu(\mathbb{T}))^2 \cap H^s(\mathbb{T})$, $u \in H^s(\mathbb{T})$ one has

$$\|\mathcal{R}(\varphi, V)u\|_{H^{s+\rho}} \lesssim_{s, \rho} (\|\varphi\|_{H^s}^q + \|V\|_{H^s}) \|u\|_{H^s},$$

$$\|d_\varphi(\mathcal{R}(\varphi, V))(u)h\|_{H^{s+\rho}} \lesssim_{s, \rho} (\|\varphi\|_{H^s}^{q-1} + \|V\|_{H^s}) \|u\|_{H^s} \|h\|_{H^s}, \tag{2.35}$$

$$\|d_V(\mathcal{R}(\varphi, V))(u)h\|_{H^{s+\rho}} \lesssim_{s, \rho} \|u\|_{H^s} \|h\|_{H^s}.$$

(iii) We denote by $\Sigma_1^{-\rho}[r, 3]$ the space of operators $R(\varphi)[\cdot]$ of the form

$$R(\varphi) = R_1(\varphi) + R_2(\varphi) + R_{\geq 3}(\varphi, V),$$

where $R_j \in \mathbf{M}_j^{-\rho}$, $j = 1, 2$, and $R_{\geq 3} \in \mathbf{NH}_3^{-\rho}$.

We now prove that the operators defined above extend as continuous maps on the Sobolev spaces.

Lemma 2.4 *Let $\rho \in \mathbb{R}$. Consider a multilinear operator $R \in \mathbf{M}_p^{-\rho}$ and μ introduced in Def. 2.3-(i). Then, there is $s_0 = s_0(\mu) > 1/2$ such that, for $s \geq s_0 + |\rho|$, the map R in (2.30) with*

coefficients satisfying (2.34) extends as a continuous map from $(H^s(\mathbb{T}; \mathbb{C}))^p \times H^s(\mathbb{T}; \mathbb{C})$ to $H^{s+\rho}(\mathbb{T}; \mathbb{C})$. Moreover one has

$$\| R(u_1, \ldots, u_p)[u_{p+1}] \|_{H^{s+\rho}} \lesssim \| u_{p+1} \|_{H^s} \prod_{i=1}^{p} \| u_i \|_{H^s} . \tag{2.36}$$

Proof We follow Lemma 2.5 in [33]. We give the proof only in the case $p = 2$. The case $p = 1$ will follow similarly. By (2.3) we have

$$\| R(u_1, u_2)[u_3] \|_{H^{s+\rho}}^2$$

$$\leq \sum_{\xi \in \mathbb{Z}} \langle \xi \rangle^{2(s+\rho)} \left(\sum_{\substack{j_1, j_2, j_3 \in \mathbb{Z} \\ j_1+j_2+j_3=\xi}} |r(\xi, j_1, j_2, j_3)| |\widehat{u}_1(j_1)| |\widehat{u}_2(j_2)| |\widehat{u}_3(j_3)| \right)^2$$

$$\overset{(2.34)}{\lesssim} \sum_{\xi \in \mathbb{Z}} \left(\sum_{\substack{j_1, j_2, j_3 \in \mathbb{Z} \\ j_1+j_2+j_3=\xi}} \langle \xi \rangle^s \max_2 \{ \langle j_1 \rangle, \langle j_2 \rangle, \langle j_3 \rangle \}^{\mu+|\rho|} |\widehat{u}_1(j_1)| |\widehat{u}_2(j_2)| |\widehat{u}_3(j_3)| \right)^2$$

$$:= I + II + III , \tag{2.37}$$

where I, II, III are the terms in (2.37) which are supported respectively on indexes such that

$$\max_1 \{ \langle j_1 \rangle, \langle j_2 \rangle, \langle j_3 \rangle \} = \langle j_1 \rangle, \quad \text{or} \quad \max_1 \{ \langle j_1 \rangle, \langle j_2 \rangle, \langle j_3 \rangle \} = \langle j_2 \rangle, \quad \text{or}$$
$$\max_1 \{ \langle j_1 \rangle, \langle j_2 \rangle, \langle j_3 \rangle \} = \langle j_3 \rangle.$$

Consider for instance the term III. By using the Young inequality for sequences we deduce

$$III \lesssim \| (\langle j_1 \rangle^{\mu+|\rho|} \widehat{u}_1(j_1)) * (\langle j_2 \rangle^\mu \widehat{u}_2(j_2)) * (\langle j_3 \rangle^s \widehat{u}_3(j_3)) \|_{\ell^2}$$
$$\lesssim \| u_1 \|_{H^{1/2+\epsilon+\mu+|\rho|}} \| u_2 \|_{H^{1/2+\epsilon+\mu}} \| u_3 \|_{H^s} ,$$

for $\epsilon > 0$, which is the is the (2.36) taking $s_0 > \mu + 1/2$. The bounds of I and II are similar. □

Now we introduce class of linear operators obtained by linearizing multilinear and non-homogenous maps at a C^∞-function φ. Such operators are pseudo-differential according to the definition given in subsection 2.2.1.

Definition 2.5 Let $m \in \mathbb{R}$, $s > 1/2$, $q \in \mathbb{N}$ and $0 < r \leq 1$.
(i) **(Linear symbols).** We denote by $\mathbf{SM}_1^m$ the class of maps $\varphi \in C^\infty(\mathbb{T}; \mathbb{C}) \to a_1(\varphi; \cdot) \in \mathcal{N}_s^m$ such that it has the form

$$a_1(\varphi; x, \xi) = \sum_{j \in \mathbb{Z}, \sigma \in \{\pm\}} a_1^\sigma(j, \xi) \varphi_j^\sigma e^{\sigma i j x} , \tag{2.38}$$

and

$$|\partial_\xi^\beta (a_1)^\sigma(j, \xi)| \lesssim \langle j \rangle^\mu \langle \xi \rangle^{m-\beta} , \tag{2.39}$$

for some $\mu \geq 0$.

(*ii*) **(Quadratic symbols).** We denote by $\mathbf{SM}_2^m$ the class of maps $\varphi \in C^\infty(\mathbb{T}; \mathbb{C}) \to$ $a_2(\varphi; \cdot) \in \mathcal{N}_s^m$ such that it has the form

$$
a_2(\varphi; x, \xi) = \sum_{j_1, j_2 \in \mathbb{Z}, \sigma \in \{\pm\}} a_2^{\sigma\sigma}(j_1, j_2, \xi) \varphi_{j_1}^\sigma \varphi_{j_2}^\sigma e^{i\sigma(j_1+j_2)x}
$$
$$
+ \sum_{j_1, j_2 \in \mathbb{Z}} a_2^{+-}(j_1, j_2, \xi) \varphi_{j_1} \overline{\varphi_{j_2}} e^{i(j_1-j_2)x} ,
\tag{2.40}
$$

and

$$
|\partial_\xi^\beta (a_2)^{\sigma_1\sigma_2}(j_1, j_2, \xi)| \lesssim \max\{\langle j_1 \rangle, \langle j_2 \rangle\}^\mu \langle \xi \rangle^{m-\beta}
\tag{2.41}
$$

for some $\mu \geq 0$.

(*iii*) **(Non-homogeneous symbols).** Let $r \in (0, 1)$. We denote by $\mathbf{SNH}_q^m$ the class of maps

$$
(\varphi, V) \in B_r(H^\mu(\mathbb{T}))^2 \to a_{\geq q}(\varphi, V; \cdot) \in \mathcal{N}_s^m \qquad \text{for some } \mu > 1/2,
$$

such that, for any $s \geq s_0 + \mu$, if $(\varphi, V) \in (B_r(H^\mu(\mathbb{T})) \cap H^s(\mathbb{T}))^2$ one has

$$
|a_{\geq q}(\varphi, V; \cdot)|_{\mathcal{N}_s^m} \lesssim_s \|\varphi\|_{H^{s+\mu}}^q + \|V\|_{H^{s+\mu}} ,
$$
$$
|d_\varphi a_{\geq q}(\varphi, V, \cdot)[h]|_{\mathcal{N}_s^m} \lesssim_s (\|\varphi\|_{H^{s+\mu}}^{q-1} + \|V\|_{H^{s+\mu}}) \|h\|_{H^{s+\mu}} , \qquad \forall h \in H^{s+\mu}(\mathbb{T}) , \quad (2.42)
$$
$$
|d_V a_{\geq q}(\varphi, V, \cdot)[h]|_{\mathcal{N}_s^m} \lesssim_s \|h\|_{H^{s+\mu}} .
$$

(*iv*) **(Symbols).** We denote by $\mathbf{S\Sigma}_1^m[r, 3]$ the space of symbols $a(\varphi, V; x, \xi)$ of the form

$$
a(\varphi, V; x, \xi) = a_1(\varphi; x, \xi) + a_2(\varphi; x, \xi) + a_{\geq 3}(\varphi, V; x, \xi),
$$

where $a_j \in \mathbf{SM}_j^m$, $j = 1, 2$, and $a_{\geq 3} \in \mathbf{SNH}_3^m$.

Lemma 2.6 *Let* $a_j = a_j(\varphi; \cdot) \in \mathbf{SM}_j^m$, $j = 1, 2$. *Then, there exist* $\mu \geq 0$, $s_0 = s_0(\mu) > 1/2$ *such that, for* $s \geq s_0$, *the maps* $\varphi \to a_j(\varphi; \cdot)$ *extend as continuous maps* $H^{s+\mu}(\mathbb{T}) \to \mathcal{N}_s^m$. *Moreover*

$$
|a_1(\varphi)|_{\mathcal{N}_s^m} \lesssim_s \|\varphi\|_{H^{s+\mu}}, \qquad |a_2(\varphi)|_{\mathcal{N}_s^m} \quad \lesssim_s \|\varphi\|_{H^{s+\mu}}^2.
$$

Proof The proof is straightforward using (2.16) and the bounds on the coefficients of a_j in Def. 2.5. ∎

Lemma 2.7 *Let* $m \in \mathbb{R}$, $s_0 > 1$. *The following holds.*
(*i*) *Let* $a = a(\varphi, \cdot) \in \mathbf{SM}_j^m$, $j = 1, 2$. *Then* $Op^{\mathrm{BW}}(a) \in \mathbf{M}_j^m$ *and there is* $\mu \geq 0$ *such that*

$$
\|Op^{\mathrm{BW}}(a)h\|_{H^{s-m}} \lesssim_s \|\varphi\|_{H^{s_0+\mu}}^j \|h\|_{H^s}.
$$

(*ii*) *Let* $a = a(\varphi, \cdot) \in \mathbf{SNH}_3^m$. *Then, for all* $h \in H^s(\mathbb{T})$,

$$
\|Op^{\mathrm{BW}}(a)[h]\|_{H^{s-m}} \lesssim_s (\|\varphi\|_{H^{s_0+\mu}}^3 + \|V\|_{H^{s_0+\mu}}) \|h\|_{H^s} .
\tag{2.43}
$$

(*iii*) *Let* $a = a(\varphi, \cdot) \in \mathbf{S\Sigma}_1^m$. *Then* $Op^{\mathrm{BW}}(a) \in \Sigma_1^m[r, 3]$.

Proof (*i*) It follows by Lemmata 2.2 and 2.6. (*ii*) The bound (2.43) is consequence of Lemma 2.2 and bounds (2.42). (*iii*) If the symbol $a(\varphi, \cdot)$ is multilinear (see Def. 2.5), then the (2.34) follows by (2.23), (2.20), (2.38), (2.39), (2.40) and (2.41) with $\beta = 0$ and by explicit computations. Assume that the symbol $a(\varphi, \cdot)$ is non-homogeneous. We first notice that

$$
d_{\varphi, V} Op^{\mathrm{W}}(a) = Op^{\mathrm{W}}(d_{\varphi, V} a).
$$

Then (2.42) and Lemma 2.2 imply the (2.35). This concludes the proof. ∎

Lemma 2.8 *Let m, ρ_1, $\rho_2 \in \mathbb{R}$. We have the following:*
(i) $\mathbf{SM}_j^m \subset \mathbf{SNH}_j^m$ *for $j = 1, 2$;*
(ii) $\mathbf{M}_j^m \subset \mathbf{NH}_j^m$ *for $j = 1, 2$;*
(iii) If $R_i \in \mathbf{M}_i^{-\rho_1}$, $Q_j \in \mathbf{M}_j^{-\rho_2}$ for some $i, j \in \mathbb{N}$, then $R_i \circ Q_j \in \mathbf{M}_{i+j}^{-\min\{\rho_1,\rho_2\}}$. If $R \in \Sigma_1^{-\rho_1}[r, 3]$, $Q \in \Sigma_1^{-\rho_2}[r, 3]$ then $R \circ Q \in \Sigma_1^{-\min\{\rho_1,\rho_2\}}[r, 3]$.

Proof *(i)* Using formulæ (2.38), (2.40) one can write explicitly the differential of the symbol in the variable φ. Then the bounds (2.42) follows by using (2.39), (2.41) and Lemma 2.6. *(ii)* One can reason as done for item *(i)* using (2.30)–(2.32), (2.34) and reasoning as in the proof of Lemma 2.4. *(iii)* Since Definitions 2.3, 2.5 are equivalent to the classes of symbols and operators introduced in [9, 14] we refer to these papers for a detailed proof of the third bullet.

Lemma 2.9 *Let $m \in \mathbb{R}$, $i, k \in \mathbb{N}$, $a \in \mathbf{SM}_i^m$, $B = B(\varphi) \in \mathbf{M}_k^n$ and $c(\varphi; x, \xi) := a(B(\varphi)\varphi; x, \xi)$. Then the following holds.*
(i) If $i = k = 1$ then $c \in \mathbf{SM}_2^m$.
(ii) If i or k are strictly greater than 1 then $c \in \mathbf{SNH}_{i+k}^m$.

Proof *(i)* By (2.38), (2.30), (2.31) we have

$$c(\varphi; x, \xi) = \sum_{j \in \mathbb{Z}} a^\sigma(j, \xi)(B(\varphi)\varphi)_j^\sigma \, e^{i\sigma j x} = \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z}} a^\sigma(j, \xi) \left(\sum_{j_2 \in \mathbb{Z}} (B(\varphi))_{\sigma j}^{\sigma_2 j_2} \varphi_{j_2}^{\sigma_2} \right) e^{i\sigma j x}$$

$$= \frac{1}{2\pi} \sum_{\substack{j, j_1, j_2 \in \mathbb{Z}, \\ \sigma_1 j_1 + \sigma_2 j_2 = \sigma j}} a^\sigma(j, \xi) b^{\sigma_1}(j, j_1, j_2) \varphi_{j_1}^{\sigma_1} \varphi_{j_2}^{\sigma_2} \, e^{i(\sigma_1 j_1 + \sigma_2 j_2)x} .$$

Then setting, for $\sigma j = \sigma_1 j_1 + \sigma_2 j_2$,

$$c^{\sigma_1, \sigma_2}(j_1, j_2, \xi) := \frac{1}{2\pi} a^\sigma(j, \xi) b^{\sigma_1}(j, j_1, j_2),$$

and using the fact that $\varphi \in C^\infty(\mathbb{T}; \mathbb{C})$ we have (recall bounds (2.39), (2.34))

$$|\partial_\xi^\beta c^{\sigma_1, \sigma_2}(j_1, j_2, \xi)| \lesssim |\partial_\xi^\beta a^\sigma(j, \xi)| \max_2 \{\langle j_1 \rangle, \langle j_2 \rangle\}^{\mu + |n|} \max_1 \{\langle j_1 \rangle, \langle j_2 \rangle\}^n$$

$$\lesssim \max\{\langle j_1 \rangle, \langle j_2 \rangle\}^{\tilde{\mu}} \langle \xi \rangle^{m - \beta} ,$$

for some $\tilde{\mu} \geq 0$. This concludes the proof of *(i)*. To prove *(ii)* we first notice that, reasoning as before, one can prove that c is homogenous of degree $i + k \geq 3$. More precisely

$$c(\varphi; x, \xi) = \sum_{\substack{j_p \in \mathbb{Z}, p = 1, \ldots, i+k, \\ \sigma_1 j_1 + \cdots + \sigma_{i+k} j_{i+k} = \sigma j}} a^\sigma(j, \xi) b^{\sigma_1, \ldots, \sigma_{i+k}}(j, j_1, \ldots, j_{i+k}) \varphi_{j_1}^{\sigma_1} \cdots \varphi_{j_{i+k}}^{\sigma_{i+k}} \, e^{i\sigma j x}.$$

Without loss of generality we can set $\langle j_1 \rangle = \max_{p=1,\ldots,i+k}\{\langle j_p \rangle\}$. Therefore, for $s \geq \beta \geq 0$ we have

$$\|\partial_\xi^\beta c(\varphi; x, \xi)\|_{H^s}^2 \lesssim \sum_{j \in \mathbb{Z}} |\partial_\xi^\beta a^\sigma(j, \xi)|^2 \left(\sum_{\sigma_1 j_1 + \cdots + \sigma_{i+k} j_{i+k} = \sigma j} |\varphi_{j_1}^{\sigma_1} \cdots \varphi_{j_{i+k}}^{\sigma_{i+k}}| \langle j_1 \rangle^{s + |n| + \mu} \right)^2$$

$$\lesssim \sum_{j \in \mathbb{Z}} |(\partial_\xi^\beta a^\sigma(j, \xi)) \, (\langle j_1 \rangle^{s + |n| + \mu} |\varphi_{j_1}| * \cdots * |\varphi_{j_{i+k}}|)_j|^2$$

$$\lesssim |a|_{\mathcal{N}_s^m} \|\varphi\|_{H^{s+\tilde{\mu}}}^{2(i+k)} ,$$

for some $\tilde{\mu} \geq 0$. Since c is a multilinear symbol its differential can be computed explicitly and we obtain the bounds (2.42).

Lemma 2.10 *Let* $a \in \mathbf{SM}_i^m$, $Q \in \mathbf{M}_j^{-\rho}$ *and* $A := Op^{\mathrm{BW}}(a) \circ Q$, $Q \circ Op^{\mathrm{BW}}(a)$. *Then the following holds.*
(i) If $i = j = 1$ *then* $A \in \mathbf{M}_2^{m-\rho}$.
(ii) If i *or* j *are greater than 1 then* $A \in \mathbf{NHM}_{i+j}^{m-\rho}$.

Proof Let us study the case $A = Op^{\mathrm{BW}}(a) \circ Q$. The other is similar.
(i) By using (2.20), (2.22)–(2.23), (2.38) and (2.30) (with $p = 1$), (2.31) one deduces that the operator A can be expanded as in (2.30) with $p = 2$ with coefficients as in (2.32) where

$$r_2^{\sigma_1,\sigma_2}(j, j_1, j_2, j_3) := a^{\sigma_1}(\sigma_1 j_1, \frac{j - \sigma_1 j_1}{2}) \chi\left(\frac{2|j_1|}{\langle j - \sigma_1 j_1 \rangle}\right) Q^{\sigma_2}(j - \sigma_1 j_1, j_3, j_2)$$

$$\sigma_1 j_1 + \sigma_2 j_2 + j_3 = j,$$

where $\alpha^{\sigma_1}(p, j)$, $Q^{\sigma_2}(j, p_1, p_2)$, $j, p, p_1, p_2 \in \mathbb{Z}$ are respectively the coefficients of the symbol $a(x, \xi)$ and of the operator Q, and satisfy the bounds (2.39) and (2.34). Recalling (2.21), one can note that

$$\chi\left(\frac{2|j_1|}{\langle j - \sigma_1 j_1 \rangle}\right) \neq 0 \quad \Rightarrow \quad |j_1| \ll |j| \sim |j - \sigma_1 j_1| \leq |j_2| + |j_3| \lesssim \max_2\{\langle j_1 \rangle, \langle j_2 \rangle, \langle j_3 \rangle\}.$$

Therefore, by (2.39) and (2.34), we get

$$|r_2^{\sigma_1,\sigma_2}(j, j_1, j_2, j_3)| \lesssim |j_1|^{\mu} \langle j - \sigma_1 j_1 \rangle^m \frac{\max_2\{\langle j_2 \rangle, \langle j_3 \rangle\}^{\mu+\rho}}{\max_1\{\langle j_2 \rangle, \langle j_3 \rangle\}^{\rho}}$$

$$\lesssim \frac{\max_2\{\langle j_1 \rangle, \langle j_2 \rangle, \langle j_3 \rangle\}^{\tilde{\mu}+\rho}}{\max_1\{\langle j_1 \rangle, \langle j_2 \rangle, \langle j_3 \rangle\}^{\rho-m}},$$

for some $\tilde{\mu} \geq \mu$. This implies $A \in \mathbf{M}_2^{m-\rho}$.
(ii) It follows by Lemmata 2.2, 2.6, 2.7 and 2.4, 2.8.

Remark 2.11 Given $a \in \mathbf{S\Sigma}_1^{m_1}[r, 3]$ and $b \in \mathbf{S\Sigma}_1^{m_2}[r, 3]$ and reasoning as in [9, 14], one can note that the remainder $R(a, b) \in \Sigma_1^{m_1+m_2-3}[r, 3]$.

2.2.3 Matrices of Operators

We define some special classes of linear operators on spaces of functions.

Definition 2.12 Let $\rho \in \mathbb{R}$, $p = 1, 2$, $q \geq 3$. We denote by $\mathbf{M}_p^{-\rho} \otimes \mathcal{M}_2(\mathbb{C})$ the space of 2×2 matrices whose entries are operators in $\mathbf{M}_p^{-\rho}$. We denote by $\mathbf{SM}_p^m \otimes \mathcal{M}_2(\mathbb{C})$ the space of 2×2 matrices whose entries are symbols in $\mathbf{SM}_p^m$. Given a matrix of operators $Q \in \mathbf{M}_p^{-\rho} \otimes \mathcal{M}_2(\mathbb{C})$ (resp. $\mathbf{NH}_q^{-\rho}[r] \otimes \mathcal{M}_2(\mathbb{C})$, $\Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$) we shall write

$$Q := (Q_\sigma^{\sigma'})_{\sigma,\sigma' \in \{\pm\}} := \begin{pmatrix} Q_+^+ & Q_+^- \\ Q_-^+ & Q_-^- \end{pmatrix} \tag{2.44}$$

where $Q_\sigma^{\sigma'} \in \mathbf{M}_p^{-\rho}$ denote the entries of the matrix Q.
Similar definitions and notations are used when we consider $\mathbf{NH}_q^{-\rho}[r]$, $\Sigma_1^{-\rho}[r, 3]$ instead of $\mathbf{M}_p^{-\rho}$.

Along the paper we shall consider special subspaces of matrix valued operators introduced above.

Definition 2.13 Let $\rho \in \mathbb{R}$, $p \in \mathbb{N} \cup \{0\}$, $0 < r \le 1$ and let $Q \in \mathbf{M}_p^{-\rho}$.
(i) We define the operator $\overline{Q}$ as

$$\overline{Q}[h] := \overline{Q[\overline{h}]}, \qquad h \in H^s(\mathbb{T}). \tag{2.45}$$

(ii) We say that a matrix of linear operators $\mathcal{Q} \in \mathbf{M}_p^{-\rho} \otimes \mathcal{M}_2(\mathbb{C})$ is *real-to-real* (or reality preserving) if it has the form (see (2.44))

$$\mathcal{Q} := \begin{pmatrix} Q_+^+ & Q_+^- \\ Q_+^- & Q_+^+ \end{pmatrix}, \quad \text{i.e.} \quad Q_{-\sigma}^{-\sigma'} = \overline{Q_\sigma^{\sigma'}}, \quad \forall \sigma, \sigma' \in \{\pm\}, \tag{2.46}$$

for some $Q_+^+, Q_+^- \in \mathbf{M}_p^{-\rho}$.
(iii) We say that a matrix of linear operators $\mathcal{Q} \in \mathbf{M}_p^{-\rho} \otimes \mathcal{M}_2(\mathbb{C})$ is *real* if it has the form

$$\mathcal{Q} := \begin{pmatrix} Q^+ & Q^- \\ Q^- & Q^+ \end{pmatrix}, \quad \text{where} \quad Q^+ := Q_+^+, \quad Q^- := Q_+^-, \quad Q^\sigma = \overline{Q^\sigma} \tag{2.47}$$

for some $Q_+^+, Q_+^- \in \mathbf{M}_p^{-\rho}$.
Similar definitions and notations are used when we consider $\mathbf{NH}_q^{-\rho}[r]$, $\Sigma_1^{-\rho}[r, 3]$ instead of $\mathbf{M}_p^{-\rho}$.

Recall the spaces $H^s(\mathbb{T})$ defined in (1.5). One can easily check that a *real-to-real* linear operator $\mathcal{Q}$ preserves the spaces $H^s(\mathbb{T})$. We also observe that a real matrix is real-to-real. On the space $H^0(\mathbb{T})$ we define the scalar product

$$(U, V)_{L^2} := \int_{\mathbb{T}} U \cdot \overline{V} dx, \qquad U = \begin{bmatrix} u \\ \overline{u} \end{bmatrix}, \quad V = \begin{bmatrix} v \\ \overline{v} \end{bmatrix}. \tag{2.48}$$

Given an operator $\mathcal{Q}$ of the form (2.47) we denote by $\mathcal{Q}^*$ its adjoint with respect to the scalar product (2.48), i.e.

$$(\mathcal{Q}U, V)_{L^2} = (U, \mathcal{Q}^* V)_{L^2}, \quad \forall\, U, V \in H^0(\mathbb{T}). \tag{2.49}$$

One can check that

$$\mathcal{Q}^* := \begin{pmatrix} (Q^+)^* & (\overline{Q^-})^* \\ (Q^-)^* & (\overline{Q^+})^* \end{pmatrix} \stackrel{(2.47)}{=} \begin{pmatrix} (Q^+)^* & (Q^-)^* \\ (Q^-)^* & (Q^+)^* \end{pmatrix}, \tag{2.50}$$

where $(Q^+)^*$ and $(Q^-)^*$ are respectively the adjoints of the operators Q^+ and Q^- with respect to the standard complex scalar product on $L^2(\mathbb{T}; \mathbb{C})$.

Definition 2.14 Let $\mathcal{Q} \in \mathbf{M}_p^{-\rho} \otimes \mathcal{M}_2(\mathbb{C})$ be a *real* linear operator of the form (2.47). We say that $\mathcal{Q}$ is:

- *self-adjoint* if

$$(Q^+)^* = Q^+, \qquad Q^- = (Q^-)^*. \tag{2.51}$$

- *Hamiltonian* if

$$\mathcal{Q} = iE\mathcal{A}, \qquad E = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{2.52}$$

where $\mathcal{A} \in \mathbf{M}_p^{-\rho} \otimes \mathcal{M}_2(\mathbb{C})$ is a self-adjoint operator matrix.

Similar definitions are used when we consider operators in $\mathbf{NH}_q^{-\rho}[r] \otimes \mathcal{M}_2(\mathbb{C})$, $\Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$.

Consider now a symbol $a(x, \xi) \in \mathbf{S}\Sigma_1^m[r, 3]$ and set $A := Op^W(a(x, \xi))$. Using (2.20) one can check that (recall (2.45))

$$\overline{A} = Op^W(\overline{a(x, -\xi)}), \qquad A^* = Op^W(\overline{a(x, \xi)}). \qquad (2.53)$$

By (2.53) we deduce that the operator A is self-adjoint with respect to the standard scalar product on $L^2(\mathbb{T}; \mathbb{C})$ if and only if the symbol $a(x, \xi)$ is real valued.

We need the following definition. Consider two symbols $a, b \in \mathcal{N}_s^m$ and the matrix

$$A := A(x, \xi) := \begin{pmatrix} a(x, \xi) & b(x, \xi) \\ \overline{b(x, -\xi)} & \overline{a(x, -\xi)} \end{pmatrix}.$$

Define the operator

$$M := Op^W(A(x, \xi)) := \begin{pmatrix} Op^W(a(x, \xi)) & Op^W(b(x, \xi)) \\ Op^W(\overline{b(x, -\xi)}) & Op^W(\overline{a(x, -\xi)}) \end{pmatrix}. \qquad (2.54)$$

For the matrix of pseudo-differential operators defined above the following facts hold:

- *Real-to-real*: by (2.53) we have that the operator M in (2.54) has the form (2.46), hence it is real-to-real;
- *Self-adjointeness*: using (2.53) the operator M in (2.54) is self-adjoint with respect to the scalar product (2.48) if and only if (recall (2.51))

$$a(x, \xi) = \overline{a(x, \xi)}, \qquad b(x, -\xi) = b(x, \xi). \qquad (2.55)$$

- *Reality*: if both the symbols $a(x, \xi)$, $b(x, \xi)$ are real valued we have that the operator M in (2.54) has the form (2.47), hence it is *real* and *self-adjoint*.

Definition 2.15 (*Symplectic map*) Let $\mathcal{Q} \in \mathbf{M}_p^{-\rho} \otimes \mathcal{M}_2(\mathbb{C})$ (resp. $\mathbf{NH}_p^{-\rho}[r] \otimes \mathcal{M}_2(\mathbb{C})$ or $\Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$) be a real-to-real (or *real*) linear operator of the form (2.46) (or (2.47)). We say that $\mathcal{Q}$ is *symplectic* if

$$\mathcal{Q}^*(iE)\mathcal{Q} = iE. \qquad (2.56)$$

3 Birkhoff Normal Form

In this section we construct a suitable normal form for the Klein–Gordon Hamiltonian H in (2.12). This is the content of Proposition 3.9. This normal form procedure presents *small divisors* problems. Estimates on such small divisors are provided by Lemma 3.8. In Sect. 3.1 we recall some properties of homogeneous Hamiltonian functions. In Sect. 3.3 we construct the oscillating function which approximate the dynamics of the Klein–Gordon.

3.1 Homogenous Hamiltonians

Let F be a homogenous Hamiltonian of the form

$$F(Z) = \sum_{\pi(\vec{\sigma}, \vec{j})=0} F_{\vec{\sigma}, \vec{j}} z_{j_1}^{\sigma_n} \ldots z_{j_n}^{\sigma_n}, \qquad Z = \begin{bmatrix} z \\ \overline{z} \end{bmatrix}, \qquad (3.1)$$

where $\vec{\sigma} := (\sigma_1, \ldots, \sigma_n)$ is a n-dimensional vector of signs $\{\pm\}$, $\vec{j} := (j_1, \ldots, j_n) \in \mathbb{Z}^n$, $z_j^+ := z_j$, $z_j^- := \overline{z_j}$, $F_{\vec{\sigma}, \vec{j}} \in \mathbb{C}$ and

$$\pi(\vec{\sigma}, \vec{j}) := \sum_{i=1}^{n} \sigma_i j_i . \tag{3.2}$$

The equality $\pi(\vec{\sigma}, \vec{j}) = 0$ encodes the fact that F commutes with the momentum Hamiltonian

$$M := -\frac{i}{2}(Z_x, \overline{Z})_{H^0(\mathbb{T})} = \sum_{j \in \mathbb{Z}} j \, |z_j|^2.$$

If we want to highlight the degree of homogeneity of the Hamiltonian F we shall write $F = F^{(n)}$.

Fixed a subset S of $\mathbb{Z}$, we define for $0 \leq k \leq n$

$$\mathcal{A}_{n,k} := \left\{ (\vec{\sigma}, \vec{j}) \in \{\pm\}^n \times \mathbb{Z}^n : \text{there are exactly indexes } j_i \text{ that belong to S, } \pi(\vec{\sigma}, \vec{j}) = 0 \right\}.$$

We define also

$$\mathcal{A}_{n, \leq k} := \bigcup_{i \leq k} \mathcal{A}_{n, i}, \qquad \mathcal{A}_{n, \geq k} := \bigcup_{i \geq k} \mathcal{A}_{n, i},$$

and

$$F^{(n,k)} = \sum_{\mathcal{A}_{n,k}} F_{\vec{\sigma}, \vec{j}} \, z_{j_1}^{\sigma_n} \cdots z_{j_n}^{\sigma_n}.$$

Analogously one can define $F^{(n, \leq k)}$ and $F^{(n, \geq k)}$ replacing $\mathcal{A}_{n,k}$ with $\mathcal{A}_{n, \leq k}$ and $\mathcal{A}_{n, \geq k}$ respectively. We define the projections

$$\Pi^{(n,k)} F = F^{(n,k)}, \qquad \Pi^{(n, \leq k)} F = F^{(n, \leq k)}, \qquad \Pi^{(n, \geq k)} F = F^{(n, \geq k)}. \tag{3.3}$$

Remark 3.1 By the conservation of momentum, if S is a finite set then $\mathcal{A}_{n, \leq 1}$ is finite, for all $n \in \mathbb{N}$.

Definition 3.2 *(Resonances)* Let $n \geq 0$. We say that $(\vec{\sigma}, \vec{j}) \in \{\pm\}^n \times \mathbb{Z}^n$ is a resonance, or a n-resonance, if

$$\Omega(m; \vec{\sigma}, \vec{j}) := \Omega(\vec{\sigma}, \vec{j}) := \sum_{i=1}^{n} \sigma_i \Lambda(j_i) = 0, \qquad \pi(\vec{\sigma}, \vec{j}) = 0. \tag{3.4}$$

We say that a n-resonance $(\vec{\sigma}, \vec{j})$ is *trivial* if if n is even and, up to permutations, one has

$$\vec{j} = (j, j, k, k, \ldots), \qquad \vec{\sigma} = (+, -, +, -, \ldots).$$

Definition 3.3 *(Resonant monomials)*
• We say that $z_{j_1}^{\sigma_1} \ldots z_{j_n}^{\sigma_n}$ is a resonant monomial if $(\vec{\sigma}, \vec{j})$ is a resonance.
• A resonant monomial $z_{j_1}^{\sigma_1} \ldots z_{j_n}^{\sigma_n}$ is said *action-preserving* if it depends only on the square modulus (or actions) $|z_{j_i}|^2$.
• Given a homogenous Hamiltonian F of degree n as in (3.1) we denote by F_{res} the projection of F on the space of resonant monomials.

Remark 3.4 Using the conservation of momentum, one can prove that actually all 4-resonances are trivial. This means that the only resonant monomials of degree 4 are action-preserving. We prove this in Lemma 3.8-(iv).

Definition 3.5 We denote by $\mathrm{ad}_{H^{(2)}}(\cdot) := \{\cdot, H^{(2)}\}$ the adjoint action of the Hamiltonian $H^{(2)}$. We denote by Π_{Ker} and Π_{Rg} the projections on the kernel and the range of the adjoint action respectively. We observe that Π_{Ker} is the projection on the space of resonant monomials.

Definition 3.6 Let E be a subset of $\mathbb{Z}$. Let us define the subspace

$$\mathbf{V}_E := \left\{ Z = \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z}} Z_j \, e^{ijx} : Z_j = 0 \ \ j \notin E \right\}. \tag{3.5}$$

We denote with $\Pi_{\mathbf{V}_E}$ the projection on the subspace $\mathbf{V}_E$.

If E is a finite subset and a vector field X has image contained in $\mathbf{V}_E$ then we say that X has finite rank or it is a finite rank vector field.

In the following lemma we provide estimates for the vector fields associated to the generators of the Birkhoff maps. With an abuse of notation we denote $L^\infty(\mathbb{T}) \times L^\infty(\mathbb{T})$ by $L^\infty(\mathbb{T})$.

Lemma 3.7 *Let $n \geq 3$ and $S \subset \mathbb{Z}$ be a finite set. Let F be a homogenous Hamiltonian of degree n as in* (3.1) *such that*

$$F = F^{(n, \leq 1)} \quad \text{and} \quad \llbracket F \rrbracket := \sup_{(\vec{\sigma}, \vec{j})} |F_{\vec{\sigma}, \vec{j}}| < \infty. \tag{3.6}$$

Then the following holds:

(i) We have the estimates

$$\|X_F(Z)\|_{H^s} \lesssim_s \llbracket F \rrbracket \|Z\|_{L^\infty}^{n-2} \|Z\|_{H^0}, \quad \|X_F(Z)\|_{L^\infty} \lesssim \llbracket F \rrbracket \|Z\|_{L^\infty}^{n-1},$$
$$\forall Z \in H^0(\mathbb{T}) \cap L^\infty(\mathbb{T}),$$
$$\|(dX_F)(Z)[\widetilde{Z}]\|_{H^s} \lesssim_s \llbracket F \rrbracket \|Z\|_{L^\infty}^{n-2} \|\widetilde{Z}\|_{H^0}, \quad \forall \widetilde{Z} \in H^0(\mathbb{T}),$$
$$\|(dX_F)(Z)[\widetilde{Z}]\|_{L^\infty} \lesssim \llbracket F \rrbracket \|Z\|_{L^\infty}^{n-2} \|\widetilde{Z}\|_{L^\infty}, \quad \forall \widetilde{Z} \in L^\infty(\mathbb{T}).$$

(ii) If G is another homogenous Hamiltonian of degree m then $\{F, G\}$ is a homogenous Hamiltonian of degree $n + m - 2$ such that its vector field has finite rank and the following estimate holds

$$\|X_{\{F,G\}}\|_{H^s} \lesssim_s \llbracket F \rrbracket \llbracket G \rrbracket \|Z\|_{L^\infty}^{n+m-4} \|Z\|_{H^0} \quad \forall Z \in H^0(\mathbb{T}) \cap L^\infty(\mathbb{T}).$$

Proof By Remark 3.1 the assumption (3.6) guarantees that the vector field X_F has finite rank. In particular its image is contained in $\mathbf{V}_{E_n}$ (recall (3.5)) where $E_n := (n-1) S$. Then $\|X_F(Z)\|_{H^s} \leq C \|X_F(Z)\|_{H^0}$ where $C > 0$ is a constant depending on the index s and on the set S. For some constant $C_* > 0$ we have

$$\|X_F(Z)\|_{H^0} \leq C_* \llbracket F \rrbracket \|\widehat{Z} * \cdots * \widehat{Z}\|_{\ell^2} = C_* \llbracket F \rrbracket \|Z^{n-1}\|_{H^0} \leq C_* \llbracket F \rrbracket \|Z\|_{L^\infty}^{n-2} \|Z\|_{H^0},$$

where we denoted by $\widehat{Z}$ the sequence of the Fourier coefficients of Z. This concludes the proof of the first bound in item (i). The third follows similarly using the fact that X_F is a multilinear operator (see (3.1)).

Regarding the second bound, we have that $\|Z\|_{L^\infty} \leq \|Z\|_{\ell^1}$, where

$$\|Z\|_{\ell^1} := \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z}} |Z_j|.$$

Therefore by Young inequality

$$\|X_F(Z)\|_{L^\infty} \leq \|X_F(Z)\|_{\ell^1} \lesssim \frac{[\![F]\!]}{\sqrt{2\pi}} \sum_{j \in E} | \sum_{\sum_{i=1}^{n-1} \sigma_i j_i = j} z_{j_1}^{\sigma_1} \dots z_{j_{n-1}}^{\sigma_{n-1}}| \lesssim [\![F]\!] \Big(\frac{1}{\sqrt{2\pi}} \sum_{j \in E_n} |Z_j|\Big)^{n-1},$$

for some $E \subseteq E_n$. By noticing that $|Z_j| \leq \sqrt{2\pi}\|Z\|_{L^\infty}$ we get the second bound in (i). The fourth bound is obtained in the same way using that dX_F is a homogenous polynomial of degree $n-1$.

Item (ii) follows using formula (2.15) and reasoning as in the proof of item (i). We observe that the support of the Hamiltonian $\{F, G\}$ is contained in the support of F, which is finite, hence $X_{\{F,G\}}$ has finite rank. $\qquad \square$

The following lemma provides lower bounds on the function $\Omega(\vec{\sigma}, \vec{j})$ in (3.4).

Lemma 3.8 (Small divisors) *Let S be a finite subset of $\mathbb{Z}$. There exists a full measure set $\mathcal{M} \subset [1, 2]$ such that for all $m \in \mathcal{M}$ the following holds:*

(i) if $n = 3, 5$ then $\Omega(\vec{\sigma}, \vec{j}) \neq 0$ for all $(\vec{\sigma}, \vec{j}) \in \mathcal{A}_{n, \leq 1}$;
(ii) if $n = 4$ and $(\vec{\sigma}, \vec{j}) \in \mathcal{A}_{n, \leq 1}$ is such that $\Omega(\vec{\sigma}, \vec{j}) = 0$ then $(\vec{\sigma}, \vec{j}) \in \mathcal{A}_{n,0}$. Moreover the 4-resonances are trivial, namely they have the following form (up to permutations)

$$\vec{j} = (j, j, k, k), \quad \vec{\sigma} = (+, -, +, -). \tag{3.7}$$

(iii) If $n = 3$ and $\pi(\vec{\sigma}, \vec{j}) = 0$ then

$$|\Omega(\vec{\sigma}, \vec{j})| \geq \frac{m}{2\sqrt{j^2 + m}}, \tag{3.8}$$

where $|j| = \min_{i=1,2,3} |j_i|$.
(iv) If $n = 4$, $\pi(\vec{\sigma}, \vec{j}) = 0$ and $(\vec{\sigma}, \vec{j})$ is not a trivial resonance (recall Def. 3.2) then

$$|\Omega(\vec{\sigma}, \vec{j})| \geq c_* \frac{m}{(\sqrt{j^2 + m})^3}, \tag{3.9}$$

where $|j| = \min_{i=1,2,3,4} |j_i|$ and for some absolute constant $c_* > 0$.
(v) For any $m \in \mathcal{M}$ and for $n = 5$ one has

$$|\Omega(m; \vec{\sigma}, \vec{j})| > \gamma, \tag{3.10}$$

for some $\gamma > 0$ depending only on S, for any $(\vec{\sigma}, \vec{j}) \in \mathcal{A}_{5, \leq 1}$.

Proof First of all, given $(\vec{\sigma}, \vec{j}) \in \{\pm\}^n \times \mathbb{Z}^n$, $n \geq 3$, without loss of generality we can consider $|j_1| \geq \cdots \geq |j_n| \geq 0$. We have to give lower bounds for the resonant combinations of the j_i's with any possible choice of the signs σ_i. To simplify the notation we shall write (recall (3.2), (3.4))

$$\Omega_{\sigma_1,\dots,\sigma_n}(m) := \Omega(m; \vec{\sigma}, \vec{j}), \qquad \pi_{\sigma_1,\dots,\sigma_n} := \pi(\vec{\sigma}, \vec{j}).$$

Proof of item (iii). First we observe that $\Omega_{++-}(m), \Omega_{+-+}(m) \geq \Omega_{-++}(m)$. Hence we can find a bound from below just for $\Omega_{-++}(m)$. We have First we note that

$$\Omega_{-++}(0) = -|j_1| + |j_2| + |j_3| \geq 0,$$

since $\pi_{-++} = -j_1 + j_2 + j_3 = 0$. Secondly we also note that

$$\Omega'_{-++}(m) = \frac{1}{2}\left(-\frac{1}{\Lambda(j_1)} + \frac{1}{\Lambda(j_2)} + \frac{1}{\Lambda(j_3)}\right) \geq \frac{1}{2\Lambda(j_3)},$$

where $\Lambda(j)$ is in (2.5). Then we have

$$\Omega_{-++}(m) = \Omega_{-++}(0) + \int_0^m \Omega'_{-++}(\tilde{m})\,d\tilde{m} \geq \frac{m}{2\sqrt{j_3^2 + m}},$$

which is the bound (3.8).

Proof of item (iv). The interesting cases are when there are 3 and 2 plus signs. For the first case we reason as before. We have $\Omega_{+-++}(m), \Omega_{++-+}(m), \Omega_{+++-}(m) \geq \Omega_{-+++}(m)$. Hence we study the latter case. Recalling again that

$$\pi_{-+++} = -j_1 + j_2 + j_3 + j_4 = 0,$$

we observe that

$$\Omega_{-+++}(0) = -|j_1| + |j_2| + |j_3| + |j_4| \geq 0, \quad \Omega'_{-+++}(m) \geq \frac{1}{2\Lambda(j_4)}.$$

Then we obtain

$$\Omega_{-+++}(m) \geq \frac{m}{2\sqrt{j_4^2 + m}}.$$

For the case of 2 plus signs we have $\Omega_{+-+-}(m), \Omega_{++--}(m) \geq \Omega_{+--+}(m)$. All the other cases are obtained by a global change of sign. Let us now write

$$\Omega_{+--+}(m) := \Omega(m) = \sqrt{j_1^2 + m} - \sqrt{j_2^2 + m} - \sqrt{j_3^2 + m} + \sqrt{j_4^2 + m}.$$

We notice that $\Omega_{+--+}(m)$ vanishes identically on 4-resonances $(\vec{\sigma}, \vec{j})$ of the form

$$\vec{j} = (j, \pm j, k, \pm k), \qquad \vec{\sigma} = (+, -, -, +),$$

up to permutations. Actually we claim that the above resonances are trivial. Indeed, up to permutations, we have the following cases.

The first case is when $\vec{j} = (j, j, k, k)$ or $\vec{j} = (j, j, -k, -k)$, $\vec{\sigma} = (+, -, -, +)$ which are obviously trivial. The second case is when $\vec{j} = (j, -j, k, -k)$, $\vec{\sigma} = (+, -, -, +)$. Then (recall (3.2)), by the conservation of momentum, we have $\pi(\vec{\sigma}, \vec{j}) = 2j - 2k = 0$, which implies that $j = k$. Then $(\vec{\sigma}, \vec{j})$ has the form (3.7). The other case is $\vec{j} = (j, -j, k, k)$, $\vec{\sigma} = (+, -, -, +)$. Then $\pi(\vec{\sigma}, \vec{j}) = 2j = 0$. Again $(\vec{\sigma}, \vec{j})$ has the form (3.7). This proves the claim.

Hence, recalling the momentum conservation $j_1 - j_2 - j_3 + j_4 = 0$, and assuming that $(\vec{\sigma}, \vec{j})$ is not a trivial resonance (see Def. 3.2), we can follow Lemma 7.2 in [7] to obtain the lower bound

$$\Omega(m) \geq c_* \frac{m}{(\sqrt{j_4^2 + m})^3}$$

for some absolute constant $c_* > 0$. This implies (3.9) and concludes the proof of item (iv).

Proof of item (ii). By item (iv) proved above we know that $\Omega(\vec{\sigma}, \vec{j}) = 0$ only if $(\vec{\sigma}, \vec{j})$ is a trivial resonance, i.e. if it has the form (3.7). This implies item (ii)

Proof of items (i) and (v). By section 5 in [22] we have that for *almost every mass m*, there are $c > 0$ and $N \in \mathbb{N}$ such that (see (3.4))

$$|\Omega(m; \vec{\sigma}, \vec{j})| \geq c(1 + |j_1|^2 + \cdots + |j_n|^2)^{-N}$$

for any $(\vec{\sigma}, \vec{j}) \in \{\pm\}^n \times \mathbb{Z}^n$ and n odd. We are only interested in the case $n = 3, 5$. Moreover, taking $(\vec{\sigma}, \vec{j}) \in \mathcal{A}_{n, \leq 1}, n = 3, 5$, we have, by Remark 3.1, that only a finite number of choices of $(\vec{\sigma}, \vec{j})$ are possible. Hence, the bound above implies that actually $|\Omega(m; \vec{\sigma}, \vec{j})| \geq \gamma$ for some constant γ depending only on the set of indexes S. Therefore bound (3.10) and item (i) follow. This concludes the proof.

3.2 Partial Birkhoff Normal Form

In this subsection we prove the following result.

Proposition 3.9 *Let S be a finite subset of $\mathbb{Z}$ and $\mathcal{M}$ the set defined in Lemma 3.8. Then for all $m \in \mathcal{M}$ and for all $s \geq 0$ there exists $C(s) > 0$ such that for all $r > 0$ satisfying*

$$C(s)\, r \leq 1 \,, \tag{3.11}$$

the following holds. There exists a finite set $E \subset \mathbb{Z}$ such that $S \subset E$ and an analytic, invertible, symplectic change of variables $\Phi_B : B_1(H^s(\mathbb{T})) \cap B_r(L^\infty(\mathbb{T})) \to B_2(H^s(\mathbb{T})) \cap B_{2r}(L^\infty(\mathbb{T}))$ of the form

$$\Phi_B = \mathrm{I} + \Pi_{\mathbf{V}_E} \Phi \Pi_{\mathbf{V}_E} \,, \qquad \Phi_B(Z) =: W \,, \tag{3.12}$$

with

$$\|\Phi_B(Z) - Z\|_{H^s} \lesssim_s r\|Z\|_{H^0} \,, \tag{3.13}$$

$$\|(d\Phi_B(Z))\widetilde{Z} - \widetilde{Z}\|_{H^s} \lesssim_s r\|\widetilde{Z}\|_{H^0} \,, \qquad \forall\, \widetilde{Z} \in H^0(\mathbb{T}) \,, \tag{3.14}$$

such that the Hamiltonian H in (2.12) is transformed in

$$\mathcal{H} := H \circ \Phi_B^{-1} = H^{(2)} + H^{(3, \geq 2)} + \mathcal{H}_{res}^{(4,0)} + \mathcal{H}^{(4, \geq 2)} + \mathcal{H}^{(5, \geq 2)} + \mathfrak{R}^{(\geq 6)} \,, \tag{3.15}$$

where $\mathcal{H}^{(4, \geq 2)}, \mathcal{H}^{(5, \geq 2)}$ and $\mathfrak{R}^{(\geq 6)}$ generate a finite rank vector field. Moreover

(i) *$\mathcal{H}_{res}^{(4,0)}$ is a action-preserving Hamiltonian of the form*

$$\mathcal{H}_{res}^{(4,0)}(W) = \frac{1}{2} \sum_{j,k \in S} c_{jk}|w_j|^2|w_k|^2 \,, \qquad c_{jk} \in \mathbb{R}, \quad c_{jk} = c_{kj} \,. \tag{3.16}$$

(ii) *The remainder $X_{\mathfrak{R}^{(\geq 6)}}$ is a smooth function with a zero at the origin of order 6 and*

$$\|X_{\mathfrak{R}^{(\geq 6)}}(W)\|_{H^s} \lesssim_s \|W\|_{L^\infty}^4 \|W\|_{H^s} \,, \qquad \forall\, W \in H^s(\mathbb{T}) \,. \tag{3.17}$$

Moreover the vector fields $X_{\mathcal{H}^{(4, \geq 2)}}, X_{\mathcal{H}^{(5, \geq 2)}}$ are homogenous of degree 4 and 5 respectively and satisfy

$$\|X_{\mathcal{H}^{(4, \geq 2)}}(W)\|_{H^s} \lesssim_s \|W\|_{L^\infty}^2 \|W\|_{H^s} \,,$$
$$\|X_{\mathcal{H}^{(5, \geq 2)}}(W)\|_{H^s} \lesssim_s \|W\|_{L^\infty}^3 \|W\|_{H^s} \,. \tag{3.18}$$

(iii) *Let $\varphi \in B_r(H^s(\mathbb{T}))$. Then the linearized operator $(dX_{\mathfrak{R}^{(\geq 6)}})(\varphi)[\cdot]$ belongs to the class $\mathbf{NH}_4^{-\rho}[r]$ (see Def. 2.3) for any $\rho \geq 0$. In particular one has*

$$\|(dX_{\mathfrak{R}^{(\geq 6)}})(\varphi)[W]\|_{H^s} \lesssim_s \|\varphi\|_{H^s}^4 \|W\|_{H^s} \,,$$
$$\|(d^2 X_{\mathfrak{R}^{(\geq 6)}})(\varphi)[W_1, W_2]\|_{H^s} \lesssim_s \|\varphi\|_{H^s}^3 \|W_1\|_{H^s} \|W_2\|_{H^s} \,, \tag{3.19}$$

for any W, W_1, $W_2 \in H^s(\mathbb{T})$.

The inverse map Φ_B^{-1} fulfils bounds like (3.13)–(3.14).

Proof The map Φ_B is constructed in three steps.

First step: Let us consider the homogenous Hamiltonian (recall (3.4))

$$F^{(3)} = \sum_{\mathcal{A}_{3,\leq 1}} F_{\vec{\sigma},\vec{j}}^{(3)} z_{j_1}^{\sigma_1} z_{j_2}^{\sigma_2} z_{j_3}^{\sigma_3}, \qquad F_{\vec{\sigma},\vec{j}}^{(3)} := \begin{cases} \dfrac{H_{\vec{\sigma},\vec{j}}^{(3,\leq 1)}}{i\Omega(\vec{\sigma},\vec{j})} & \text{if } \Omega(\vec{\sigma},\vec{j}) \neq 0, \\ 0 & \text{if } \Omega(\vec{\sigma},\vec{j}) = 0. \end{cases}$$

By definition the Hamiltonian $F^{(3)}$ solves the homological equation

$$\{F^{(3)}, H^{(2)}\} + H^{(3,\leq 1)} = \Pi_{\mathrm{Ker}} H^{(3,\leq 1)}. \tag{3.20}$$

By item (i) of Lemma 3.8 we have that $\Pi_{\mathrm{Ker}} H^{(3,\leq 1)} = 0$. We observe that by Remark 3.1 the equation $\dot{Z} = X_{F^{(3)}}(Z)$ is an analytic finite dimensional ODE. In particular the image of $X_{F^{(3)}}$ is contained in a subspace of the form (3.5). Then it is flow Φ_3^t is well defined, at least for small times, and analytic. Since $F^{(3)}$ is a homogenous Hamiltonian of degree 3, then the vector field $X_{F^{(3)}}$ is a homogenous function of degree 2. Moreover by the lower bound (3.8) on the 3-resonances and the fact that $[\![H^{(3)}]\!] \lesssim 1$ we have

$$[\![F^{(3)}]\!] \lesssim 1. \tag{3.21}$$

Since $F^{(3)}$ is homogenous, in a sufficiently small neighbourhood of the origin the flow Φ_3^t is defined for times $t \in [0, 1]$. We call $\Phi_3 := \Phi_{F^{(3)}}^1$. We claim that

$$\|\Phi_3^t(Z)\|_{H^s} \leq 2\|Z\|_{H^s}, \qquad \|\Phi_3^t(Z)\|_{L^\infty} \leq 2\|Z\|_{L^\infty}, \qquad \forall t \in [0, 1], \tag{3.22}$$

for $Z \in B_1(H^s(\mathbb{T})) \cap B_r(L^\infty(\mathbb{T}))$ for $r > 0$ small enough.

We start by proving the bound for the L^∞-norm. We use a bootstrap argument. Let us call

$$T_* := \sup\{T \geq 0 : \|\Phi_3^t(Z)\|_{L^\infty} \leq 2\|Z\|_{L^\infty}\}.$$

We observe that $T_* > 0$. Assume that $T_* \geq 1$. We shall prove that at time $t = 1$ a better estimate on $\|\Phi_3^t(Z)\|_{L^\infty}$ holds and then T_* has to be greater than 1. We have that

$$\Phi_3^t(Z) = Z + \int_0^t X_{F^{(3)}}(\Phi_3^\tau(Z)) \, d\tau, \tag{3.23}$$

then, by Lemma 3.7-(i) and using that $Z \in B_1(H^s(\mathbb{T})) \cap B_r(L^\infty(\mathbb{T}))$, we have for $t \in [0, T_*)$

$$\|\Phi_3^t(Z)\|_{L^\infty} \leq \|Z\|_{L^\infty} + 2\,C\,[\![F]\!]r \int_0^t \|\Phi_{F^{(3)}}^\tau(Z)\|_{L^\infty} \, d\tau,$$

for some universal constant $C > 0$. By Gronwall Lemma

$$\|\Phi_3^t(Z)\|_{L^\infty} \leq \|Z\|_{L^\infty} \, \exp(2C[\![F]\!]rt).$$

Then taking $r \ll (2C[\![F^{(3)}]\!])^{-1}$ we have that $\|\Phi_{F^{(3)}}^1(Z)\|_{L^\infty} \leq (3/2)\|Z\|_{L^\infty}$. Hence we proved the claim on the bound of the L^∞-norm. Now we prove the one on the H^s norm. By Cauchy–Schwarz and Lemma 3.7-(i) we have, for $Z \in B_1(H^s(\mathbb{T})) \cap B_r(L^\infty(\mathbb{T}))$,

$$\frac{d}{dt}\|\Phi_3^t(Z)\|_{H^0}^2 \lesssim \|X_{F^{(3)}}(\Phi_3^t(Z))\|_{H^0} \|\Phi_3^t(Z)\|_{H^0}$$

$$\lesssim \|\Phi_3^t(Z)\|_{L^\infty} \|\Phi_3^t(Z)\|_{H^0}^2 \lesssim 2r\|\Phi_3^t(Z)\|_{H^0}^2.$$

Again by Gronwall Lemma we have $\|\Phi_3^t(Z)\|_{H^0}^2 \leq \|Z\|_{H^0}^2 \exp(2rCt)$ for all $t \in [0, T_*]$. Taking $r \ll (2rC)^{-1}$ we get $\|\Phi_3^t(Z)\|_{H^0} \leq 2\|Z\|_{H^0}$ for all $t \in [0, 1]$. By (3.23) and Lemma 3.7-(i)

$$\|\Phi_3^t(Z)\|_{H^s} \leq \|Z\|_{H^s} + C(s)[\![F^{(3)}]\!] \int_0^t \|\Phi_3^\tau(Z)\|_{L^\infty} \|\Phi_3^\tau(Z)\|_{H^0} \, d\tau$$

$$\leq \|Z\|_{H^s} + 2 C(s)[\![F^{(3)}]\!]r \int_0^t \|\Phi_3^\tau(Z)\|_{H^s} \, d\tau \, ,$$

for some constant $C(s) > 0$ depending on s. Reasoning as for the bound of the L^∞-norm we get the estimate $\|\Phi_3^t(Z)\|_{H^s} \leq 2\|Z\|_{H^s}$ for all $t \in [0, 1]$.

Now we claim that

$$\|d\Phi_3^t(Z)[\widetilde{Z}]\|_{H^s} \leq 2\|\widetilde{Z}\|_{H^s} \, , \qquad \|d\Phi_3^t(Z)[\widetilde{Z}]\|_{L^\infty} \leq 2\|\widetilde{Z}\|_{L^\infty} \qquad \forall t \in [0, 1] \, , \quad (3.24)$$

for $Z \in B_1(H^s(\mathbb{T})) \cap B_r(L^\infty(\mathbb{T}))$ for $r > 0$ small enough. We have

$$d\Phi_3^t(Z)[\widetilde{Z}] = \widetilde{Z} + \int_0^t dX_{F^{(3)}}(\Phi_{F^{(3)}}^\tau(Z)) \, [d\Phi_{F^{(3)}}^\tau(Z)[\widetilde{Z}]] \, d\tau \, . \quad (3.25)$$

By Lemma 3.7-(i) and (3.22) we have

$$\|d\Phi_3^t(Z)[\widetilde{Z}]\|_{L^\infty} \leq \|\widetilde{Z}\|_{L^\infty} + 2C[\![F^{(3)}]\!]r \int_0^t \|d\Phi_{F^{(3)}}^\tau(Z)[\widetilde{Z}]\|_{L^\infty} \, d\tau$$

for some universal constant $C > 0$. Taking r small enough and using Gronwall Lemma we obtain the bound for the L^∞-norm in (3.24). In the same way one can prove the bound for the H^s-norm.

By Lemma 3.7, bounds (3.21), (3.22) and the expression (3.23) we have for all $Z \in B_1(H^s(\mathbb{T})) \cap B_r(L^\infty(\mathbb{T}))$

$$\|\Phi_3(Z) - Z\|_{H^s} \leq \sup_{t \in [0,1]} \|X_{F^{(3)}}(\Phi_3^t(Z))\|_{H^s}$$
$$\lesssim_s \sup_{t \in [0,1]} \|\Phi_3^t(Z)\|_{L^\infty} \|\Phi_3^t(Z)\|_{H^0} \lesssim_s r\|Z\|_{H^0} \, , \quad (3.26)$$

which is a bound like (3.13). By using the estimates (3.24) we obtain a similar bound for the differential $d\Phi_3^t$. For the inverse of Φ_3 a similar estimate holds.

We obtain the new Hamiltonian by Taylor expanding $H \circ \Phi_3^t$ at $t = 0$. Hence

$$H_1 := H \circ \Phi_3^{-1} = H + \{F^{(3)}, H\} + \sum_{p=2}^3 \frac{1}{p!} \underbrace{\{F^{(3)}, \{F^{(3)}, \ldots, H\} \cdots \}}_{p-times} + \mathfrak{R}_1^{(\geq 6)}$$

$$\overset{(3.20)}{=} H^{(2)} + H^{(3,\geq 2)} + H_1^{(4)} + H_1^{(5)} + \mathfrak{R}_1^{(\geq 6)}$$

with

$$H_1^{(4)} := \frac{1}{2}\{F^{(3)}, H^{(3,\leq 1)}\} + \{F^{(3)}, H^{(3,\geq 2)}\} \, ,$$

$$\mathfrak{R}_1^{(\geq 6)} := \frac{1}{3!} \int_0^1 (1-t)^4 \{F^{(3)}, \{F^{(3)}, \{F^{(3)}, \{F^{(3)}, H\}\}\}\} \circ \Phi_3^{-t} \, dt \, ,$$

and where $H_1^{(5)}$ collect all the terms of degree of homogeneity 5. By (3.21) and the fact that $X_{F^{(3)}}$ is finite rank we have

$$[\![H_1^{(5)}]\!] \lesssim 1 \, , \qquad [\![\{F^{(3)}, \{F^{(3)}, \{F^{(3)}, \{F^{(3)}, H\}\}\}\}]\!] \lesssim 1 \, .$$

We observe that $H_1 - H^{(2)} - H^{(3,\geq 2)}$ generates a finite rank vector field. Moreover, using that

$$X_{\mathfrak{R}_1^{(\geq 6)}} = \frac{1}{3!} \int_0^1 (1-t)^4 d\Phi_3^t [X_{F^{(3)}}, [X_{F^{(3)}}, [X_{F^{(3)}}, X_H]]] \circ \Phi_3^{-t} \, dt,$$

Lemma 3.7-(i) and the estimates (3.22), (3.24) on the map Φ_3 and its differential one can check that $X_{\mathfrak{R}_1^{(\geq 6)}}$ has a zero at the origin of order at least 5 and it satisfies (3.17) and (3.19).

Second step: Now we look for a transformation that normalizes the term $H_1^{(4,\leq 1)}$. Let us consider the homogenous Hamiltonian

$$F^{(4)} = \sum_{\mathcal{A}_{4,\leq 1}} F_{\vec{\sigma},\vec{j}}^{(4)} z_{j_1}^{\sigma_1} z_{j_2}^{\sigma_2} z_{j_3}^{\sigma_3} z_{j_4}^{\sigma_4}, \qquad F_{\vec{\sigma},\vec{j}}^{(4)} := \begin{cases} \dfrac{H_{\vec{\sigma},\vec{j}}^{(4,\leq 1)}}{i\Omega(\vec{\sigma},\vec{j})} & \text{if } \Omega(\vec{\sigma},\vec{j}) \neq 0 \\ 0 & \text{if } \Omega(\vec{\sigma},\vec{j}) = 0. \end{cases}$$

By definition the Hamiltonian $F^{(4)}$ solves the homological equation

$$\{F^{(4)}, H^{(2)}\} + H_1^{(4,\leq 1)} = \Pi_{\mathrm{Ker}} H_1^{(4,\leq 1)}. \tag{3.27}$$

By item (ii) of Lemma 3.8 we have that $\Pi_{\mathrm{Ker}} H_1^{(4,\leq 1)} = \Pi_{\mathrm{Ker}} H_1^{(4,0)}$. By Remark 3.4 $\Pi_{\mathrm{Ker}} H_1^{(4,0)}$ is a homogenous Hamiltonian action-preserving of degree 4, then it has the form (3.16). By Lemma 3.8-(iv) and by considering that $\mathcal{A}_{4,\leq 1}$ is finite we have

$$[\![F^{(4)}]\!] \lesssim 1.$$

Reasoning as in the first step and using item (i) of Lemma 3.7 we have that the flow

$$\Phi_{F^{(4)}}^t : B_1(H^s(\mathbb{T})) \cap B_r(L^\infty(\mathbb{T})) \to B_2(H^s(\mathbb{T})) \cap B_{2r}(L^\infty(\mathbb{T}))$$

for times $t \in [0,1]$ and $r > 0$ small enough (satisfying a condition like (3.11)), has estimates as in (3.22) (and similar for its differential as in (3.24)). We call $\Phi_4 := \Phi_{F^{(4)}}^1$ and we have, reasoning as in (3.26),

$$\|\Phi_4(Z) - Z\|_{H^s} \lesssim_s r^2 \|Z\|_{H^0}, \qquad \forall Z \in B_1(H^s(\mathbb{T})) \cap B_r(L^\infty(\mathbb{T})). \tag{3.28}$$

Using the smallness condition (3.11), we have that (3.28) implies a bound like (3.13). The bound (3.14) follows by reasoning as in the previous step. We obtain the new Hamiltonian by Taylor expanding $H \circ \Phi_{F^{(4)}}^{-t}$ at $t = 0$

$$H_2 := H \circ \Phi_4^{-1} = H_1 + \{F^{(4)}, H_1\} + \int_0^1 (1-t)\{F^{(4)}, \{F^{(4)}, H_1\}\} \circ \Phi_{F^{(4)}}^{-t} \, dt$$

$$\overset{(3.27)}{=} H^{(2)} + H^{(3,\geq 2)} + \Pi_{\mathrm{Ker}} H_1^{(4,0)} + H_1^{(4,\geq 2)} + H_2^{(5)} + \mathfrak{R}_2^{(\geq 6)}$$

with

$$H_2^{(5)} = \{F^{(4)}, H^{(3,\geq 2)}\} + H_1^{(5)},$$

$$\mathfrak{R}_2^{(\geq 6)} := \mathfrak{R}_1^{(\geq 6)} + \{F^{(4)}, H_1^{(4)} + \mathfrak{R}_1^{(\geq 6)}\} + \int_0^1 (1-t)\{F^{(4)}, \{F^{(4)}, H_1\}\} \circ \Phi_{F^{(4)}}^t \, dt.$$

We have

$$[\![H_2^{(5)}]\!] \lesssim L^2, \qquad [\![\{F^{(4)}, \{F^{(4)}, H_1\}\}]\!] \lesssim 1.$$

We observe that $H_1^{(4,\geq 2)}$, $H_2^{(5)}$, $\mathfrak{R}_2^{(\geq 6)}$ generate finite rank vector fields. Moreover, reasoning as in the previous step, one can check that $X_{\mathfrak{R}_2^{(\geq 6)}}$ has a zero at the origin of order at least 5 and satisfies (3.17), (3.19).

Third step: We want to normalize $H_2^{(5,\leq 1)}$. Let us consider the homogenous Hamiltonian

$$F^{(5)} = \sum_{\mathcal{A}_{5,\leq 1}} F_{\vec{\sigma},\vec{j}}^{(5)} z_{j_1}^{\sigma_1} z_{j_2}^{\sigma_2} z_{j_3}^{\sigma_3} z_{j_4}^{\sigma_4} z_{j_5}^{\sigma_5}, \qquad F_{\vec{\sigma},\vec{j}}^{(5)} := \begin{cases} \dfrac{H_{\vec{\sigma},\vec{j}}^{(5,\leq 1)}}{i\Omega(\vec{\sigma},\vec{j})} & \text{if } \Omega(\vec{\sigma},\vec{j}) \neq 0 \\[2ex] 0 & \text{if } \Omega(\vec{\sigma},\vec{j}) = 0. \end{cases}$$

By definition the Hamiltonian $F^{(5)}$ solves the homological equation

$$\{F^{(5)}, H^{(2)}\} + H_2^{(5,\leq 1)} = \Pi_{\mathrm{Ker}} H_2^{(5,\leq 1)}.$$

By item (ii) of Lemma 3.8 we have that $\Pi_{\mathrm{Ker}} H_2^{(5,\leq 1)} = 0$. By item (iii) of Lemma 3.8 and by considering that $\mathcal{A}_{5,\leq 1}$ is finite we have

$$[\![F^{(5)}]\!] \lesssim 1.$$

Reasoning as in the first step and using item (i) of Lemma 3.7 we have that the flow

$$\Phi_{F^{(5)}}^t : B_1(H^s(\mathbb{T})) \cap B_r(L^\infty(\mathbb{T})) \to B_2(H^s(\mathbb{T})) \cap B_{2r}(L^\infty(\mathbb{T}))$$

for times $t \in [0,1]$ and $r > 0$ satisfying a condition like (3.11), has estimates as in (3.22) (and similar for its differential as in (3.24)). We call $\Phi_5 := \Phi_{F^{(5)}}^1$ and we have, reasoning as in (3.26),

$$\|\Phi_5(Z) - Z\|_{H^s} \lesssim_s r^3 \|Z\|_{H^0}, \qquad \forall Z \in B_r(L^\infty(\mathbb{T})). \tag{3.29}$$

Using the smallness condition (3.11) (taking $\mathtt{a} > 0$ large enough), we have that (3.29) implies a bound like (3.13). The bound (3.14) follows by reasoning as in the previous step using the fact that Φ_5 is the time one flow map generated by the vector field $X_{F^{(5)}}$. We obtain the new Hamiltonian by Taylor expanding $H \circ \Phi_{F^{(5)}}^{-t}$ at $t = 0$

$$H_3 := H \circ \Phi_5^{-1} = H_2 + \{F^{(5)}, H_2\} + \int_0^1 (1-t)\{F^{(5)}, \{F^{(5)}, H_2\}\} \circ \Phi_{F^{(5)}}^{-t}\, dt$$

$$\overset{(3.27)}{=} H^{(2)} + H^{(3,\geq 2)} + \Pi_{\mathrm{Ker}} H_1^{(4,0)} + H_1^{(4,\geq 2)} + H_2^{(5,\geq 2)} + \mathfrak{R}_3^{(\geq 6)}$$

with

$$\mathfrak{R}_3^{(\geq 6)} := \mathfrak{R}_2^{(\geq 6)} + \{F^{(5)}, H_1^{(4)} + H_2^{(5)} + \mathfrak{R}_2^{(\geq 6)}\} + \int_0^1 (1-t)\{F^{(5)}, \{F^{(5)}, H_2\}\} \circ \Phi_{F^{(5)}}^{-t}\, dt.$$

We observe that $H_1^{(4,\geq 2)}$, $H_2^{(5,\geq 2)}$, $\mathfrak{R}_3^{(\geq 6)}$ generate finite rank vector fields. Moreover, reasoning as in the previous step, one can check that $X_{\mathfrak{R}_3^{(\geq 6)}}$ has a zero at the origin of order at least 5 and satisfies (3.17), (3.19).

Now we define

$$\Phi_B := \Phi_5 \circ \Phi_4 \circ \Phi_3, \qquad \mathfrak{R}^{(\geq 6)} := \mathfrak{R}_3^{(\geq 6)}.$$

Then formula (3.15) holds. Since each Birkhoff map has the form (3.12) then Φ_B has the same form. By the estimates (3.26), (3.28), (3.29) and using the smallness condition (3.11) we obtain (3.13). The bound (3.14) follows by composition since the maps Φ_i, $i = 3, 4, 5$ satisfies the same properties.

3.3 The Dynamics of the Normalized Hamiltonian

Let $S \subset \mathbb{Z}$ be any finite set. In this section we study the dynamics of the first order of the Hamiltonian which has been normalized following the procedure described in Proposition 3.9. We call

$$\mathcal{N} = H^{(2)} + H^{(3,\geq 2)} + \mathcal{H}_{res}^{(4,0)} + \mathcal{H}^{(4,\geq 2)} + \mathcal{H}^{(5,\geq 2)} \tag{3.30}$$

the Hamiltonian in the Birkhoff coordinates Φ_B provided by Proposition 3.9, up to the remainder $\mathfrak{R}^{(\geq 6)}$. We observe that the vector field of $H^{(3,\geq 2)} + \mathcal{H}^{(4,\geq 2)} + \mathcal{H}^{(5,\geq 2)}$ vanishes on the finite dimensional subspace

$$\mathcal{U}_S = \{w_n = 0 \ n \notin S\}.$$

The restricted Hamiltonian $\mathcal{N}_{|\mathcal{U}_S}$ coincides then with

$$\mathcal{H}_{res} = H^{(2)} + \mathcal{H}_{res}^{(4,0)}. \tag{3.31}$$

Thus, the equations of motion read as

$$\dot{w}_n = \mathrm{i} \left(\Lambda(n) + \sum_{k \in S} C_{nk} |w_k|^2 \right) w_n, \qquad n \in S. \tag{3.32}$$

For all subsets $\tilde{S} \subseteq S$ the subspace

$$\mathcal{V}_{\tilde{S}} = \{w_n = 0 \ n \notin \tilde{S}\}$$

is invariant by the flow of (3.32). The actions $|w_j|^2$, $j \in S$ are all conserved quantities. Then the complementary of $\bigcup_{\tilde{S} \subset S} \mathcal{V}_{\tilde{S}}$ is foliated by maximal (n-dimensional) invariant tori $\{|w_j|^2 = \xi_j, \ j \in S\}$ with $\xi_j > 0$, $j \in S$. They support quasi-periodic motions given by

$$\varphi(\xi, \theta; t, x) = \sum_{j \in S} \sqrt{\xi_j} \, e^{\mathrm{i}(\theta_j + \omega_j(\xi)t + jx)}, \qquad \xi = (\xi_j)_{j \in S}, \qquad \theta = (\theta_j)_{j \in S} \qquad \theta_j \in \mathbb{T}, \tag{3.33}$$

where the frequency vector has components

$$\omega_j(\xi) = \Lambda(j) + \sum_{k \in S} C_{jk} \xi_k \qquad j \in S. \tag{3.34}$$

We observe that these frequencies can be very close to resonant, depending on the choice of the ξ_j's. The subspaces $\mathcal{V}_{\tilde{S}}$ are foliated by lower dimensional invariant tori supporting the quasi-periodic motions

$$\varphi(\xi, \theta; t, x) = \sum_{j \in \tilde{S}} \sqrt{\xi_j} \, e^{\mathrm{i}(\theta_j + \omega_j(\xi)t + jx)}, \qquad \theta_j \in \mathbb{T}.$$

4 Setting of the Problem in the Birkhoff Coordinates

Let us fix some $N \in \mathbb{N}$ and consider a finite symmetric subset such that

$$\begin{aligned} S = \{j_1, \ldots, j_N\} \subset \mathbb{Z}, \qquad j \in S \quad &\Rightarrow \quad -j \in S, \\ j \notin S \quad &\Leftrightarrow \quad |j| > \max_{k \in S}\{|k|\}. \end{aligned} \tag{4.1}$$

We need the above properties of the set S only for then construction of the modified energy on high frequencies in Sect. 5.2.2, see Lemma 5.16. Consider the new variables

$$W := \begin{bmatrix} w \\ \bar{w} \end{bmatrix} := \Phi_B(Z)$$

where Φ_B is the Birkhoff map constructed in Proposition 3.9 applied with S as in (4.1). By the symplectic nature of the Birkhoff map we have that, the equation $\partial_t Z = X_H(Z)$ with H in (2.12), becomes

$$\partial_t W = X_{\mathcal{H}}(W), \tag{4.2}$$

where $X_{\mathcal{H}}$ is the vector field of the new Hamiltonian $\mathcal{H} = H \circ \Phi_B^{-1}$ in (3.15).

4.1 The Approximate Solution

We consider a solution $\varphi(t, x)$ of the normalized Hamiltonian system $\mathcal{N}$ with initial condition on the subspace $\mathcal{V}_S$. We could consider also solutions with initial data in $\mathcal{V}_{\tilde{S}}$ with $\tilde{S} \subset S$ symmetric without relevant changes in the proof.

The function $\varphi(t, x) = \varphi(\xi, \theta; t, x)$ has the form (3.33), where $\xi = (\xi_j)_{j \in S}$ can be any vector with positive components and $\theta = (\theta_j)_{j \in S}$ is any vector of angles. Since the normalized change of coordinates Φ_B provided in Proposition 3.9 is well defined only on a small neighbourhood of the origin of L^∞, it is convenient to rescale the actions $\xi_j \mapsto \varepsilon^2 \xi_j$, with $\varepsilon > 0$ small enough such that if $\varepsilon = r$ condition (3.11) is satisfied. Then the rescaled solution has the form

$$\varepsilon\varphi(t, x) = \varepsilon \sum_{j \in S} \sqrt{\xi_j}\, e^{\mathrm{i}(\theta_j + \omega_j(\xi)t + jx)}. \tag{4.3}$$

We observe that

$$\sup_{t \in \mathbb{R}} \|\varepsilon\varphi\|_{H^s} \lesssim \varepsilon \qquad \forall s \geq 0. \tag{4.4}$$

In the following we shall assume further smallness conditions on the parameter ε. Such conditions can all be written in the following form

$$C_s \varepsilon < 1, \tag{4.5}$$

for some constant $C_s > 0$ depending on $s > 1/2$ and the set S.

The next lemma shows that the functions $\varepsilon\varphi$ constructed in (4.3) are approximate solutions of the Hamiltonian equation given by $\mathcal{H}$.

Lemma 4.1 *Let $s > 1/2$. There $C_s > 1$ such that if (4.5) holds, then the residual*

$$\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi) := -\varepsilon\partial_t\varphi + X_{\mathcal{H}}(\varepsilon\varphi) \tag{4.6}$$

satisfies

$$\sup_{t \in [0,T]} \|\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)\|_{H^s} \lesssim_s \varepsilon^5. \tag{4.7}$$

Proof We have that $X_{\mathcal{H}} = X_{\mathcal{N}} + X_{\mathfrak{R}^{(\geq 6)}}$. Since $\varepsilon\varphi$ is a solution of the Hamiltonian system given by $\mathcal{N}$ we have

$$\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi) = X_{\mathfrak{R}^{(\geq 6)}}(\varepsilon\varphi).$$

Then the thesis follows by estimates (3.17) and (4.4) taking ε small enough.

4.2 The Error Function in Birkhoff Coordinates

We set

$$\varepsilon^\beta V := W - \varepsilon\varphi, \qquad \text{for } \beta > 2, \tag{4.8}$$

where $\varepsilon\varphi$ is the approximate solution of the form (4.3), supported on a set S as in (4.1), and with frequencies of oscillation given in (3.34). We also recall that estimate (4.4) holds true. Then the error function V in the Birkhoff coordinates solves the equation (recall the definition of $\mathrm{Res}_\mathcal{H}$ in (4.6))

$$\dot{V} = dX_\mathcal{H}(\varepsilon\varphi)[V] + \varepsilon^\beta \mathcal{Q}(\varepsilon\varphi)[V, V] + \varepsilon^{-\beta}\mathrm{Res}_\mathcal{H}(\varepsilon\varphi), \tag{4.9}$$

where

$$\mathcal{Q}(\varepsilon\varphi)[V, V] := \int_0^1 (1 - t)\, d^2 X_\mathcal{H}(\varepsilon\varphi + t\varepsilon^\beta V)[V, V]\, dt. \tag{4.10}$$

In the next proposition we show that the control on the function V implies the main result Theorem 1.1.

Proposition 4.2 *There is $C_s > 1$ such that if*

$$\sup_{t\in[0,T]} \varepsilon^\beta \|V\|_{H^s} \le 2\varepsilon^\beta, \quad \beta > 2, \tag{4.11}$$
$$\sup_{t\in[0,T]} \|Z\|_{H^s} \lesssim_s \varepsilon, \quad T = c_0\varepsilon^{-2-\sigma}, \quad \sigma > 0,$$

for some $c_0 > 0$ and with ε satisfying (4.5), then the error $\varepsilon^\beta R := Z - \varepsilon\varphi$ satisfies the bound

$$\sup_{t\in[0,T]} \varepsilon^\beta \|R\|_{H^s} \lesssim_s \varepsilon^2. \tag{4.12}$$

Proof Recall that $Z = \Phi_B^{-1}(W)$. Then, by estimate (3.13) in Prop. 3.9, which holds also for Φ_B^{-1}, and the assumptions (4.11), (4.5) with $C_s > 1$ large enough, we deduce that

$$\|W\|_{H^s} \lesssim_s \varepsilon, \quad \text{and} \quad \|Z - W\|_{H^s} = \|\Phi_B^{-1}(W) - W\|_{H^s} \lesssim_s \varepsilon\|W\|_{H^s} \lesssim_s \varepsilon^2,$$

uniformly in $t \in [0, T]$. Then, recalling $\beta > 2$, one has

$$\|Z - \varepsilon\varphi\|_{H^s} \overset{(4.8)}{\le} \|Z - W\|_{H^s} + \|\varepsilon^\beta V\|_{H^s} \overset{(4.2)}{\lesssim_s} \varepsilon^2 + \varepsilon^\beta \lesssim_s \varepsilon^2,$$

uniformly in $t \in [0, T]$. This implies (4.12). $\qquad\blacksquare$

The result above guarantees that in order to obtain our main result we must show that the solution V of (4.9) satisfies the (4.11). To do this we will provide some *à priori* estimates on V.

Recalling (4.9) the main issues are the following:

- Show that the term $dX_\mathcal{H}(\varepsilon\varphi)[V] + \varepsilon^\beta \mathcal{Q}(\varepsilon\varphi)[V, V]$ has a pseudo-differential structure. This is the content of Sect. 4.3.
- In Sects. 5–6 we provide the energy estimates for the flow.

4.3 Pseudo-Differential Structure of the Equation for the Remainder

As explained before, in the next sections we shall provide a priori bounds (see (4.11)) on the solution V of the problem (4.9). In Sect. 3.3 we provided suitable upper bounds on the non-homogenenous term $\varepsilon^{-\beta} \text{Res}_{\mathcal{H}}(\varepsilon\varphi)$ appearing in the right hand side of (4.9) (see Lemma 4.1). Hence, we shall prove that the vector field

$$\mathcal{Y}(\varepsilon\varphi, V) := dX_{\mathcal{H}}(\varepsilon\varphi)[V] + \mathcal{Q}(\varepsilon\varphi)[V, V], \tag{4.13}$$

which is nonlinear in V, generates a well-defined flow on the spaces H^s. In this subsection we show that actually $\mathcal{Y}(\varepsilon\varphi, V)$ has a *pseudo-differential* structure. We deal with the two summands in (4.13) separately.

Define the following real symbols (recall (1.2)–(1.3))

$$
\begin{aligned}
a(U; x) &:= -\tfrac{1}{2}(\partial_{\phi_{xx}} f_2)(\phi, \phi_x, \phi_{xx}), \qquad \phi = \tfrac{\Lambda^{-\frac{1}{2}}}{\sqrt{2}}(u + \overline{u}), \quad U = \begin{bmatrix} u \\ \overline{u} \end{bmatrix}, \\
b(U; x) &:= (\partial_{\phi_x} f_2)(\phi, \phi_x, \phi_{xx}), \\
c(U; x) &:= (\partial_{\phi} f_2)(\phi, \phi_x, \phi_{xx}) \\
d(U; x) &:= -ma(U; x) - \tfrac{3}{8}a_{xx}(U; x) - \tfrac{1}{2}b_x(U; x) + c(U; x).
\end{aligned}
\tag{4.14}
$$

Consider the functions in (4.8) and the symbols in (4.14) with $U \rightsquigarrow \varepsilon\varphi, V$. Let us define

$$
\begin{aligned}
\mathtt{f}(\varepsilon\varphi, V; x) &:= a(\varepsilon\varphi; x) + \tfrac{1}{2}a(V; x), \\
\mathtt{g}(\varepsilon\varphi, V; x) &:= d(\varepsilon\varphi; x) + \tfrac{1}{2}d(V; x).
\end{aligned}
\tag{4.15}
$$

As a consequence of Propositions 4.4, 4.8 we deduce the following result.

Proposition 4.3 *The vector field $\mathcal{Y}(\varepsilon\varphi, V)$ in (4.13) has the form*

$$
\begin{aligned}
\mathcal{Y}(\varepsilon\varphi, V) &= \mathrm{i} E Op^{\mathrm{BW}}\left(\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\mathtt{f}(\varepsilon\varphi, V; x)\right)\Lambda(\xi)\right)V \\
&\quad + \mathrm{i} E Op^{\mathrm{BW}}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\mathtt{g}(\varepsilon\varphi, V; x)\Lambda^{-1}(\xi)\right)V + \mathrm{i} E \mathfrak{Q}(\varepsilon\varphi)V + R(V),
\end{aligned}
\tag{4.16}
$$

where the remainder (recall Def. 2.3)

$$\mathfrak{Q}(\varepsilon\varphi) = \mathfrak{Q}_1(\varepsilon\varphi) + \mathfrak{Q}_2(\varepsilon\varphi) + \mathfrak{Q}_{\geq 3}(\varepsilon\varphi) \in \Sigma_1^{-2}[r, 3] \otimes \mathcal{M}_2(\mathbb{C}),$$

is real *according to Def. 2.13. Moreover the non-homogeneous component $\mathfrak{Q}_{\geq 3}(\varepsilon\varphi) \in \mathbf{NH}_3^{-2}[r] \otimes \mathcal{M}_2(\mathbb{C})$ satisfy bounds like (2.35). Moreover for s large enough the remainder $R(V)$ has the form*

$$(R^+(V), \overline{R^+(V)})^T$$

and it satisfies

$$\|R(V)\|_{H^{s+2}} \lesssim_s \|V\|_{H^s}^2 \qquad \forall V \in B_\varepsilon(H^s). \tag{4.17}$$

Eventually, for s large enough and ε satisfying (4.5), one has

$$\|\mathcal{Y}(\varepsilon\varphi, V)\|_{H^{s-1}} \lesssim_s \|V\|_{H^s}, \qquad \forall V \in B_{\varepsilon\sqrt{L}}(H^s). \tag{4.18}$$

4.3.1 Para-Differential Structure of the Linearized Operator

Consider the function $\varepsilon\varphi$ in (4.8) and recall the bound (4.4). The aim of this section is to prove the following result.

Proposition 4.4 *There exists $s_0 > 0$ large enough such that for all $s \geq s_0$ there is $C_s > 0$ such that the following holds. If ε satisfies (4.5) and $\varepsilon\varphi$ defined in (4.8) satisfies (4.4) then we have that $a(\varepsilon\varphi; x), d(\varepsilon\varphi; x) \in \mathbf{SM}_1^0$ (see Def. 2.5) are independent of $\xi \in \mathbb{R}$ and real valued, and*

$$
\begin{aligned}
(dX_{\mathcal{H}})(\varepsilon\varphi)[V] = {}& \mathrm{i}E\,Op^{\mathrm{BW}}\left(\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} a(\varepsilon\varphi; x)\right)\Lambda(\xi)\right)V \\
& + \mathrm{i}E\,Op^{\mathrm{BW}}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} d(\varepsilon\varphi; x)\Lambda^{-1}(\xi)\right) + \mathrm{i}E\mathfrak{Q}(\varepsilon\varphi)V\,,
\end{aligned}
\tag{4.19}
$$

where $\mathcal{H}$ in (3.15), $\Lambda(\xi)$ is in (2.5) and where the remainder

$$
\mathfrak{Q}(\varepsilon\varphi) = \mathfrak{Q}_1(\varepsilon\varphi) + \mathfrak{Q}_2(\varepsilon\varphi) + \mathfrak{Q}_{\geq 3}(\varepsilon\varphi) \in \Sigma_1^{-2}[r, 3] \otimes \mathcal{M}_2(\mathbb{C}),
$$

is real *according to Def. 2.13. Moreover the non-homogeneous component $\mathfrak{Q}_{\geq 3}(\varepsilon\varphi) \in \mathbf{NH}_3^{-2}[r] \otimes \mathcal{M}_2(\mathbb{C})$ satisfy bounds like (2.35). Finally the following estimate holds: for $s \geq s_0$*

$$
\|dX_{\mathcal{H}}(\varepsilon\varphi)[Z]\|_{H^{s-1}} \lesssim_s \|Z\|_{H^s}(1 + C_s\|\varepsilon\varphi\|_{H^s})\,,
\tag{4.20}
$$

for any $Z \in H^s$.

Remark 4.5 We remark that the operator $(dX_{\mathcal{H}})(\varepsilon\varphi)[\cdot]$ is *Hamiltonian* according to Definition 2.14.

We have the following.

Lemma 4.6 *Recall (3.15) and let $X := X_{\mathcal{H}_{res}^{(4,0)}} + X_{\mathcal{H}^{(4,\geq 2)}} + X_{\mathcal{H}^{(5,\geq 2)}}$. Then we have*

$$
(dX)(\varepsilon\varphi)[\cdot] := \sum_{j=2}^{3} \mathcal{S}_j(\varepsilon\varphi)[\cdot],
$$

for some real $\mathcal{S}_j(\varphi) \in \mathbf{M}_j^{-\rho} \otimes \mathcal{M}_2(\mathbb{C})$, for any $\rho \geq 0$. Moreover the coefficients of $\mathcal{S}_2(\varepsilon\varphi)$ and $\mathcal{S}_3(\varepsilon\varphi)$ satisfy (2.34).

Proof It follows recalling that X is sum of finite rank vector fields (recall Def. 3.6). ∎

In view of Lemma 4.6, we have that the linearized operator of the vector field $X_{\mathcal{H}}$ has the form

$$
(dX_{\mathcal{H}})(\varepsilon\varphi)[V] = \mathrm{i}E\Lambda V + (dX_{H^{(3,\geq 2)}})(\varepsilon\varphi)[V] + \mathcal{S}(\varepsilon\varphi)[V]\,,
\tag{4.21}
$$

for some $\mathcal{S}(\varepsilon\varphi) \in \Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ (see Def. 2.3) for any $\rho \geq 0$. To study the contribution coming from the Hamiltonian $H^{(3,\geq 2)}$ we first analyse the linearized operator of $X_{H^{(3)}}$ where $H^{(3)}$ is the Hamiltonian in (2.12). We have the following Lemma.

Lemma 4.7 *Consider the symbols $a(\varepsilon\varphi; x), d(\varepsilon\varphi; x)$ in (4.14). Then we have that $a(\varepsilon\varphi; x), d(\varepsilon\varphi; x)$ belong to $\mathbf{SM}_1^0$ (see Def. 2.5), are independent of $\xi \in \mathbb{R}$ and real valued, and*

$$
(dX_{H^{(3)}})(\varepsilon\varphi)[V] = \mathrm{i}E\,Op^{\mathrm{BW}}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} a(\varepsilon\varphi; x)\Lambda(\xi) + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} d(\varepsilon\varphi; x)\Lambda^{-1}(\xi)\right)V + \widetilde{\mathcal{S}}(\varepsilon\varphi)V
\tag{4.22}
$$

for some real $\widetilde{\mathcal{S}}(\varepsilon\varphi) \in \mathbf{M}_1^{-2} \otimes \mathcal{M}_2(\mathbb{C})$, where $\Lambda(\xi)$ is in (2.5).

Proof Recall that

$$
X_{H^{(3)}} = \begin{pmatrix} X_{H^{(3)}}^+ \\ \overline{X_{H^{(3)}}^+} \end{pmatrix}, \qquad X_{H^{(3)}}^+ (v) \overset{(1.4)}{=} \frac{\mathrm{i}}{\sqrt{2}} \Lambda^{-1/2} N \left(\Lambda^{-1/2} \left(\frac{v + \overline{v}}{\sqrt{2}} \right) \right)
$$

where N is the nonlinearity in (1.2), (1.3). Therefore, by linearizing at (recall $\mathcal{C}$ in (2.10)) $\varepsilon \varphi = \mathcal{C} \begin{bmatrix} u \\ \overline{u} \end{bmatrix}$, we get

$$
\begin{aligned}
(dX_{H^{(3)}}^+)(\varepsilon\varphi)[V] &= \frac{\mathrm{i}\Lambda^{-1/2}}{\sqrt{2}} \big((\partial_u f_2)(u, u_x, u_{xx}) h \big) \\
&\quad + \frac{\mathrm{i}\Lambda^{-1/2}}{\sqrt{2}} \big((\partial_{u_x} f_2)(u, u_x, u_{xx}) h_x + (\partial_{u_{xx}} f_2)(u, u_x, u_{xx}) h_{xx} \big)
\end{aligned}
\tag{4.23}
$$

where $h := \Lambda^{-1/2}(v + \overline{v})/\sqrt{2}$ and $V = \begin{bmatrix} v \\ \overline{v} \end{bmatrix}$. We now expand in decreasing symbols the above operators. Since the function G in (1.3) is C^∞ and cubic it is easy to check that the functions $\partial_u f_2, \partial_{u_x} f_2, \partial_{u_{xx}} f_2$ are symbols independent of $\xi \in \mathbb{R}$ and in the class $\mathbf{SM}_1^0$ (see Def. 2.5). As a consequence all the symbols defined in (4.14) belongs to $\mathbf{SM}_1^0$ as well.

We write

$$
a(\varepsilon\varphi; x) = a_\chi(\varepsilon\varphi; x, \xi) + B(\varepsilon\varphi; x, \xi), \qquad B(\varepsilon\varphi; x, \xi) := \mathcal{F}^{-1}\Big(\widehat{a}(\eta)\big(1 - \chi(|\eta|/\langle\xi\rangle)\big) \Big).
\tag{4.24}
$$

The operator $Op^W(B)$ has the form (2.30), (2.31) with

$$
r^\sigma(j, j_1, j_2) = a\left(j_1, \frac{j_1 + 2j_2}{2} \right) \left(1 - \chi\left(\frac{2j_1}{j_1 + 2j_2} \right) \right), \qquad j = j_1 + j_2.
$$

Let us consider $Op^W(\widetilde{B}) = Op^W(\widetilde{B}(\varepsilon\varphi; x, \xi))$ as the multilinear operator with the form (2.30), (2.31) with coefficients

$$
\tilde{r}^\sigma(j, j_1, j_2) = \begin{cases} r^\sigma(j, j_1, j_2) & \text{if } j_1 \in S, \\ 0 & \text{if } j_1 \notin S. \end{cases}
$$

Since $r^\sigma(j, j_1, j_2) \neq 0$ implies that $\langle j_1 \rangle \geq \langle j + j_2 \rangle$ and $S \subset \mathbb{Z}$ is finite, we have $Op^W(\widetilde{B}) \in \mathbf{M}_1^{-2}$.

Since $\varepsilon\varphi$ is compactly Fourier supported (its Fourier support is contained in S) then $Op^W(B) = Op^W(\widetilde{B})$. By abuse of notation we say that $Op^W(B) \in \mathbf{M}_1^{-2}$.

One can reason in the same way to deal with the term depending on the symbol $d(\varepsilon\varphi; x)$. By the symbolic calculus in Lemma 2.2 we get that

$$
\begin{aligned}
&\Lambda^{-1/2} Op^{BW}\big(\partial_{u_{xx}} f_2\big) \partial_{xx} \Lambda^{-1/2} \\
&= Op^{BW}\Big(-\xi^2 \partial_{u_{xx}} f_2 \Lambda^{-1}(\xi) - \frac{1}{2\mathrm{i}} \{\Lambda^{-1/2}(\xi), \xi^2 \Lambda^{-1/2}(\xi) \partial_{u_{xx}} f_2\} \Big) \\
&\quad - Op^{BW}\Big(\frac{\Lambda^{-1/2}(\xi)}{2\mathrm{i}} \{\partial_{u_{xx}} f_2, \xi^2 \Lambda^{-1/2}(\xi)\} + \frac{\Lambda^{-\frac{1}{2}}(\xi)}{8} \sigma(\partial_{u_{xx}} f_2, \xi^2 \Lambda^{-\frac{1}{2}}(\xi)) \Big) \\
&\quad + Op^{BW}\Big(\frac{1}{8} \sigma\big(\Lambda^{-\frac{1}{2}}(\xi), \partial_{u_{xx}} f_2 \xi^2 \Lambda^{-\frac{1}{2}}(\xi)\big) + \frac{1}{4} \{\Lambda^{-\frac{1}{2}}(\xi), \{\partial_{u_{xx}} f_2, \xi^2 \Lambda^{-\frac{1}{2}}(\xi)\}\} \Big) \\
&= Op^{BW}\big(-\xi^2 \partial_{u_{xx}} f_2 \Lambda^{-1}(\xi) - (\partial_{u_{xx}} f_2)_x \mathrm{i}\xi \Lambda^{-1}(\xi) + a^{(-1)}(x, \xi) \big)
\end{aligned}
$$

up to smoothing remainders in the class $\mathbf{M}_1^{-2}$ and where

$$
\begin{aligned}
\mathrm{a}^{(-1)}(x,\xi) := & -\frac{\Lambda^{-\frac{1}{2}}(\xi)}{8}\sigma\left(\partial_{u_{xx}}f_2, \xi^2\Lambda^{-\frac{1}{2}}(\xi)\right) + \frac{1}{8}\sigma\left(\Lambda^{-\frac{1}{2}}(\xi), \partial_{u_{xx}}f_2\xi^2\Lambda^{-\frac{1}{2}}(\xi)\right) \\
& + \frac{1}{4}\left\{\Lambda^{-\frac{1}{2}}(\xi), \{\partial_{u_{xx}}f_2, \xi^2\Lambda^{-\frac{1}{2}}(\xi)\}\right\}.
\end{aligned}
$$

(4.25)

Similarly we have

$$
\begin{aligned}
& \Lambda^{-1/2}Op^{\mathrm{BW}}\left(\partial_u f_2\right)\Lambda^{-1/2} + \Lambda^{-1/2}Op^{\mathrm{BW}}\left(\partial_{u_x}f_2\right)\partial_x\Lambda^{-1/2} \\
& = Op^{\mathrm{BW}}\left(\mathrm{i}\xi\Lambda^{-1}(\xi)(\partial_{u_x}f_2) + \mathrm{b}^{(-1)}(x,\xi)\right)
\end{aligned}
$$

up to smoothing remainders in the class $\mathbf{M}_1^{-2}$ and where

$$
\mathrm{b}^{(-1)}(x,\xi) := \frac{\Lambda^{-\frac{1}{2}}(\xi)}{2\mathrm{i}}\left\{\partial_{u_x}f_2, \mathrm{i}\xi\Lambda^{-\frac{1}{2}}(\xi)\right\} + \frac{1}{2\mathrm{i}}\left\{\Lambda^{-\frac{1}{2}}(\xi), \partial_{u_x}f_2\mathrm{i}\xi\Lambda^{-\frac{1}{2}}(\xi)\right\} + \partial_u f_2\Lambda^{-1}(\xi).
$$

(4.26)

Therefore

$$
\begin{aligned}
(dX_{H^{(3)}}^+)(\varepsilon\varphi)[V] = & \frac{\mathrm{i}}{2}Op^{\mathrm{BW}}\left(-\xi^2\partial_{u_{xx}}f_2\Lambda^{-1}(\xi) + \mathrm{i}\xi\Lambda^{-1}(\xi)\left(\partial_{u_x}f_2 - (\partial_{u_{xx}}f_2)_x\right)\right)[v+\overline{v}] \\
& + \frac{\mathrm{i}}{2}Op^{\mathrm{BW}}\left(\mathrm{a}^{(-1)}(x,\xi) + \mathrm{b}^{(-1)}(x,\xi)\right)[v+\overline{v}]
\end{aligned}
$$

(4.27)

with $\mathrm{a}^{(-1)}, \mathrm{b}^{(-1)}$ in (4.25), (4.26), up to some remainder in $\mathbf{M}_1^{-2}$. Now we show that, thanks to the Hamiltonian structure of the nonlinearity, there is a cancelation of the terms of order zero. Indeed, by (1.2)–(1.3), one has

$$
\begin{aligned}
f_2 = & (\partial_u G)(u, u_x) - \partial_{uu_x}G(u, u_x)u_x - \partial_{u_xu_x}G(u, u_x)u_{xx}, \\
\partial_{u_x}f_2 = & (\partial_{uu_x}G)(u, u_x) - (\partial_{u_xu}G)(u, u_x) \\
& - (\partial_{uu_xu_x}G)(u, u_x)u_x - (\partial_{u_xu_xu_x}G)(u, u_x)u_{xx}, \\
\partial_{u_{xx}}f_2 = & -(\partial_{u_xu_x}G)(u, u_x), \\
(\partial_{u_{xx}}f_2)_x = & \partial_{u_x}f_2.
\end{aligned}
$$

Moreover, using formulæ (2.18)–(2.19), one gets that (recall (4.25), (4.26))

$$
\begin{aligned}
\mathrm{a}^{(-1)}(x,\xi) + \mathrm{b}^{(-1)}(x,\xi) = & (\partial_{u_{xx}}f_2)_{xx}\Lambda^{-1}(\xi)\left(\frac{1}{4} - \frac{\xi^2\Lambda^{-2}(\xi)}{8} + \frac{\xi^4\Lambda^{-4}(\xi)}{4}\right) \\
& - \frac{1}{2}(\partial_{u_x}f_2)_x\Lambda^{-1}(\xi) + (\partial_u f_2)\Lambda^{-1}(\xi).
\end{aligned}
$$

(4.28)

Notice also that (recall (2.5))

$$
\begin{aligned}
-\xi^2\Lambda^{-1}(\xi) = & \frac{-\xi^2}{\sqrt{|\xi|^2+m}} = -\sqrt{|\xi|^2+m} + \frac{m}{\sqrt{|\xi|^2+m}} = -\Lambda(\xi) + m\Lambda^{-1}(\xi), \\
\xi^2\Lambda^{-2}(\xi) = & 1 - m\Lambda^{-2}(\xi).
\end{aligned}
$$

By the discussion above we have that (4.27) can be written as

$$
(dX^+_{H^{(3)}})(\varepsilon\varphi)[V] = \frac{i}{2} Op^{BW}\Big(-\partial_{u_{xx}} f_2 \Lambda(\xi)\Big)[v + \overline{v}]
$$
$$
+ \frac{i}{2} Op^{BW}\Big(\big(m(\partial_{u_{xx}} f_2) + \frac{3}{8}(\partial_{u_{xx}} f_2)_{xx} - \frac{1}{2}(\partial_{u_x} f_2)_x + (\partial_u f_2)\big)\Lambda^{-1}(\xi)\Big)[v + \overline{v}],
$$

(4.29)

up to some remainder in $\mathbf{M}_1^{-2}$. In the latter equation we have used that pseudo-differential operators of order -2 are 2-smoothing multilinear operators by Lemma 2.8. Recalling (4.14), we have that the (4.29) implies (4.22).

Proof of Proposition 4.4 Notice that $H^{(3)} = H^{(3, \leq 1)} + H^{(3, \geq 2)}$. Hence

$$
(dX_{H^{(3, \geq 2)}})(\varepsilon\varphi)[\cdot] = (dX_{H^{(3)}})(\varepsilon\varphi)[\cdot] - (dX_{H^{(3, \leq 1)}})(\varepsilon\varphi)[\cdot].
$$

Since $X_{H^{(3, \leq 1)}}$ is finite rank and quadratic we have $(dX_{H^{(3, \leq 1)}})(\varepsilon\varphi) \in \mathbf{M}_1^{-2} \otimes \mathcal{M}_2(\mathbb{C})$. Then the structure of the linearized operator $(dX_{H^{(3, \geq 2)}})(\varphi)[\cdot]$ is essentially determined by Lemma 4.7. Then (4.19) follows by (4.21).

The bound (4.20) follows from Lemmata 4.6, 2.4 to estimate the smoothing remainders, and from Lemma 4.7 and Lemma 2.7 (recall also Def. 2.5) to estimate the unbounded pseudo-differential terms.

4.3.2 Para-Linearization of the Nonlinear Term

We consider the nonlinear term $\mathcal{Q}(\varepsilon\varphi)[V, V]$ appearing in (4.9). Recalling (4.10) we write

$$
\mathcal{Q}(\varepsilon\varphi)[V, V] := \widetilde{\mathcal{Q}}(\varepsilon\varphi)[V, V] + \mathcal{Q}^\star(\varepsilon\varphi)[V, V], \quad \text{where}
$$
$$
\widetilde{\mathcal{Q}}(\varepsilon\varphi)[V, V] := \frac{1}{2} X_{H^{(3, \geq 2)}}(V, V),
$$
$$
\mathcal{Q}^\star(\varepsilon\varphi)[V, V] := \int_0^1 (1 - t)\, d^2 X_{\mathcal{H} - H^{(3, \geq 2)}}(\varepsilon\varphi + tV)[V, V]\, dt.
$$

(4.30)

We note that $\mathcal{Q}^\star$ has finite rank and the following estimate holds

$$
\|\mathcal{Q}^\star(\varepsilon\varphi)[V, V]\|_{H^s} \lesssim_s \|\varepsilon\varphi\|_{L^\infty} \|V\|^2_{H^0} \qquad \forall\, V \in B_\varepsilon(H^s).
$$

(4.31)

The (4.31) follows recalling (3.15), (3.18), (3.19) estimate (4.4) and the smallness (4.5). The aim of this section is to show that term $\widetilde{\mathcal{Q}}$ in (4.30) has a para-differential structure.

We have the following.

Proposition 4.8 *Let $s > 2 + 1/2$, $V \in H^s$ and let $\mathcal{Q}(\varepsilon\varphi)[V, V]$ be the nonlinearity in (4.30) with $\varepsilon\varphi$ satisfying (4.4). We have that*

$$
\mathcal{Q}(\varepsilon\varphi)[V, V] = \frac{i}{2} E\, Op^{BW}\Big(\big[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\big] a(V; x)\Lambda(\xi) + \big[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\big] d(V; x)\Lambda^{-1}(\xi)\Big)V + R(V), \quad (4.32)
$$

where $a(V; x), d(V; x)$ are in (4.14). The function $R(V)$ has the form $(R^+(V), \overline{R^+(V)})^T$ and it satisfies a bound like (4.17). Moreover we have that $a \in \mathbf{SM}_1^0$ (see Def. 2.5) and

$$
|a(V)|_{\mathcal{N}_p^0} \lesssim \|V\|_{H^{p+s_0+2}}, \quad \forall\, p + s_0 + 2 \leq s, \quad p \in \mathbb{N},
$$
$$
|d(V)|_{\mathcal{N}_p^0} \lesssim \|V\|_{H^{p+s_0+4}}, \quad \forall\, p + s_0 + 4 \leq s, \quad p \in \mathbb{N},
$$

(4.33)

where $s_0 > 1/2$.

Proof By (4.30) we write

$$\mathcal{Q}(\varepsilon\varphi)[V, V] = \frac{1}{2} X_{H^{(3)}}(V, V) - \frac{1}{2} X_{H^{(3,\leq 1)}}(V, V) + \mathcal{Q}^\star(\varepsilon\varphi)[V, V], \qquad (4.34)$$

where $H^{(3)}$ is the Hamiltonian function in (2.12). The vector field $X_{H^{(3,\leq 1)}}$ is finite rank (recall Remark 3.1). Then the last two summands in the r.h.s. of (4.34) satisfy a bound like (4.17) using also the estimate (4.31). By applying the Bony-paralinearization formula (see [48, 53]) to the vector field $X_{H^{(3)}}$, using symbolic calculus (see Lemma 2.7) and reasoning as in Lemma 4.7, one gets

$$X_{H^{(3)}}(V, V) = \mathrm{i} E O p^{\mathrm{BW}}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} a(V; x)\Lambda(\xi) + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} d(V; x)\Lambda^{-1}(\xi)\right)$$

up to some smoothing remainder satisfying (4.17). For more details we refer the reader to Proposition 3.6 in [33]. Therefore formula (4.32) follows. The bound (4.33) follows using the explicit definition of the symbols $a(V; x), d(V; x)$ in (4.14) and by Lemma 2.6.

Proof of Proposition 4.3 It follows by by Propositions 4.4 and 4.8.

Remark 4.9 (Hamiltonian structure 1). Recalling (4.15), Remark 4.5 and Proposition 4.4, we notice that the operator

$$\mathrm{i} E O p^{\mathrm{BW}}\left(\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathtt{f}(\varepsilon\varphi, V; x)\right)\Lambda(\xi) + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathtt{g}(\varepsilon\varphi, V; x)\Lambda^{-1}(\xi)\right)_{|V\equiv 0}[\cdot] + \mathrm{i} E \mathfrak{Q}(\varepsilon\varphi)[\cdot]$$

is *Hamiltonian* according to Definition 2.14.

5 The Estimates on the Error Function: High Frequencies

Consider the remainder V in (4.8) which solves (4.9). We shall provide a priori bounds on the norm of V as long as (4.11) holds. This section concerns the study of the *high* frequencies of V. In particular we consider the equation

$$\varepsilon^\beta \partial_t \Pi_S^\perp \dot{V} = \Pi_S^\perp d X_{\mathcal{H}}(\varepsilon\varphi)[\varepsilon^\beta V] + \Pi_S^\perp \mathcal{Q}(\varepsilon\varphi)[\varepsilon^\beta V, \varepsilon^\beta V] + \Pi_S^\perp \mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi), \qquad (5.1)$$

where S is in (4.1). The main result of this section is the following.

Theorem 5.1 (Estimates of high modes) *There is $s_0 \gg 1$ such that for any $s \geq s_0$ there is a constant $C_s > 0$ such that if ε satisfies (4.5) the following holds. Consider $\varepsilon\varphi$ in (4.8) and let V be a solution of (4.9) defined for $t \in [0, T]$ for some $T > 0$. Then, if (4.11) and (4.4) hold true, one has*

$$\varepsilon^{2\beta} \|\Pi_S^\perp V\|_{H^s}^2 \leq (1 + \varepsilon C_s)\varepsilon^{2\beta} \|\Pi_S^\perp V(0)\|_{H^s}^2$$
$$+ C_s T \sup_{t\in[0,T]} \|\varepsilon\varphi\|_{H^s}^3 \sup_{t\in[0,T]} \|\varepsilon^\beta V\|_{H^s}^2$$
$$+ C_s T \sup_{t\in[0,T]} \|\varepsilon^\beta V\|_{H^s}^3 + C_s T \varepsilon^5 \sup_{t\in[0,T]} \|\varepsilon^\beta V\|_{H^s}, \qquad (5.2)$$

uniformly in $t \in [0, T]$.

The proof of the above result involves several arguments. We start by rewriting equation (5.1) in a more suitable way. Let us define

$$U^\perp := \varepsilon^\beta \Pi_S^\perp V, \qquad U^\perp = \begin{pmatrix} u^\perp \\ \overline{u^\perp} \end{pmatrix}, \qquad u^\perp := \varepsilon^\beta \Pi_S^\perp v. \qquad (5.3)$$

In the following lemma we provide some properties of the vector field in the r.h.s. of (5.1).

Lemma 5.2 (*i*) *The function* $U^\perp$ *in* (5.3) *satisfies the problem*

$$\partial_t U^\perp = \Pi_S^\perp \mathrm{i} E \, Op^{\mathrm{BW}}\left(\left(\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \mathtt{f}(\varepsilon\varphi, \varepsilon^\beta V; x)\right)\Lambda(\xi)\right) U^\perp$$
$$+ \Pi_S^\perp \mathrm{i} E \, Op^{\mathrm{BW}}\left(\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \mathtt{g}(\varepsilon\varphi, \varepsilon^\beta V; x)\Lambda^{-1}(\xi)\right) U^\perp \tag{5.4}$$
$$+ \Pi_S^\perp \mathrm{i} E \mathfrak{Q}(\varepsilon\varphi) U^\perp + \Pi_S^\perp \widetilde{R}(\varepsilon\varphi, \varepsilon^\beta V) + \Pi_S^\perp \mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi),$$

where the symbols $\mathtt{f}$, $\mathtt{g}$ *and the remainder* $\mathfrak{Q}$ *are given by Proposition* 4.3. *Moreover the remainder*

$$\widetilde{R}(\varepsilon\varphi; \varepsilon^\beta V) = \left(\widetilde{R}^+(\varepsilon\varphi; \varepsilon^\beta V), \overline{\widetilde{R}^+(\varepsilon\varphi; \varepsilon^\beta V)}\right)^T,$$

and there is $s_0 > 1$ *such that, for* $s \geq s_0$, *it satisfies the bound*

$$\|\widetilde{R}(\varepsilon\varphi, \varepsilon^\beta V)\|_{H^s} \lesssim_s \varepsilon^{2\beta} \|V\|_{H^s}^2. \tag{5.5}$$

(*ii*) *The vector field*

$$\Pi_S^\perp \mathrm{i} E \, Op^{\mathrm{BW}}\left(\left(\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \mathtt{f}(\varepsilon\varphi, \varepsilon^\beta V; x)\right)\Lambda(\xi)\right)_{|V \equiv 0} \Pi_S^\perp + \Pi_S^\perp \mathrm{i} E \mathfrak{Q}(\varepsilon\varphi) \Pi_S^\perp$$
$$+ \Pi_S^\perp \mathrm{i} E \, Op^{\mathrm{BW}}\left(\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \mathtt{g}(\varepsilon\varphi, \varepsilon^\beta V; x, \xi)\Lambda^{-1}(\xi)\right)_{|V \equiv 0}$$
$$= \Pi_S^\perp (dX_{\mathcal{H}})(\varepsilon\varphi) \Pi_S^\perp$$

is Hamiltonian *according to Def.* 2.14.

Proof By Proposition 4.3 and (5.3) (see also (4.9), (4.13) and (4.16)) we have that equation (5.1) can be written as

$$\partial_t U^\perp = \Pi_S^\perp \mathrm{i} E \, Op^{\mathrm{BW}}\left(\left(\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \mathtt{f}(\varepsilon\varphi, \varepsilon^\beta V; x)\right)\Lambda(\xi)\right)[\varepsilon^\beta V]$$
$$+ \Pi_S^\perp \mathrm{i} E \, Op^{\mathrm{BW}}\left(\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \mathtt{g}(\varepsilon\varphi, \varepsilon^\beta V; x)\Lambda^{-1}(\xi)\right)[\varepsilon^\beta V] \tag{5.6}$$
$$+ \Pi_S^\perp \mathrm{i} E \mathfrak{Q}(\varepsilon\varphi)[\varepsilon^\beta V] + \Pi_S^\perp R(\varepsilon^\beta V) + \Pi_S^\perp \mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi).$$

Then (5.4) comes from the fact that $\varepsilon^\beta V = \Pi_S \varepsilon^\beta V + U^\perp$ and that the operators in the right hand side of (5.6) composed with Π_S are bounded and satisfy (5.5). Recalling Propositions 4.4, 4.3 we notice that

$$\Pi_S^\perp (dX_{\mathcal{H}})(\varepsilon\varphi)[\varepsilon^\beta V] = \Pi_S^\perp dX_{\mathcal{H}}(\varepsilon\varphi)[\Pi_S^\perp \varepsilon^\beta V], \tag{5.7}$$

where we used that, since $\varepsilon\varphi$ is Fourier supported only on the tangential set S in (4.1) then

$$\Pi_S^\perp dX_{\mathcal{H}_{res}^{(4,0)}}(\varepsilon\varphi) = \Pi_S^\perp dX_{\mathcal{H}^{(n, \geq 2)}}(\varepsilon\varphi)[\Pi_S V] = 0.$$

This proves item (*ii*).

5.1 Block-Diagonalization and Basic Energy Estimates

In this section we construct a (linear) change of coordinates which block-diagonalize the system (5.4) up to smoothing remainders. We first introduce a map, given in terms of a suitable para-product, in such a way one can diagonalize the matrix of symbols

$$E\left(\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \mathtt{f}(\varepsilon\varphi, \varepsilon^\beta V; x)\right), \tag{5.8}$$

where $\mathtt{f}$ is the symbol in (4.15). This is the content of Sect. 5.1.1. In Sect. 5.1.2 we provide a correction which give us a map which is symplectic on the restriction of the space $H^s(\mathbb{T};\mathbb{C}) \times H^s(\mathbb{T};\mathbb{C})$ to functions supported only on S^c. Finally in Sect. 5.1.3 we diagonalize at the highest order the system (5.4).

5.1.1 Construction of an Approximate Diagonalizating Matrix

Let us define the matrices[2]

$$G := G(\varepsilon\varphi, \varepsilon^\beta V) := \begin{pmatrix} \mathtt{g}^+ & \mathtt{g}^- \\ \mathtt{g}^- & \mathtt{g}^+ \end{pmatrix}, \qquad G^{-1} := G^{-1}(\varepsilon\varphi, \varepsilon^\beta V) := \begin{pmatrix} \mathtt{g}^+ & -\mathtt{g}^- \\ -\mathtt{g}^- & \mathtt{g}^+ \end{pmatrix},$$

(5.9)

where the symbols $\mathtt{g}^+$, $\mathtt{g}^-$ are defined as

$$\mathtt{g}^+ := \mathtt{g}^+(\varepsilon\varphi, \varepsilon^\beta V) := \frac{1 + \mathtt{f} + \lambda}{\sqrt{2\lambda(1 + \mathtt{f} + \lambda)}},$$

$$\mathtt{g}^- := \mathtt{g}^-(\varepsilon\varphi, \varepsilon^\beta V) := \frac{-\mathtt{f}}{\sqrt{2\lambda(1 + \mathtt{f} + \lambda)}},$$

(5.10)

and

$$\lambda := \lambda(\varepsilon\varphi, \varepsilon^\beta V) := \sqrt{(1 + \mathtt{f})^2 - (\mathtt{f})^2} = \sqrt{1 + 2\mathtt{f}}.$$

(5.11)

The first result of this section is the following.

Lemma 5.3 *Recall the symbol $\mathtt{f}$ in (4.15) (see also (4.14)), and the functions $\varepsilon\varphi$, V in (4.8). Assume the (4.4) and (4.11). Then the following holds.*
(i) the symbol $\lambda - 1$ in (5.11) is independent of $\xi \in \mathbb{R}$ and belongs to (see Def. 2.5) $\mathbf{S\Sigma}_1^m[r, 3]$ with estimates uniform in $t \in [0, T]$.
(ii) the symbols $\mathtt{g}^+ - 1$, $\mathtt{g}^-$ in (5.10) are independent of $\xi \in \mathbb{R}$ and belong to (see Def. 2.5) $\mathbf{S\Sigma}_1^m[r, 3]$ with estimates uniform in $t \in [0, T]$.
(iii) One has that

$$\partial_t \mathtt{g}^\sigma = \widetilde{\mathfrak{g}}_1^\sigma(\varepsilon\varphi) + \widetilde{\mathfrak{g}}_2^\sigma(\varepsilon\varphi, \varepsilon\varphi) + \widetilde{\mathfrak{g}}_{\geq 3}^\sigma(\varepsilon\varphi, \varepsilon^\beta V), \quad \sigma \in \{\pm\},$$

(5.12)

where $\widetilde{\mathfrak{g}}_j^\sigma \in \mathbf{SM}_j^0$, $j = 1, 2$, and $\widetilde{\mathfrak{g}}_{\geq 3}^\sigma$, $\sigma \in \{\pm\}$, are in $\mathcal{N}_p^0$ with

$$|\widetilde{\mathfrak{g}}_{\geq 3}^\sigma|_{\mathcal{N}_p^0} \lesssim_s \|\varepsilon\varphi\|_{H^{p+\mu}}^3 + \|\varepsilon^\beta V\|_{H^{p+\mu}} + \varepsilon^5,$$

(5.13)

$\sigma \in \{\pm\}$, for any $p + \mu \leq s$, $p \in \mathbb{N}$, for some $\mu > 1/2$, with estimates uniform in $t \in [0, T]$.
(iv) Recall (5.8), (5.9). One has that

$$G^{-1}(\varepsilon\varphi, \varepsilon^\beta V)E\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\mathtt{f}(\varepsilon\varphi, \varepsilon^\beta V; x)\right)G(\varepsilon\varphi, \varepsilon^\beta V) = E\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\lambda(\varepsilon\varphi, \varepsilon^\beta V).$$

(5.14)

Proof Item (i). By Taylor expanding the function λ in (5.11) we get

$$\lambda = 1 + \mathtt{f} - \frac{\mathtt{f}^2}{2} + \frac{3\mathtt{f}^3}{2}\int_0^1 (1 + 2t\mathtt{f})^{-\frac{5}{2}}(1 - t)^2 dt,$$

(5.15)

---

[2] To simplify the notation, from now on we will omit the dependence on (x, ξ) of the symbols.

from which we deduce, using also (4.15),

$$\lambda = 1 + a(\varepsilon\varphi) - \frac{1}{2}[a(\varepsilon\varphi)]^2 + \frac{1}{2}a(\varepsilon^\beta V) - \frac{1}{2}a(\varepsilon\varphi)a(\varepsilon^\beta V) - \frac{1}{8}[a(\varepsilon^\beta V)]^2$$

$$+ \frac{3\mathtt{f}^3}{2}\int_0^1 (1 + 2t\mathtt{f})^{-\frac{5}{2}}(1-t)^2 dt.$$

(5.16)

Using (4.33), (4.4), the fact that the symbol $a(\varepsilon\varphi)$ in (4.14) is in $\mathbf{SM}_1^0$, and Lemma 2.8 we deduce that $\lambda - 1$ belongs to the class $\mathbf{S\Sigma}_1^m[r, 3]$ of Definition 2.5.

Item (ii). It follows using the explicit formulæ (5.10) and reasoning as done for Item (i).

Item (iii). Recalling $\mathcal{H}$ in (3.15) and (4.6), we notice that the function $\varepsilon\varphi$ in (4.8) solves the equation

$$\partial_t \varepsilon\varphi = X_{\mathcal{H}}(\varepsilon\varphi) - \mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi).$$

By the properties of $\mathcal{H}$ given by Proposition 3.9 we can write

$$\partial_t \varepsilon\varphi = \mathrm{i}EOp^W([\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}]\Lambda(\xi))\varepsilon\varphi + \mathrm{i}EM(\varepsilon\varphi)\varepsilon\varphi - \mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)$$
$$M(\varepsilon\varphi) = M_1(\varepsilon\varphi) + M_2(\varepsilon\varphi) + M_{\geq 3}(\varepsilon\varphi)$$

(5.17)

for some real and self-adjoint matrix of operators $M(\varphi)$ where $M_i \in \mathbf{M}_i^1 \otimes \mathcal{M}_2(\mathbb{C}), i = 1, 2$ and $M_{\geq 3} \in \mathbf{NH}_3^1[r] \otimes \mathcal{M}_2(\mathbb{C})$. Recall also that V solves the equation (4.9) that, by Proposition 4.3, can be written as

$$\varepsilon^\beta \partial_t V = \mathcal{Y}(\varepsilon\varphi, \varepsilon^\beta V) + \mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)$$

(5.18)

where $\mathcal{Y}(\varepsilon\varphi, V)$ is in (4.16). We now compute the time derivative of the symbol g^+ in (5.10). The case of g^- is similar. By item (ii) (recall Def. 2.5) we have that g^+ admits the expansion

$$\mathsf{g}^+ = 1 + \mathsf{g}_1^+(\varepsilon\varphi) + \mathsf{g}_2^+(\varepsilon\varphi, \varepsilon\varphi) + \mathsf{g}_{\geq 3}^+(\varepsilon\varphi, \varepsilon^\beta V), \quad \mathsf{g}_j^+ \in \mathbf{SM}_j^0, \quad j = 1, 2,$$

(5.19)

and $\mathsf{g}_{\geq 3}^+, \sigma \in \{\pm\}$, are in $\mathcal{N}_p^0$ with

$$|\mathsf{g}_{\geq 3}^+|_{\mathcal{N}_p^0} \lesssim_s \|\varepsilon\varphi\|_{H^{p+\mu}}^3 + \|\varepsilon^\beta V\|_{H^{p+\mu}},$$

(5.20)

for any $p + \mu \leq s$, $p \in \mathbb{N}$ and some $\mu > 1/2$, with estimates uniform in $t \in [0, T]$. Moreover one also has (recall estimates (2.42))

$$|d_{\varepsilon\varphi}\mathsf{g}_{\geq 3}^+(\varepsilon\varphi, \varepsilon^\beta V)[h]|_{\mathcal{N}_p^0} \lesssim_s \|\varepsilon\varphi\|_{H^{p+\mu}}^2 \|h\|_{H^{p+\mu}} + \|\varepsilon^\beta V\|_{H^{p+\mu}}\|h\|_{H^{p+\mu}},$$
$$|d_{\varepsilon^\beta V}\mathsf{g}_{\geq 3}^+(\varepsilon\varphi, \varepsilon^\beta V)[h]|_{\mathcal{N}_p^0} \lesssim_s \|h\|_{H^{p+\mu}},$$

(5.21)

for any $p + \mu \leq s$, $p \in \mathbb{N}$, for any $h \in H^s$.

Consider the symbol g_1^+ in (5.19) which is linear in $\varepsilon\varphi$. Therefore we have

$$\partial_t \mathsf{g}_1^+(\varepsilon\varphi) \overset{(5.17)}{=} \mathsf{g}_1^+(\mathrm{i}EOp^W([\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}]\Lambda(\xi))\varepsilon\varphi) + \mathsf{g}_1^+(\mathrm{i}EM_1(\varepsilon\varphi)\varepsilon\varphi) + \mathsf{g}_1^+(\mathrm{i}EM_2(\varepsilon\varphi)\varepsilon\varphi)$$
$$+ \mathsf{g}_1^+(\mathrm{i}EM_{\geq 3}(\varepsilon\varphi)\varepsilon\varphi) + \mathsf{g}_1^+(-\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)).$$

Using the properties of the matrices of operators $M_i(\varepsilon\varphi), i = 1, 2, M_{\geq 3}(\varepsilon\varphi)$, the composition properties in Lemma 2.9 we deduce

$$\partial_t \mathsf{g}_1^+(\varepsilon\varphi) = \mathsf{a}_1(\varepsilon\varphi) + \mathsf{a}_2(\varepsilon\varphi) + \mathsf{a}_{\geq 3}(\varepsilon\varphi) + \mathsf{g}_1^+(-\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)),$$

for some symbols $\mathsf{a}_j \in \mathbf{SM}_j^0$, $j = 1, 2$, and $\mathsf{a}_{\geq 3} \in \mathbf{SNH}_3^m[r]$. The term $\mathsf{g}_1^+(-\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi))$ can be estimated (as a symbol of order 0) using Lemmas 2.6 and 4.1 to estimate the residual. One concludes that $\mathsf{g}_1^+(-\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi))$ can be absorbed in the term $\widetilde{\mathsf{g}}_{\geq 3}^+$ in (5.12) satisfying

(5.13). Concerning the term g_2^+ in (5.19) (which is quadratic in $\varepsilon\varphi$), one can reason as done for g_1^+. Consider now $g_{\geq 3}^+(\varepsilon\varphi, \varepsilon^\beta V)$ in (5.19). One has

$$\partial_t g_{\geq 3}^+(\varepsilon\varphi, \varepsilon^\beta V) = (d_{\varepsilon\varphi} g_{\geq 3}^+(\varepsilon\varphi, \varepsilon^\beta V))[\varepsilon\partial_t\varphi] + (d_{\varepsilon^\beta V} g_{\geq 3}^+(\varepsilon\varphi, \varepsilon^\beta V))[\varepsilon^\beta \partial_t V].$$

The estimate (5.13) follows using (5.21), (5.17), (5.18), and (4.7), Proposition 3.9 to estimate the norms of $\varepsilon\partial_t\varphi$, the (4.18) and Lemma 4.1 to estimate $\varepsilon^\beta \partial_t V$.

Item (iv). The (5.14) follows by and explicit computation using (5.9), (5.10) and (5.8).

The idea now is that a possible map which block-diagonalizes the system (5.1) at highest order would be $Op(G^{-1})$ with G^{-1} as in (5.9). This is a consequence of Item (iv) in Lemma 5.3. However the linear map $Z \mapsto Op(G_\chi^{-1}(\varepsilon\varphi, \varepsilon^\beta V))Z$ is not symplectic. In the following we will show how to construct a symplectic correction of such map.

Lemma 5.4 *Under the hypotheses of Lemma 5.3 there exists a real valued symbol* $\mathrm{m} := \mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)$ *such that the following holds:*
(i) *one has*

$$\begin{aligned} g^+(\varepsilon\varphi, \varepsilon^\beta V) &= \cosh(|\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)|)\,, \\ g^-(\varepsilon\varphi, \varepsilon^\beta V) &= \frac{\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)}{|\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)|} \sinh(|\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)|)\,, \end{aligned} \tag{5.22}$$

where g^σ, $\sigma \in \{\pm\}$, *are given in* (5.10).
(ii) *the symbol* m *is in (see Def. 2.5)* $\mathbf{S}\Sigma_1^m[r, 3]$.
(iii) *One has that*

$$\partial_t \mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V) = \widetilde{\mathrm{m}}_1(\varepsilon\varphi) + \widetilde{\mathrm{m}}_2(\varepsilon\varphi, \varepsilon\varphi) + \widetilde{\mathrm{m}}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V)\,, \tag{5.23}$$

where $\widetilde{\mathrm{m}}_j \in \mathbf{SM}_j^0$, $j = 1, 2$, *and* $\widetilde{\mathrm{m}}_{\geq 3}$ *is in* $\mathcal{N}_p^0$ *with*

$$|\widetilde{\mathrm{m}}_{\geq 3}|_{\mathcal{N}_p^0} \lesssim_s \|\varepsilon\varphi\|_{H^{p+\mu}}^3 + \|\varepsilon^\beta V\|_{H^{p+\mu}} + \varepsilon^5\,, \tag{5.24}$$

for any $p + \mu \leq s$, $p \in \mathbb{N}$, $\mu > 1/2$, *with estimates uniform in* $t \in [0, T]$.

Proof We look for a solution of the equations (5.22). First of all notice that, using the estimates on the symbol f (see (4.15) and (4.33)), the estimate on $\lambda \in \mathbf{S}\Sigma_1^m[r, 3]$ (see item (i) of Lemma 5.3), estimates (4.4) and (4.11) on the functions $\varepsilon\varphi$, V and the smallness assumption on ε, we get that

$$(g^+(\varepsilon\varphi, \varepsilon^\beta V))^2 - 1 \stackrel{(5.10),(5.11)}{=} \frac{f^2}{2\lambda(1 + f + \lambda)} \geq 0.$$

Therefore we set

$$\begin{aligned} |\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)| &:= \mathrm{arccosh}(g^+(\varepsilon\varphi, \varepsilon^\beta V)) \\ &= \ln\left(g^+(\varepsilon\varphi, \varepsilon^\beta V) + \sqrt{(g^+(\varepsilon\varphi, \varepsilon^\beta V))^2 - 1}\right). \end{aligned} \tag{5.25}$$

Consider now the second equation (5.22). We first observe that

$$\frac{\sinh(|\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)|)}{|\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)|} = 1 + \sum_{k\geq 0} \frac{(\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V))^{2k}}{(2k + 1)!} \geq 1.$$

Hence we set

$$\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V) := g^-(\varepsilon\varphi, \varepsilon^\beta V)\frac{|\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)|}{\sinh(|\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)|)}\,. \tag{5.26}$$

Using the properties of the symbols g^{+}, g^{-} (given by Lemma 5.3), and by Taylor expanding formulæ (5.25), (5.26), reasoning as in item (i) of Lemma 5.3 and using the (5.19)–(5.20), one obtains that (see Def. 2.5) $\mathbf{S}\Sigma_1^m[r, 3]$. This proves item (ii).

Item (iii) follows by an explicit computation using (5.26), estimates (5.21) and reasoning as done in the proof if item (iii) of Lemma 5.3.

We need the following technical Lemma, which is essentially a consequence of Proposition 3.6 in [35].

Lemma 5.5 *Under the hypotheses of Lemma 5.3 the following holds. For any $k \geq 1$ one has that the operator*

$$Q^{(k)}(\varepsilon\varphi, \varepsilon^\beta V) := \left(Op^{\mathrm{BW}}(\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)) \right)^k - Op^{\mathrm{BW}}\left((\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V))^k \right)$$

belongs to the class $\Sigma_1^{-3}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ (see Def. 2.3). More precisely there is $\mathsf{C} = \mathsf{C}(s)$ independent of $k \geq 1$, such that

$$\|Q^{(k)}(\varepsilon\varphi, \varepsilon^\beta V)Z\|_{H^{s+\rho}} \leq \mathsf{C}^k (\|\varepsilon\varphi\|_{H^s}^k + \|\varepsilon^\beta V\|_{H^s}^k)\|Z\|_{H^s},$$

$$\|d_{\varepsilon\varphi}\left(Q^{(k)}(\varepsilon\varphi, \varepsilon^\beta V)\right)(Z)h\|_{H^{s+\rho}} \leq \mathsf{C}^k (\|\varphi\|_{H^s}^{k-1} + \|V\|_{H^s}^k)\|Z\|_{H^s}\|h\|_{H^s}, \qquad (5.27)$$

$$\|d_{\varepsilon^\beta V}\left(Q^{(k)}(\varepsilon\varphi, \varepsilon^\beta V)\right)(Z)h\|_{H^{s+\rho}} \leq \mathsf{C}^k \|V\|_{H^s}^{k-1}\|Z\|_{H^s}\|h\|_{H^s}.$$

any $Z = \left[\begin{smallmatrix} z \\ \bar{z} \end{smallmatrix} \right] \in H^s$.

Proof The case $k = 1$ is trivial. Let us consider $k = 2$. This case is essentially a consequence of (2.28) in Remark 2.2. Indeed, notice that

$$Q^{(2)}(\varepsilon\varphi, \varepsilon^\beta V)h = \left(Op^{\mathrm{BW}}(\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)) \circ Op^{\mathrm{BW}}(\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)) - Op^{\mathrm{BW}}\left((\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V))^2\right) \right)h$$

$$= \sum_{\xi \in \mathbb{Z}} e^{i\xi x} (\widehat{Q^{(2)}h})(\xi),$$

$$(\widehat{Q^{(2)}h})(\xi) = (2\pi)^{-\frac{3}{2}} \sum_{\eta, \theta \in \mathbb{Z}} (r_1 - r_2)(\xi, \theta, \eta)\widehat{\mathrm{m}}(\xi - \theta)\widehat{\mathrm{m}}(\theta - \eta)\widehat{h}(\eta)$$

$$(5.28)$$

where

$$r_1(\xi, \theta, \eta) := \chi\left(\frac{|\xi - \theta|}{\langle\xi + \theta\rangle}\right) \chi\left(\frac{|\theta - \eta|}{\langle\theta + \eta\rangle}\right), \qquad r_2(\xi, \eta) := \chi\left(\frac{|\xi - \eta|}{\langle\xi + \eta\rangle}\right) \qquad (5.29)$$

and $\widehat{\mathrm{m}}(\xi)$ denotes the Fourier transform in x of the function $(\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V))(x)$ evaluated at $\xi \in \mathbb{Z}$. We remark that the remainder $Q^{(2)}$ is bilinear in the symbol $\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)$. Recall that, by item (ii) of Lemma 5.4, the symbol m belongs to $\mathbf{S}\Sigma_1^m[r, 3]$. Therefore, using (5.28), we can expand $Q^{(2)}$ as the sum of two remainders, the first depending only on $\varepsilon\varphi$ and the second depending at least linearly in the variable $\varepsilon^\beta V$. The estimates (5.27) can be deduced by following almost word by word the proof of Proposition 2.5 in [36], and using formulæ (5.28) and (5.29). In order to get the result for any $k \geq 3$, one reasons by induction following the proof of Proposition 3.6 in [35]. The Taylor expansions in $\varepsilon\varphi$ and $\varepsilon^\beta V$ follow again recalling that the remainders $Q^{(k)}$ are multilinear in the symbol $\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V) \in \mathbf{S}\Sigma_1^m[r, 3]$.

Consider the matrix of symbols

$$\mathbf{M} := \mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V) = \begin{pmatrix} 0 & \mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V) \\ \mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V) & 0 \end{pmatrix}, \qquad (5.30)$$

where $\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V)$ is given by Lemma 5.4. We shall study, for $\tau \in [0, 1]$, the properties of the flow map of the following problem:

$$\begin{cases} \partial_\tau \Phi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot] = Op^{\mathrm{BW}}(\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))\Phi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot], \\ \Phi_{\mathrm{m}}^0(\varepsilon\varphi, \varepsilon^\beta V)[\cdot] = \mathbb{1}[\cdot], \end{cases} \tag{5.31}$$

where $\mathbf{M}_\chi$ is the matrix whose entries are given in terms of the symbol $\mathrm{m}_\chi(\varepsilon\varphi, \varepsilon^\beta V)$ defined as in (4.15). More precisely we prove the following.

Lemma 5.6 *Under the hypotheses of Lemma 5.3 the following holds. The flow map of (5.31) is well-posed for $\tau \in [0, 1]$ and have the form*

$$\Phi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot] - \mathbb{1} = Q(\tau; \varepsilon\varphi, \varepsilon^\beta V) \quad \tau \in [0, 1], \tag{5.32}$$

for some real operator $Q \in \Sigma_1^0[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ (see Def. 2.3), with estimates uniform in $\tau \in [0, 1]$. The same holds for the inverse map $(\Phi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V))^{-1}[\cdot]$. Moreover for the time one flow map we have the following expansion

$$\Phi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot]_{|\tau=1} := Op^{\mathrm{BW}}(G^{-1}(\varepsilon\varphi, \varepsilon^\beta V))[\cdot] + R(\varepsilon\varphi, \varepsilon^\beta V)[\cdot], \tag{5.33}$$

$$(\Phi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V))^{-1}[\cdot]_{|\tau=1} := Op^{\mathrm{BW}}(G(\varepsilon\varphi, \varepsilon^\beta V))[\cdot] + R'(\varepsilon\varphi, \varepsilon^\beta V)[\cdot], \tag{5.34}$$

where G, G^{-1} are in (5.9), the operators R, R' are real and belong to $\Sigma_1^{-3}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ (see Def. 2.3). Finally

$$(\partial_t \Phi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot])_{|\tau=1} := Op^{\mathrm{BW}}(\widetilde{G}(\varepsilon\varphi, \varepsilon^\beta V))[\cdot] + \widetilde{R}_{\leq 2}(\varepsilon\varphi)[\cdot] + \widetilde{R}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V)[\cdot], \tag{5.35}$$

where $\widetilde{G}(\varepsilon\varphi, \varepsilon^\beta V)$ is a 2×2 matrix of real valued symbols, independent of $\xi \in \mathbb{R}$, of the form

$$\widetilde{G}(\varepsilon\varphi, \varepsilon^\beta V) := \widetilde{G}_1(\varepsilon\varphi) + \widetilde{G}_2(\tau; \varepsilon\varphi, \varepsilon\varphi) + \widetilde{G}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V), \tag{5.36}$$

where $\widetilde{G}_j \in \mathbf{SM}_j^0 \otimes \mathcal{M}_2(\mathbb{C})$, $j = 1, 2$, and where $\widetilde{G}_{\geq 3}(\tau; \varepsilon\varphi, \varepsilon^\beta V)$ is in $\mathcal{N}_p^0$ with

$$|\widetilde{G}_{\geq 3}|_{\mathcal{N}_p^0} \lesssim_s \|\varepsilon\varphi\|_{H^{p+\mu}}^3 + \|\varepsilon^\beta V\|_{H^{p+\mu}} + \varepsilon^5, \tag{5.37}$$

for any $p + \mu \leq s$, $p \in \mathbb{N}$, for some $\mu > 1/2$, with estimates uniform in $\tau \in [0, 1]$. Moreover the operator $\widetilde{R}_{\leq 2}$ has the form

$$\widetilde{R}_{\leq 2}(\tau; \varepsilon\varphi) := \widetilde{R}_1(\tau; \varepsilon\varphi) + \widetilde{R}_2(\tau; \varepsilon\varphi), \quad \widetilde{R}_j, \in \mathbf{M}_j^{-3} \otimes \mathcal{M}_2(\mathbb{C}), \quad j = 1, 2,$$

and $\widetilde{R}_{\geq 3}$ satisfies

$$\|\widetilde{R}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V)Z\|_{H^{s+3}} \lesssim_s \|Z\|_{H^s}(\|\varepsilon\varphi\|_{H^s}^3 + \|\varepsilon^\beta V\|_{H^s} + \varepsilon^5). \tag{5.38}$$

Proof By Lemma 2.2 (see (2.24)) and the bounds for the semi-norm of the symbol $\mathrm{m}(\varepsilon\varphi, \varepsilon^\beta V) \in \mathbf{S}\Sigma_1^0[r, 3]$, we have that $Op^{\mathrm{BW}}(\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))$ is a bounded operator on $\mathbf{H}^s$. Then by standard theory of Banach space ODE we have that the flow Φ_{m}^τ is well-defined. In particular we have

$$\Phi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot] = \exp\left(\tau Op^{\mathrm{BW}}(\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))\right) = \sum_{k \geq 0} \frac{\tau^k}{k!}(Op^{\mathrm{BW}}(\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V)))^k.$$

Hence by Lemma 5.5 one has

$$\Phi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot] = Op^{\mathrm{BW}}\left(\exp(\tau\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))\right) + \sum_{k \geq 0} \frac{\tau^k}{k!} Q^{(k)}(\varepsilon\varphi, \varepsilon^\beta V).$$

Reasoning as in Corollary 3.1 in [35], one deduces that

$$\Phi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot] = Op^{BW}\Big(\exp(\tau \mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V)) \Big) + R^\tau(\varepsilon\varphi, \varepsilon^\beta V) \tag{5.39}$$

where $R^\tau \in \Sigma_1^{-3}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$, with estimates uniform in $\tau \in [0, 1]$. The latter assertion follows by estimates (5.27) and using the smallness assumption on ε to have the convergence of the series. Furthermore, by (5.30), we have

$$
\exp(\tau \mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V)) = \sum_{n=0}^{\infty} \frac{\tau^{2n}}{(2n)!} \begin{pmatrix} |m|^{2n} & 0 \\ 0 & |m|^{2n} \end{pmatrix} + \sum_{n=0}^{\infty} \frac{\tau^{2n+1}}{(2n+1)!} \begin{pmatrix} 0 & |m|^{2n}m \\ |m|^{2n}m & 0 \end{pmatrix}
$$
$$
= \begin{pmatrix} \cosh(\tau|m|) & \frac{m}{|m|} \sinh(\tau|m|) \\ \frac{m}{|m|} \sinh(\tau|m|) & \cosh(\tau|m|) \end{pmatrix}.
$$
$$\tag{5.40}$$

We claim that there exists two real valued symbols $K^\sigma(\tau; \varepsilon\varphi, \varepsilon^\beta V)$, $\sigma \in \{\pm\}$, independent of $\xi \in \mathbb{R}$, such that

$$\exp(\tau \mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V)) = \begin{pmatrix} K^+(\tau; \varepsilon\varphi, \varepsilon^\beta V) & K^-(\tau; \varepsilon\varphi, \varepsilon^\beta V) \\ K^-(\tau; \varepsilon\varphi, \varepsilon^\beta V) & K^+(\tau; \varepsilon\varphi, \varepsilon^\beta V) \end{pmatrix} \in \mathbf{S}\Sigma_1^0[r, 3] \otimes \mathcal{M}_2(\mathbb{C}),$$
$$\tag{5.41}$$

with estimates uniform in $\tau \in [0, 1]$. Indeed $\cosh(\tau|m|)$ and $(m/|m|)\sinh(\tau|m|)$ are analytic functions of the symbol $m(\varepsilon\varphi, \varepsilon^\beta V) \in \mathbf{S}\Sigma_1^0[r, 3]$ (see Lemma 5.4 and Def. 2.5). Therefore the multilinear expansions in $\varepsilon\varphi$ and estimates (2.42) follow by explicit computations. This implies the claim (5.41). By the discussion above we have that the (5.39) reads

$$\Phi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot] = Op^{BW}\begin{pmatrix} K^+(\tau; \varepsilon\varphi, \varepsilon^\beta V) & K^-(\tau; \varepsilon\varphi, \varepsilon^\beta V) \\ K^-(\tau; \varepsilon\varphi, \varepsilon^\beta V) & K^+(\tau; \varepsilon\varphi, \varepsilon^\beta V) \end{pmatrix} + R^\tau(\varepsilon\varphi, \varepsilon^\beta V). \tag{5.42}$$

This implies (5.32) recalling Lemma 2.7-(iii) and the fact that $R^\tau \in \Sigma_1^{-3}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$.

Using Lemma 5.4 (see (5.22)) and recalling also (5.9)–(5.10) we deduce that formulæ (5.39), (5.40) and (5.42) imply the expansion (5.33). The expansion (5.34) follows similarly. Let us check (5.35). By differentiating in t equation (5.31) we get

$$\partial_\tau(\partial_t \Phi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V))[\cdot] = Op^{BW}(\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))(\partial_t \Phi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V))[\cdot]$$
$$+ Op^{BW}(\partial_t \mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))\Phi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot],$$

with $(\partial_t \Phi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V))[\cdot]|_{\tau=0} = 0$. By Duhamel formulation we deduce that

$$(\partial_t \Phi_m^\tau(\varepsilon\varphi, V))[\cdot] = \Phi_m^\tau(\varepsilon\varphi, V) \int_0^\tau (\Phi_m^\sigma(\varepsilon\varphi, V))^{-1} Op^{BW}(\partial_t \mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))\Phi_m^\sigma(\varepsilon\varphi, V)[\cdot] d\sigma.$$

The term $Op^W(\partial_t \mathbf{M}_\chi(\varepsilon\varphi, \varepsilon^\beta V))$ in the integral, is completely under control for any $t \in [0, T]$ thanks to item (iii) of Lemma 5.4 and recalling (5.30). Therefore, the expansion (5.35), together with (5.36)–(5.37) and estimate (5.38) follow using the expansions (5.42) for the flow $\Phi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V)$, Lemma 2.2, Lemma 2.10, Lemma 2.8 and using the expansions (5.23) and of m, (5.42) and (5.24).

Remark 5.7 (Hamiltonian structure 2). Notice that the operator $Op^{BW}(\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))_{|V\equiv 0}[\cdot]$ is Hamiltonian according to Definition 2.14. Therefore the flow map $\Phi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V)_{|V\equiv 0}[\cdot]$ is *symplectic* according to Definition 2.15.

Lemma 5.8 (Equivalence of the norms) *Consider the function $\varepsilon\varphi$ in (4.8) and let V be a solution of (4.9). For any $s \geq s_0$ for some $s_0 > 1$ there is a constant $C_s > 1$ such that if (4.11) and (4.4) hold true with $\varepsilon > 0$ satisfying (4.5) the following holds. One has that*

$$\sup_{\tau \in [0,1]} \|(\Phi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V))^{\pm 1}Z\|_{H^s} \leq \|Z\|_{H^s}(1 + C_s\|\varepsilon\varphi\|_{H^s} + C_s\|\varepsilon^{\beta}V\|_{H^s}), \quad (5.43)$$

for any $Z \in H^S$, where $\Phi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)$ is the the flow map given in Lemma 5.6. Moreover, defining the function

$$W := \left[\begin{smallmatrix} w \\ \overline{w} \end{smallmatrix}\right] := \left(\Phi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)\right)_{|\tau=1}[\varepsilon^{\beta}V], \quad (5.44)$$

one has the equivalence

$$(1 + \varepsilon C_s)^{-1}\|\varepsilon^{\beta}V\|_{H^s} \leq \|W\|_{H^s} \leq (1 + \varepsilon C_s)\|\varepsilon^{\beta}V\|_{H^s}. \quad (5.45)$$

Proof Consider the map in (5.33) and recall G^{-1} in (5.9)–(5.10). Then, using (5.20) and (2.43) to estimate the action of the pseudo-differential operator $Op^{\mathrm{BW}}\big(G^{-1}(\varepsilon\varphi, \varepsilon^{\beta}V)\big)$, using that $\mathrm{R} \in \Sigma_1^{-3}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ (see Lemma 5.6) and the smallness assumptions on $\varepsilon\varphi$, V, one gets the bound (5.43) for the map $\Phi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)$. The bound (5.43) on the inverse follows similarly using the (5.34). The equivalence (5.45) follows by (5.43) using also (4.4), (4.11) and the smallness (4.5). $\qquad\blacksquare$

5.1.2 The Symplectic Correction

The map $\Phi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)[\cdot]$ introduced in Lemma 5.6 is a linear symplectic map w.r.t. the symplectic form ω in (2.14). However we need to study the equation (5.4) which is posed on the subspace

$$H_{\perp}^s := \{U \in H^s \;:\; \Pi_S^{\perp}U = U\}. \quad (5.46)$$

Hence we need to find a correction of the map $\Phi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)[\cdot]$ which is symplectic with respect to the restricted symplectic form $\omega\big(\Pi_S^{\perp}(\cdot), \Pi_S^{\perp}(\cdot)\big)$. This fact plays an important role in order to prove item (iii) Prop. 5.11 (see also Remark 5.12). The Hamiltonian structure of these operators in the conjugated vector field will be used in Sect. 5.2 to ensure some terms in the energy forms vanish on resonances.

This is the content of the next lemma.

Lemma 5.9 *Recall (5.30), (5.31). Under the hypotheses of Lemma 5.3 the following holds. Consider the flow $\Psi_{\mathrm{m}}^{\tau} = \Psi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)$, $\tau \in [0, 1]$, defined by the system*

$$\begin{cases} \partial_{\tau}\Psi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)[\cdot] = \Pi_S^{\perp}Op^{\mathrm{BW}}(\mathbf{M}(\varepsilon\varphi, \varepsilon^{\beta}V))\Pi_S^{\perp}\Psi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)[\cdot], \\ \Psi_{\mathrm{m}}^{0}(\varepsilon\varphi, \varepsilon^{\beta}V)[\cdot] = \mathbb{1}_{\perp}[\cdot], \end{cases} \quad (5.47)$$

where $\mathbb{1}_{\perp}$ is the identity on $H_{\perp}^s$ in (5.46). There exists a real-to-real matrix of linear operators $\Theta^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)$ belonging to $\Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ for any $\rho \geq 0$ with estimates uniform in $\tau \in [0, 1]$, such that

$$\Psi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V) = \Pi_S^{\perp}\Phi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V) \circ (\mathbb{1} + \Theta^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V))\,\Pi_S^{\perp}. \quad (5.48)$$

In particular one has the expansion

$$\Psi_{\mathrm{m}}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V)[\cdot] = \Pi_S^{\perp}Op^{\mathrm{BW}}\begin{pmatrix} \mathrm{K}^+(\tau; \varepsilon\varphi, \varepsilon^{\beta}V) & \mathrm{K}^-(\tau; \varepsilon\varphi, \varepsilon^{\beta}V) \\ \mathrm{K}^-(\tau; \varepsilon\varphi, \varepsilon^{\beta}V) & \mathrm{K}^+(\tau; \varepsilon\varphi, \varepsilon^{\beta}V) \end{pmatrix}\Pi_S^{\perp} + \mathrm{R}^{\tau}(\varepsilon\varphi, \varepsilon^{\beta}V), \quad (5.49)$$

where symbols K^σ, $\sigma \in \{\pm\}$, are in (5.42) and $R^\tau \in \Sigma_1^{-3}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$. The inverse map $\Psi_m^{-\tau}(\varepsilon\varphi, \varepsilon^\beta V)[\cdot]$ admits a similar expansion as (5.49). Moreover

$$(\partial_t \Psi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot])_{|\tau=1} := Op^{BW}\big(\widetilde{G}^\perp(\varepsilon\varphi, \varepsilon^\beta V)\big)[\cdot] + \widetilde{R}_{\leq 2}^\perp(\varepsilon\varphi)[\cdot] + \widetilde{R}_{\geq 3}^\perp(\varepsilon\varphi, \varepsilon^\beta V)[\cdot], \tag{5.50}$$

where $\widetilde{G}^\perp(\varepsilon\varphi, \varepsilon^\beta V)$ is a 2×2 matrix of real valued symbols, independent of $\xi \in \mathbb{R}$, of the form

$$\widetilde{G}^\perp(\varepsilon\varphi, \varepsilon^\beta V) := \widetilde{G}_1^\perp(\varepsilon\varphi) + \widetilde{G}_2^\perp(\tau; \varepsilon\varphi, \varepsilon\varphi) + \widetilde{G}_{\geq 3}^\perp(\varepsilon\varphi, \varepsilon^\beta V), \tag{5.51}$$

where $\widetilde{G}_j^\perp \in \mathbf{SM}_j^0 \otimes \mathcal{M}_2(\mathbb{C})$, $j = 1, 2$ and where $\widetilde{G}_{\geq 3}^\perp(\tau; \varepsilon\varphi, \varepsilon^\beta V)$ is in $\mathcal{N}_p^0$ with

$$|\widetilde{G}_{\geq 3}^\perp|_{\mathcal{N}_p^0} \lesssim_s \|\varepsilon\varphi\|_{H^{p+\mu}}^3 + \|\varepsilon^\beta V\|_{H^{p+\mu}} + \varepsilon^5, \tag{5.52}$$

for any $p + s_0 + 3 \leq s$, $p \in \mathbb{N}$, and some $\mu > 1/2$, with estimates uniform in $\tau \in [0, 1]$. The operator $R_{\leq 2}^\perp$ has the form

$$\widetilde{R}_{\leq 2}^\perp(\tau; \varepsilon\varphi) := \widetilde{R}_1^\perp(\tau; \varepsilon\varphi) + \widetilde{R}_2^\perp(\tau; \varepsilon\varphi), \qquad \widetilde{R}_j^\perp \in \mathbf{M}_j^{-3} \otimes \mathcal{M}_2(\mathbb{C}), \quad j = 1, 2,$$

and $R_{\geq 3}^\perp$ satisfies

$$\|R_{\geq 3}^\perp(\varepsilon\varphi, \varepsilon^\beta V)Z\|_{H^{s+3}} \lesssim_s \|Z\|_{H^s}\big(\|\varepsilon\varphi\|_{H^s}^3 + \|\varepsilon^\beta V\|_{H^s} + \varepsilon^5\big). \tag{5.53}$$

Finally one has that $U^\perp \mapsto \Psi_m^\tau(\varepsilon\varphi, \varepsilon^\beta V)[U^\perp]$ is symplectic with respect to the restricted symplectic form $\omega\big(\Pi_S^\perp(\cdot), \Pi_S^\perp(\cdot)\big)$ (recall ω in (2.14)).

Proof To simplify the notation during this proof we denote by X the operator $Op^{BW}(\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))$. Recalling (5.30) and Lemma 5.4 we shall write

$$X = X_{\leq 2} + X_{\geq 3}, \tag{5.54}$$

$$X_{\leq 2} = X_{\leq 2}(\varepsilon\varphi) := Op^{BW}\begin{pmatrix} 0 & m_{\leq 2}(\varepsilon\varphi) \\ m_{\leq 2}(\varepsilon\varphi) & 0 \end{pmatrix},$$

$$X_{\geq 3} = X_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V) := Op^{BW}\begin{pmatrix} 0 & m_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V) \\ m_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V) & 0 \end{pmatrix},$$

where $m_{\leq} = m_1 + m_2$, $m_j \in \mathbf{SM}_j^0$, $j = 1, 2$ and $m_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V) \in \mathbf{SNH}_3^0$. To lighten the notation in the following we omit the dependence of the maps on $(\varepsilon\varphi, \varepsilon^\beta V)$. We look for a one-parameter group of bounded linear transformations $\Upsilon^\tau : H^s \to H^s$ such that

$$\Psi^\tau = \Pi_S^\perp \Phi_m^\tau \circ \Upsilon^\tau \Pi_S^\perp. \tag{5.55}$$

We differentiate both sides with respect to the parameter τ using (5.31), (5.47). One obtains

$$\Pi_S^\perp X \Pi_S^\perp \Psi^\tau = \Pi_S^\perp X \Phi_m^\tau \Upsilon^\tau \Pi_S^\perp + \Pi_S^\perp \Phi_m^\tau (\partial_\tau \Upsilon^\tau) \Pi_S^\perp,$$

which is equivalent to

$$-\Pi_S^\perp X \Pi_S \Phi_m^\tau \Upsilon^\tau \Pi_S^\perp = \Pi_S^\perp \Phi_m^\tau (\partial_\tau \Upsilon^\tau) \Pi_S^\perp,$$

by noticing that

$$\Pi_S^\perp X \Phi_m^\tau \Upsilon^\tau \Pi_S^\perp = \Pi_S^\perp X \Phi_m^\tau \Pi_S^\perp \Upsilon^\tau \Pi_S^\perp + \Pi_S^\perp X \Pi_S \Phi_m^\tau \Upsilon^\tau \Pi_S^\perp.$$

Then (5.55) is satisfied by the solution of the following Cauchy problem

$$\begin{cases} \partial_\tau \Upsilon^\tau = Y(\tau)(\Upsilon^\tau), & Y(\tau) = Y := -\Phi_m^{-\tau} X \Pi_S \Phi_m^\tau, \\ \Upsilon^0 = \mathbb{1}. \end{cases} \tag{5.56}$$

We observe that $X\Pi_S$ is a finite rank operator. In particular

$$X_{\leq 2}\Pi_S(Z) = \sum_{j\in S} Z_j \varsigma_j, \qquad X_{\geq 3}\Pi_S(Z) = \sum_{j\in S} Z_j \varrho_j,$$

with

$$\varsigma_j := X_{\leq 2}(\mathsf{e}_j), \qquad \varrho_j := X_{\geq 3}(\mathsf{e}_j), \qquad \mathsf{e}_j := \begin{pmatrix} e^{ijx} \\ e^{-ijx} \end{pmatrix}.$$

Recall that the operator Q in (5.32) is in $\Sigma_1^0[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ and so can be expanded as

$$\mathsf{Q}(\tau; \varepsilon\varphi, \varepsilon^\beta V) := \mathsf{Q}_{\leq 2}(\tau; \varepsilon\varphi) + \mathsf{Q}_{\geq 3}(\tau; \varepsilon\varphi, \varepsilon^\beta V),$$
$$\mathsf{Q}_{\leq 2}(\tau; \varepsilon\varphi) := \mathsf{Q}_1(\tau; \varepsilon\varphi) + \mathsf{Q}_2(\tau; \varepsilon\varphi) \qquad \mathsf{Q}_j \in \mathbf{M}_j^0 \otimes \mathcal{M}_2(\mathbb{C}), \quad j = 1, 2,$$

and some operator $\mathsf{Q}_{\geq 3}(\tau; \varepsilon\varphi, \varepsilon^\beta V)$ satisfying, for $s \geq s_0 \gg 1$,

$$\|\mathsf{Q}_{\geq 3}(\tau; \varepsilon\varphi, \varepsilon^\beta V)Z\|_{H^s} \lesssim_s \|Z\|_{H^s} \left(\|\varepsilon\varphi\|_{H^s}^3 + \|\varepsilon^\beta V\|_{H^s} \right), \tag{5.57}$$

uniformly in $\tau \in [0, 1]$. In turn the vector field Y is finite rank, since Φ_{m}^τ is a linear operator and $Y = -X\Pi_S + Y_{\leq 2} + Y_{\geq 3}$ with, by using (5.32) and recalling (2.48),

$$\begin{aligned} Y_{\leq 2}(\tau)[Z] := &-\sum_{j\in S}(Z, \mathsf{e}_j)_{L^2} \mathsf{Q}_{\leq 2}(-\tau; \varepsilon\varphi)[\varsigma_j] - \sum_{j\in S}(\mathsf{Q}_{\leq 2}(\tau; \varepsilon\varphi)[Z], \mathsf{e}_j)_{L^2} \varsigma_j \\ &-\sum_{j\in S}(\mathsf{Q}_{\leq 2}(\tau; \varepsilon\varphi)[Z], \mathsf{e}_j)_{L^2} \mathsf{Q}_{\leq 2}(-\tau; \varepsilon\varphi)[\varsigma_j] - \sum_{j\in S}(Z, \mathsf{e}_j)_{L^2} \mathsf{Q}_{\leq 2}(-\tau; \varepsilon\varphi)[\varrho_j] \\ &-\sum_{j\in S}(\mathsf{Q}_{\leq 2}(\tau; \varepsilon\varphi)[Z], \mathsf{e}_j)_{L^2} \varrho_j - \sum_{j\in S}(\mathsf{Q}_{\leq 2}(\tau; \varepsilon\varphi)[Z], \mathsf{e}_j)_{L^2} \mathsf{Q}_{\leq 2}(-\tau; \varepsilon\varphi)[\varrho_j], \end{aligned}$$

and $Y_{\geq 3}$ can be written explicitly as done for $Y_{\leq 2}$, but it depend on V or it is at least cubic in $\varepsilon\varphi$. We recall that by Lemma 5.6 we have that $\mathsf{Q}_1(\tau) \in \Sigma_1^0[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ (recall Definition 2.3-(iii)). Then it is evident that, by the structure of $Y_{\leq 2}$, it has an expansion in linear and quadratic terms with respect to $\varepsilon\varphi$. Moreover every term in $Y_{\leq 2}$ is an infinitely regularizing operator. Then $Y_{\leq 2}(\tau) \in \Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ for all $\rho > 0$.

By (5.57), the fact that $\mathsf{Q}_1(\tau) \in \Sigma_1^0[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ and that $Y_{\geq 3}$ is finite rank, we have

$$\|Y_{\geq 3}(\tau)[Z]\|_{H^{s+\rho}} \lesssim_s \|Z\|_{H^s} \left(\|\varepsilon\varphi\|_{H^s}^3 + \|\varepsilon^\beta V\|_{H^s} \right).$$

By classical theory of Banach space ODEs the Cauchy problem (5.56) is well-posed on $H^s(\mathbb{T})$. In particular Υ^τ is invertible and bounded on any $H^s(\mathbb{T})$ for $\tau \in [0, 1]$, namely

$$\|(\Upsilon^\tau)^{\pm 1}(Z)\|_{H^s} \lesssim \|Z\|_{H^s}, \qquad \tau \in [0, 1]. \tag{5.58}$$

By Taylor expanding Υ^τ at $\tau = 0$ we have

$$\Upsilon^\tau[\cdot] - \mathbb{1}[\cdot] = -\tau\, Y(0)[\cdot] + \int_0^\tau (1 - l)\, \partial_\tau^2 \Upsilon^l[\cdot]\, dl =: \Theta^\tau.$$

We claim that Θ^τ is in $\Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$. We observe that

$$\partial_\tau^2 \Upsilon^l = Z(l) \circ \Upsilon^l, \qquad Z(l) := \partial_\tau Y(l) + Y^2(l),$$

and

$$\partial_\tau Y = -\sum_{j\in S}(X(\Phi_{\mathrm{m}}^\tau(Z)), e^{\mathrm{i}jx})_{L^2}\,\Phi_{\mathrm{m}}^{-\tau}(\varsigma_j) + \sum_{j\in S}(\Phi_{\mathrm{m}}^\tau(Z), e^{\mathrm{i}jx})_{L^2}\,X(\Phi_{\mathrm{m}}^{-\tau}(\varsigma_j))$$
$$-\sum_{j\in S}(X(\Phi_{\mathrm{m}}^\tau(Z)), e^{\mathrm{i}jx})_{L^2}\,\Phi_{\mathrm{m}}^{-\tau}(\varrho_j) + \sum_{j\in S}(\Phi_{\mathrm{m}}^\tau(Z), e^{\mathrm{i}jx})_{L^2}\,X(\Phi_{\mathrm{m}}^{-\tau}(\varrho_j))\,,$$
$$Y^2 = Y_{\le 2}^2 + Y_{\ge 3}^2 + Y_{\le 2}Y_{\ge 3} + Y_{\ge 3}Y_{\le 2}\,.$$

Recalling (5.54), (5.32), we consider the splitting $X \circ \Phi_{\mathrm{m}}^\tau = \mathfrak{X}_{\le 2}(\tau) + \mathfrak{X}_{\ge 3}(\tau)$ where

$$\mathfrak{X}_{\le 2}(\tau) := X_{\le 2} + X_{\le 2}Q_{\le 2}(\tau), \qquad \mathfrak{X}_{\ge 3}(\tau) := X_{\ge 3} + X_{\le 2}Q_{\ge 3}(\tau) + X_{\ge 3}Q_{\le 2}(\tau) + X_{\ge 3}Q_{\ge 3}(\tau).$$

We write $Z(\tau) = Z_{\le 2}(\tau) + Z_{\ge 3}(\tau)$, where $Z_{\le 2} = Z_{\le 2}(\tau)$ is obtained by the sum of

- terms of $\partial_\tau Y$ where $X \circ \Phi_{\mathrm{m}}^\tau$ is replaced by $\mathfrak{X}_{\le 2}(\tau)$ and Φ_{m}^τ by $\mathbb{1} + Q_{\le 2}(\tau)$;
- the term $Y_{\le 2}^2$,

while $Z_{\ge 3} = Z_{\ge 3}(\tau)$ is defined by difference. We notice that $Z_{\ge 3}$ is sum of terms depending on V (plus terms with high homogeneity in $\varepsilon\varphi$), while $Z_{\le 2}$ is independent of V. Moreover, reasoning as for $Y_{\le 2}$ and $Y_{\ge 3}$, we have that, for all $\rho \ge 0$, $Z_{\le 2} \in \Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ and $Z_{\ge 3}$ satisfies

$$\|Z_{\ge 3}(\tau)[Z]\|_{H^{s+\rho}} \lesssim_s \|Z\|_{H^s}\big(\|\varepsilon\varphi\|_{H^s}^3 + \|\varepsilon^\beta V\|_{H^s}\big).$$

We set

$$\widetilde{\Theta}_1^\tau := -\tau\,X_{\le 2}\Pi_S + \tau\,Y_{\le 2}(0) + \int_0^\tau Z_{\le 2}(l) \circ \Upsilon^l\,dl,$$
$$\widetilde{\Theta}_2^\tau := -\tau\,X_{\ge 3}\Pi_S + \tau\,Y_{\ge 3}(0) + \int_0^\tau Z_{\ge 3}(l) \circ \Upsilon^l\,dl.$$

By the discussion on Υ^τ, the bound (5.58) and Lemma 5.4-(ii) we can conclude that $\widetilde{\Theta}_1^\tau + \widetilde{\Theta}_2^\tau$ belongs to the class $\Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$. Therefore formula (5.48) is proved. By (5.48), the expansions on the map Θ^τ, (5.42), Lemmata 2.8, 2.10 we deduce the expansion (5.49).

To prove (5.50) we proceed exactly as in the proof of (5.35) in Lemma 5.6. We have

$$\partial_\tau(\partial_t \Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V))[\cdot] = \Pi_S^\perp Op^{\mathrm{BW}}(\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))\Pi_S^\perp(\partial_t \Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V))[\cdot]$$
$$+ \Pi_S^\perp Op^{\mathrm{BW}}(\partial_t\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))\Pi_S^\perp \Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)[\cdot]$$

with $(\partial_t \Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V))[\cdot]_{|\tau=0} = 0$. By Duhamel formulation we deduce that

$$(\partial_t \Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V))[\cdot]$$
$$= \Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)\int_0^\tau (\Psi_{\mathrm{m}}^\sigma(\varepsilon\varphi, \varepsilon^\beta V))^{-1}\Pi_S^\perp Op^{\mathrm{BW}}(\partial_t\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))\Pi_S^\perp \Psi_{\mathrm{m}}^\sigma(\varepsilon\varphi, \varepsilon^\beta V)[\cdot]d\sigma.$$

The term $\Pi_S^\perp Op^{\mathrm{BW}}(\partial_t\mathbf{M}(\varepsilon\varphi, \varepsilon^\beta V))\Pi_S^\perp$ in the integral can be controlled for any $t \in [0, T]$ thanks to item (iii) of Lemma 5.4 and recalling (5.30). Therefore, the expansion (5.50) and estimate (5.53) follow using the expansion (5.49) for the flow $\Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)$, (5.23), the properties of m in Lemma 5.4 and Lemmata 2.2, 2.8, 2.10.

Corollary 5.10 *Consider the flow map* $\Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)$ *given in Lemma* 5.9. *Then, for any* $s \ge s_0$ *with* $s_0 \gg 1$ *one has that*

$$\sup_{\tau\in[0,1]} \|(\Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V))^{\pm 1}Z\|_{H^s} \le \|Z\|_{H^s}(1 + C_s\|\varepsilon\varphi\|_{H^s} + C_s\|\varepsilon^\beta V\|_{H^s}) \qquad (5.59)$$

for some $C_s > 0$ depending on s, for any $Z \in H^s$ such that $\Pi_S^\perp Z = Z$. Define the function

$$W^\perp := \left[\begin{smallmatrix} w^\perp \\ \overline{w}^\perp \end{smallmatrix}\right] := \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)[U^\perp], \tag{5.60}$$

where $U^\perp$ is in (5.3) and $\Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V) := \left(\Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)\right)_{|\tau=1}$. One has the equivalence

$$(1 + \varepsilon C_s)^{-1}\|\varepsilon^\beta V\|_{H^s} \leq \|W^\perp\|_{H^s} \leq (1 + \varepsilon C_s)\|\varepsilon^\beta V\|_{H^s}. \tag{5.61}$$

Proof Let us check the bound (5.59) on $\Psi_{\mathrm{m}}^\tau(\varepsilon\varphi, \varepsilon^\beta V)$. First of all, using (2.35) and Lemma 2.4 to estimate $\Theta^\tau(\varepsilon\varphi, \varepsilon^\beta V)$ (which belongs to $\Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$ for any $\rho \geq 0$) one can prove that (recall also the assumptions (4.4), (4.11) and (4.5))

$$\|\Theta^\tau(\varepsilon\varphi, \varepsilon^\beta V)Z\|_{H^{s+3}} \lesssim_s \|Z\|_{H^s}(\|\varepsilon\varphi\|_{H^s} + \|\varepsilon^\beta V\|_{H^s}), \tag{5.62}$$

for any $Z \in H^s$. The bound (5.59) follows by formula (5.48), estimates (5.43) and (5.62). Using Neumann series, estimate (5.62) and the smallness (4.5) one can prove that $(\mathbb{1} + \Theta^\tau)^{-1} - \mathbb{1}$ satisfies an estimate like (5.62). Then the estimate (5.59) on the inverse map follows using (5.43). The equivalence (5.61) follows by (5.59) using also (4.4), (4.11) and the smallness (4.5). $\quad\square$

5.1.3 Conjugation and First Energy Estimate

We are now in position to state our conjugation result.

Proposition 5.11 *Under the assumptions of Theorem* 5.1 *the following holds. Let* $U^\perp$ *in* (5.3) *be a solution of* (5.4), *then the function* $W^\perp$ *in* (5.60) *satisfies*

$$\partial_t W^\perp = iE\Pi_S^\perp Op^{\mathrm{BW}}\left(\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]\lambda(\varepsilon\varphi, \varepsilon^\beta V; x)\Lambda(\xi) + \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]a^{(0)}(\varepsilon\varphi, \varepsilon^\beta V; x, \xi)\right) W^\perp$$
$$+ iE\Pi_S^\perp \mathfrak{R}_{\leq 2}(\varepsilon\varphi)W^\perp + \Pi_S^\perp \mathfrak{R}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V) + \Pi_S^\perp \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)Res_\mathcal{H}(\varepsilon\varphi), \tag{5.63}$$

where the symbol $\lambda(\varepsilon\varphi, \varepsilon^\beta V)$ *is in* (5.11), *the symbol* $a^{(0)}$ *is real and belongs to* $\mathbf{S}\Sigma_1^0[r, 3]$ *with estimates uniform in* $t \in [0, T]$. *Moreover the following holds.*
(i) The remainder $\mathfrak{R}_{\leq 2}(\varepsilon\varphi)$ *is real (see Def.* 2.13*) and has the form*

$$\mathfrak{R}_{\leq 2}(\varepsilon\varphi) = \mathfrak{R}_1(\varepsilon\varphi) + \mathfrak{R}_2(\varepsilon\varphi, \varepsilon\varphi), \qquad \mathfrak{R}_j \in \mathbf{M}_j^{-2} \otimes \mathcal{M}_2(\mathbb{C}), \quad j = 1, 2. \tag{5.64}$$

(ii) The remainder $\mathfrak{R}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V)$ *has the form*

$$\mathfrak{R}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V) = \left(\mathfrak{R}_{\geq 3}^+(\varepsilon\varphi, \varepsilon^\beta V), \overline{\mathfrak{R}_{\geq 3}^+(\varepsilon\varphi, \varepsilon^\beta V)}\right)^T \tag{5.65}$$

and it satisfies the bound

$$\|\mathfrak{R}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V)\|_{H^s} \lesssim_s \|\varepsilon\varphi\|_{H^s(\mathbb{T}_L)}^3 \|\varepsilon^\beta V\|_{H^s} + \|\varepsilon^\beta V\|_{H^s}^2 + \varepsilon^5\|\varepsilon^\beta V\|_{H^s}. \tag{5.66}$$

(iii) The vector field

$$\Pi_S^\perp iE Op^{\mathrm{BW}}\left(\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]\lambda(\varepsilon\varphi, \varepsilon^\beta V; x)\Lambda(\xi) + \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]a^{(0)}(\varepsilon\varphi, \varepsilon^\beta V; x, \xi)\right)_{|V\equiv 0}\Pi_S^\perp$$
$$+ iE\Pi_S^\perp \mathfrak{R}_{\leq 2}(\varepsilon\varphi)\Pi_S^\perp$$

is Hamiltonian *according to Def.* 2.14.

Proof By (5.4) and (5.60) we have

$$\partial_t W^\perp = \left(\partial_t \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)\right)[U^\perp] + \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)[\partial_t U^\perp]$$

$$\overset{(5.50)}{=} Op^{\mathrm{BW}}\left(\widetilde{G}^\perp(\varepsilon\varphi, \varepsilon^\beta V)\right)[U^\perp] + \widetilde{\mathrm{R}}^\perp_{\leq 2}(\varepsilon\varphi)[U^\perp] + \widetilde{\mathrm{R}}^\perp_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V) U^\perp + \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)[\partial_t U^\perp]$$

$$\overset{(5.51),(5.4)}{=} Op^{\mathrm{BW}}\left((\widetilde{G}^\perp_{\leq 2})(\varepsilon\varphi)\right)[(\Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V))^{-1} W^\perp] \tag{5.67}$$

$$+ \widetilde{\mathrm{R}}^\perp_{\leq 2}(\varepsilon\varphi)[(\Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V))^{-1} W^\perp] \tag{5.68}$$

$$+ Op^{\mathrm{BW}}\left((\widetilde{G}^\perp_{\geq 3})(\varepsilon\varphi, \varepsilon^\beta V)\right)[U^\perp] + \widetilde{\mathrm{R}}^\perp_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V) U^\perp + \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)\Pi^\perp_S \widetilde{R}(\varepsilon\varphi; \varepsilon^\beta V)$$
$$\tag{5.69}$$

$$+ \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)\Pi^\perp_S \mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi) \tag{5.70}$$

$$+ \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)\Pi^\perp_S [iE Op^{\mathrm{BW}}\left(\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathrm{f}(\varepsilon\varphi, \varepsilon^\beta V; x)\right)\Lambda(\xi)\right)](\Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V))^{-1} W^\perp$$
$$\tag{5.71}$$

$$+ \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)\Pi^\perp_S [iE Op^{\mathrm{BW}}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathrm{g}(\varepsilon\varphi, \varepsilon^\beta V; x)\Lambda^{-1}(\xi)\right)](\Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V))^{-1} W^\perp \tag{5.72}$$

$$+ \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)\Pi^\perp_S iE\mathfrak{Q}(\varepsilon\varphi)(\Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V))^{-1} W^\perp. \tag{5.73}$$

By Lemma 5.2 and Proposition 4.3 we have $\mathfrak{Q}(\varepsilon\varphi) \in \Sigma_1^{-2}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$. Recall the expansion (5.49). By Lemmata 2.8, 2.10 we deduce that the remainder (5.73) can be absorbed in a remainder $\mathfrak{R}_{\leq 2}(\varepsilon\varphi)$ satisfying the conditions in item (i), up to a term satisfying (5.66). Similarly, using Lemmata 5.9, 5.2, one can check that the remainders in (5.69) can be absorbed in the remainder $\mathfrak{R}_{\geq 3}(\varepsilon\varphi, V)$ satisfying the conditions in item (ii). Consider now the terms in (5.71)–(5.72). Using formula (5.48), the estimates on the remainder $\Theta^\tau \in \Sigma_1^{-\rho}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$, for any $\rho \geq 0$, and the composition properties in Lemmata 2.8, 2.10 one deduces that

$$(5.71) + (5.72)$$
$$= \Pi^\perp_S \Phi_{\mathrm{m}}(\varepsilon\varphi, V)[iE Op^{\mathrm{BW}}\left(\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathrm{f}(\varepsilon\varphi, V; x)\right)\Lambda(\xi)\right)](\Phi_{\mathrm{m}}(\varepsilon\varphi, V))^{-1} W^\perp$$
$$+ \Pi^\perp_S \Phi_{\mathrm{m}}(\varepsilon\varphi, V)[iE Op^{\mathrm{BW}}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathrm{g}(\varepsilon\varphi, V; x)\Lambda^{-1}(\xi)\right)](\Phi_{\mathrm{m}}(\varepsilon\varphi, V))^{-1} W^\perp$$

up to remainders which satisfy the properties of $\mathfrak{R}_{\leq 2}, \mathfrak{R}_{\geq 3}$ in items $(i), (ii)$. Using the expansions (5.33)–(5.34), by Lemma 2.2, Remark 2.11 and the (5.14) in Lemma 5.3 one gets

$$(5.71) + (5.72) = \Pi^\perp_S iE Op^{\mathrm{BW}}\left(\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}(\lambda(\varepsilon\varphi, V))\right)\Lambda(\xi) + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \widetilde{a}^{(0)}(\varepsilon\varphi, \varepsilon^\beta V; x, \xi)\right) W^\perp$$

for some $\widetilde{a}^{(0)} \in \mathbf{S}\Sigma_1^0[r, 3]$, up to remainders that can be absorbed in terms $\mathfrak{R}_{\leq 2}, \mathfrak{R}_{\geq 3}$ satisfying items $(i), (ii)$ and where $\lambda(\varepsilon\varphi, V)$ is in Lemma 5.3.
Consider the term (5.68). By Lemma 5.9 we have that $\mathrm{R}^\perp_{\leq 2} \in \Sigma_1^{-3}[r, 3] \otimes \mathcal{M}_2(\mathbb{C})$. By Lemma 5.9 the term in (5.68) can be absorbed in terms $\mathfrak{R}_{\leq 2}, \mathfrak{R}_{\geq 3}$ satisfying items $(i), (ii)$. Finally, consider the term in (5.67). Recalling (5.51), the expansion (5.49) and Lemmata 2.2, 2.10 we deduce that (5.67) is the sum of a pseudo-differential operator of order zero (with symbols in $\mathbf{S}\Sigma_1^0[r, 3]$) plus remainders of the form $\mathfrak{R}_{\leq 2}, \mathfrak{R}_{\geq 3}$ as in items (i)–(ii). The item (iii) follows recalling item (ii) of Lemma 5.2 and the fact that the map $U^\perp \mapsto \Psi^\tau_m(\varepsilon\varphi, \varepsilon^\beta V)[U^\perp]$ is symplectic (see Lemma 5.9). This concludes the proof.

Recalling Proposition 5.11 we have the expansions of the symbols in the r.h.s. of (5.63)

$$\lambda(\varepsilon\varphi, \varepsilon^\beta V) = 1 + \lambda_{\leq 2}(\varepsilon\varphi) + \lambda_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V), \tag{5.74}$$

$$a^{(0)}(\varepsilon\varphi, \varepsilon^\beta V) = a^{(0)}_{\leq 2}(\varepsilon\varphi) + a^{(0)}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V), \tag{5.75}$$

where

$$\lambda_{\leq 2}(\varepsilon\varphi) = \lambda_1(\varepsilon\varphi) + \lambda_2(\varepsilon\varphi, \varepsilon\varphi), \qquad \lambda_j \in \mathbf{SM}^0_j, \quad j = 1, 2, \tag{5.76}$$

$$a^{(0)}_{\leq 2}(\varepsilon\varphi) = a^{(0)}_1(\varepsilon\varphi) + a^{(0)}_2(\varepsilon\varphi, \varepsilon\varphi), \qquad a^{(0)}_j \in \mathbf{SM}^0_j, \quad j = 1, 2, \tag{5.77}$$

and $\lambda_{\geq 3}, a^{(0)}_{\geq 3}$ are in $\mathcal{N}^0_p$ with

$$|\lambda_{\geq 3}|_{\mathcal{N}^0_p} \lesssim_s \|\varepsilon\varphi\|^3_{H^{p+\mu}} + \|\varepsilon^\beta V\|_{H^{p+\mu}}, \tag{5.78}$$

$$|a^{(0)}_{\geq 3}|_{\mathcal{N}^0_p} \lesssim_p \|\varepsilon\varphi\|^3_{H^{p+\mu}} + \|\varepsilon^\beta V\|_{H^{p+\mu}}, \tag{5.79}$$

for any $p + \mu \leq s$, $p \in \mathbb{N}$, and some $\mu > 1/2$.

We now expand the right hand side of (5.63) in degrees of homogeneity in $\varepsilon\varphi$. We define

$$\mathcal{M} := \begin{pmatrix} \mathcal{M}^+ & 0 \\ 0 & \overline{\mathcal{M}^+} \end{pmatrix}, \qquad \mathcal{M}^+ := \mathcal{M}^+_0 + \mathcal{M}^+_1 + \mathcal{M}^+_2,$$

$$\mathcal{Z} := \begin{pmatrix} \mathcal{Z}^+ & \mathcal{Z}^- \\ \overline{\mathcal{Z}^-} & \overline{\mathcal{Z}^+} \end{pmatrix}, \qquad \mathcal{Z}^+ := \mathcal{Z}^+_1 + \mathcal{Z}^+_2, \quad \mathcal{Z}^- := \mathcal{Z}^-_1 + \mathcal{Z}^-_2 \tag{5.80}$$

where

$$\mathcal{M}^{(+)}_0 := \Pi^\perp_S Op^W(\Lambda(\xi)), \qquad \mathcal{M}^{(+)}_j := \Pi^\perp_S Op^{BW}\left(\lambda_j(\varepsilon\varphi; x)\Lambda(\xi)\right), \qquad j = 1, 2,$$

$$\mathcal{Z}^{(\pm)}_j := \Pi^\perp_S Op^{BW}\left(a^{(0)}_j(\varepsilon\varphi; x, \xi)\right) + \Pi^\perp_S \mathfrak{R}^\pm_j(\varepsilon\varphi).$$

and

$$\mathcal{M}_> := \begin{pmatrix} \mathcal{M}^+_> & 0 \\ 0 & \overline{\mathcal{M}^+_>} \end{pmatrix}, \qquad \mathcal{Z}_> := \begin{pmatrix} \mathcal{Z}^+_> & \mathcal{Z}^-_> \\ \overline{\mathcal{Z}^-_>} & \overline{\mathcal{Z}^+_>} \end{pmatrix},$$

$$\mathcal{M}^+_> := Op^{BW}\left(\lambda_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V; x, \xi)\Lambda(\xi)\right)$$

$$\mathcal{Z}^+_> := \Pi^\perp_S Op^{BW}\left(a^{(0)}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V; x, \xi)\right), \qquad \mathcal{Z}^-_> := \Pi^\perp_S Op^{BW}\left(a^{(0)}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V; x, \xi)\right), \tag{5.81}$$

We finally define the function $\widetilde{\mathcal{Z}} := \widetilde{\mathcal{Z}}(\varepsilon\varphi, \varepsilon^\beta V)$ as

$$\widetilde{\mathcal{Z}} := \Pi^\perp_S \mathfrak{R}_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V) + \Pi^\perp_S \Psi_{\mathrm{m}}(\varepsilon\varphi, \varepsilon^\beta V)\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi), \tag{5.82}$$

where $\mathfrak{R}_{\geq 3}$ is in (5.65). Using the notation above we rewrite the equation (5.63) as

$$\partial_t W^\perp = iE(\mathcal{M} + \mathcal{Z})W^\perp + iE(\mathcal{M}_> + \mathcal{Z}_>)W^\perp + \widetilde{\mathcal{Z}}. \tag{5.83}$$

Remark 5.12 By item (iii) of Proposition 5.11 one deduces that the operators $\mathcal{M}, \mathcal{Z}$ and $\mathcal{M}_>, \mathcal{Z}_>$ are self-adjoint according to Def. 2.14. This will be used to obtain the commutator structure in (5.86).

Consider the Fourier multiplier $\langle D \rangle := Op^W(\langle\xi\rangle)$, set $\mathcal{D} := \begin{bmatrix} \langle D \rangle & 0 \\ 0 & \langle D \rangle \end{bmatrix}$ and

$$\mathcal{N}_s(W^\perp) := \|W^\perp\|^2_{H^s} = \frac{1}{2}(\mathcal{D}^{2s}W^\perp, W^\perp)_{L^2} \tag{5.84}$$

(see (2.48) and (2.3)). We have the following.

Lemma 5.13 (First energy estimates) *Under the assumptions of Proposition* 5.11 *the following holds. One has that, for $t \in [0, T]$*

$$\partial_t \mathcal{N}_s(W^\perp) = \tfrac{1}{2} (iE\mathcal{D}^{2s} \mathcal{A}(\varepsilon\varphi) W^\perp, W^\perp)_{L^2} + \mathcal{B}(\varepsilon\varphi, \varepsilon^\beta V),$$

where
- *the remainder $\mathcal{B}(\varepsilon\varphi, \varepsilon^\beta V)$ satisfies*

$$\begin{aligned}
\sup_{t \in [0,T]} \|\mathcal{B}(\varepsilon\varphi, \varepsilon^\beta V)\|_{H^s} &\lesssim_s \sup_{t \in [0,T]} \|\varepsilon\varphi\|_{H^s}^3 \sup_{t \in [0,T]} \|\varepsilon^\beta V\|_{H^s}^2 \\
&+ \sup_{t \in [0,T]} \|\varepsilon^\beta V\|_{H^s}^3 + \varepsilon^5 \sup_{t \in [0,T]} \|\varepsilon^\beta V\|_{H^s} ;
\end{aligned} \tag{5.85}$$

- *the operator $\mathcal{A}(\varepsilon\varphi)$ has the form*

$$\mathcal{A}(\varepsilon\varphi) := \begin{bmatrix} A^+ & A^- \\ A^- & A^+ \end{bmatrix} := \mathcal{D}^{-2s}\big[iE(\mathcal{M} + \mathcal{Z})(\varepsilon\varphi), iE\mathcal{D}^{2s}\big], \tag{5.86}$$

and it is real, where $\mathcal{Z}$ is in (5.80). *In particular*

$$A^\sigma(\varepsilon\varphi) = A_1^\sigma(\varepsilon\varphi) + A_2^\sigma(\varepsilon\varphi, \varepsilon\varphi), \qquad \sigma \in \{\pm\}, \tag{5.87}$$

where (recall (5.80))

$$\begin{aligned}
A_k^+(\varepsilon\varphi) &:= \langle D \rangle^{-2s}\big[i(\mathcal{M}_k^+(\varepsilon\varphi) + \mathcal{Z}_k^+(\varepsilon\varphi)), i\langle D \rangle^{2s}\big], \quad k = 1, 2, \\
A_k^-(\varepsilon\varphi) &:= \langle D \rangle^{-2s}\Big(i\mathcal{Z}_k^-(\varepsilon\varphi)i\langle D \rangle^{2s} + i\langle D \rangle^{2s}i\mathcal{Z}_k^-(\varepsilon\varphi)\Big), \quad k = 1, 2.
\end{aligned} \tag{5.88}$$

Moreover the following holds.
- *The operator A_1^+ has the form*

$$A_1^+(\varepsilon\varphi) := Op^{\mathrm{BW}}(\mathsf{a}_1(\varepsilon\varphi; x, \xi)) + \mathsf{A}_1(\varepsilon\varphi), \qquad \mathsf{a}_1 \in \mathbf{SM}_1^0, \quad \mathsf{A}_1 \in \mathbf{M}_1^{-2}, \tag{5.89}$$

and $(A_1^+)^\sigma(j, 0, j) = 0$ for any $\sigma \in \{\pm\}$, $j \in \mathbb{Z}$.
- *The operator A_2^+ has the form*

$$A_2^+(\varepsilon\varphi, \varepsilon\varphi) := Op^{\mathrm{BW}}(\mathsf{a}_2(\varepsilon\varphi; x, \xi)) + \mathsf{A}_2(\varepsilon\varphi), \qquad \mathsf{a}_2 \in \mathbf{SM}_2^0, \quad \mathsf{A}_2 \in \mathbf{M}_2^{-2}, \tag{5.90}$$

and $(A_2^+)^{\sigma,-\sigma}(j, p_1, p_2, j) = 0$, for any $\sigma \in \{\pm\}$, $j, p_1, p_2 \in \mathbb{Z}$.
- *The operators A_k^- belongs to $\mathbf{M}_k^0$, $k = 1, 2$.*

Proof By (5.83), Remark 5.12, and an explicit computation one gets

$$\begin{aligned}
\partial_t \mathcal{N}_s(W^\perp) &= \tfrac{1}{2}(\mathcal{D}^{2s} \partial_t W^\perp, W^\perp)_{L^2} + \tfrac{1}{2}(\mathcal{D}^{2s} W^\perp, \partial_t W^\perp)_{L^2} \\
&= \tfrac{1}{2}(iE\mathcal{D}^{2s} \mathcal{A}(\varepsilon\varphi) W^\perp, W^\perp)_{L^2} \tag{5.91} \\
&+ \tfrac{1}{2}\Big(iE\mathcal{D}^{2s}\big(\langle D \rangle^{-2s}\big[i\mathcal{M}_>^+, i\langle D \rangle^{2s}\big]\big)W^\perp, W^\perp\Big)_{L^2} \tag{5.92} \\
&+ \tfrac{1}{2}(iE\big[iE\mathcal{Z}_>, iE\mathcal{D}^{2s}\big]W^\perp, W^\perp)_{L^2} \tag{5.93} \\
&+ \tfrac{1}{2}(\mathcal{D}^{2s}\widetilde{\mathcal{Z}}, W^\perp)_{L^2} + (\mathcal{D}^{2s} W^\perp, \widetilde{\mathcal{Z}})_{L^2}, \tag{5.94}
\end{aligned}$$

where $\mathcal{A}(\varepsilon\varphi)$ is in (5.86) and where we used the fact that

$$([iE\mathcal{M}_0^+, iE\mathcal{D}^{2s}]W^\perp, W^\perp)_{L^2} = 0.$$

The terms in (5.94) satisfies the bound (5.85), by (5.82), Cauchy–Schwarz inequality, estimates (5.66) and (5.59), the equivalence (5.61), bounds (4.4), (4.11), the smallness (4.5) and Lemma 4.1 to estimate the residual.

By (5.81) $\mathcal{Z}_>^{\pm}$ are bounded para-differential operators with symbols in $\mathbf{SNH}_3^0[r]$. Then, by Lemma 2.2, we have that the operator $\mathcal{D}^{-2s}[iE\mathcal{Z}_>, iE\mathcal{D}^{2s}]$ is bounded. Therefore one can check that the term in (5.93) satisfies the bound (5.85).

Consider the term (5.92). By Lemmata 2.2, 2.1, and recalling the definition of $\mathcal{M}_>^+$ in (5.81), we deduce that the operator

$$\langle D\rangle^{-2s}\big[\langle D\rangle^{2s}, Op^{\mathrm{BW}}(\lambda_{\geq 3}(\varepsilon\varphi, \varepsilon^\beta V; x, \xi)\Lambda(\xi))\big]$$

is bounded and

$$\|\langle D\rangle^{-2s}\big[\langle D\rangle^{2s}, Op^{\mathrm{W}}((\lambda_{\geq 3})_\chi(\varepsilon\varphi, \varepsilon^\beta V; x, \xi)\Lambda(\xi))\big]h\|_{H^s}$$
$$\lesssim_s \|h\|_{H^s}\Big(\sup_{t\in[0,T]}\|\varepsilon\varphi\|_{H^s}^3 + \sup_{t\in[0,T]}\|\varepsilon^\beta V\|_{H^s}\Big),$$

where we used (5.78). This implies that the term (5.92) satisfies (5.85). Consider the term (5.91) which has components in (5.86)–(5.88). Since $\mathcal{M} + \mathcal{Z}$ in (5.80) is a bounded para-differential operator plus a smoothing remainder, the operator $\mathcal{M}$ is diagonal, one can note, using Lemmata 2.2, 2.1 and the commutator structure in A_k^+, that A_k^- is in $\mathbf{M}_k^0$ and A_k^+ admits the expansions (5.89)–(5.90). Since A_k^+ involves a commutator with the diagonal operator $\langle D\rangle^{2s}$ it is easy to see that $(A_1^+)^\sigma(j, p, k)$ and $(A_2^+)^{\sigma,-\sigma}(j, p_1, p_2, k)$ are zero when $j = k$.

5.2 Modified Energies

We introduce the following *energy forms* (recall (2.48), (5.84)):

$$\mathrm{E} := \tfrac{1}{2}(\mathrm{i}\,E\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi)W^\perp, W^\perp)_{L^2},$$

$$\mathcal{Q}(\varphi) := \begin{bmatrix} Q^+ & Q^- \\ Q^- & Q^+ \end{bmatrix} \in \Sigma_1^0[r, 3] \otimes \mathcal{M}_2(\mathbb{C}),\tag{5.95}$$

$$\mathcal{E}_s := \mathcal{N}_s + \mathrm{E}.\tag{5.96}$$

In particular we look for an operator $\mathcal{Q}(\varphi)$ which is self-adjoint, real-to-real and that admits an expansion of the form (recall Def. 2.3)

$$Q^\sigma(\varphi) = Q_1^\sigma(\varphi) + Q_2^\sigma(\varphi, \varphi), \qquad Q_k^\sigma \in \mathbf{M}_k^0, \quad k = 1, 2, \ \ \sigma \in \{\pm\}.\tag{5.97}$$

We shall construct the modified energies $Q_1^\sigma, Q_2^\sigma, \sigma \in \{\pm\}$, in a suitable way. More precisely we prove the following.

Proposition 5.14 *Under the assumptions of Theorem 5.1 the following holds. There exists a real, self-adjoint operator $\mathcal{Q}(\varphi)$ as in (5.95) with Q_i^σ, $i = 1, 2$, $\sigma \in \{\pm\}$, of the form (5.97) such that the energy $\mathcal{E}_s$ in (5.96)–(5.95) satisfies, for any $t \in [0, T]$ and for $s \geq s_0$ with $s_0 \gg 1$,*

$$|\partial_t \mathcal{E}_s(W^\perp)| \lesssim_s \sup_{t\in[0,T]}\|\varphi\|_{H^s}^3 \sup_{t\in[0,T]}\|\varepsilon^\beta V\|_{H^s}^2 + \sup_{t\in[0,T]}\|\varepsilon^\beta V\|_{H^s}^3 + \varepsilon^5 \sup_{t\in[0,T]}\|\varepsilon^\beta V\|_{H^s}.\tag{5.98}$$

Moreover the following holds:
- **(Off-diagonal).** *The operators Q_k^- belong to $\mathbf{M}_k^{-1}$, $k = 1, 2$.*
- **(Diagonal I).** *The operator Q_1^+ has the form*

$$Q_1^+(\varepsilon\varphi) := Op^{\mathrm{BW}}(\mathsf{q}_1(\varepsilon\varphi; x, \xi)) + \mathsf{Q}_1(\varepsilon\varphi), \qquad \mathsf{q}_1 \in \mathbf{SM}_1^0, \quad \mathsf{Q}_1 \in \mathbf{M}_1^{-2}.\tag{5.99}$$

- **(Diagonal II).** *The operator Q_2^+ has the form*

$$Q_2^+(\varepsilon\varphi) := Op^{\mathrm{BW}}(\mathsf{q}_2(\varepsilon\varphi; x, \xi)) + \mathsf{Q}_2(\varepsilon\varphi), \qquad \mathsf{q}_2 \in \mathbf{SM}_2^0, \quad \mathsf{Q}_2 \in \mathbf{M}_2^{-1}.\tag{5.100}$$

Finally (recall 2.32, (2.34) *and* (2.40)*) one has*

$$|q_2^{\sigma_1\sigma_2}(j_1, j_2, \xi)| \lesssim \gamma^{-1} \max\{\langle j_1\rangle, \langle j_2\rangle\}^{\mu}, \tag{5.101}$$

$$|(Q_2)^{\sigma_1\sigma_2}(\xi, j_1, j_2, j_3)| + |(Q_2^-)^{\sigma_1\sigma_2}(\xi, j_1, j_2, j_3)| \lesssim \gamma^{-1} \frac{\max_2\{\langle j_1\rangle, \langle j_2\rangle, \langle j_3\rangle\}^{\mu+1}}{\max\{\langle j_1\rangle, \langle j_2\rangle, \langle j_3\rangle\}}, \tag{5.102}$$

for some $\mu > 0$.

In Sect. 5.2.1 we construct the operators Q_1^σ, $\sigma \in \{\pm\}$, while in Sect. 5.2.2 we construct the operators Q_2^σ, $\sigma \in \{\pm\}$, Then in Sect. 5.2.3 we prove Proposition 5.14 and then we conclude the proof of Theorem 5.1.

5.2.1 First Order Homological Equation

In this section we prove the following result.

Lemma 5.15 (Homological equation 1) *There exist operators Q_1^+ satisfying* (5.99) *such that*

$$A_1^+ + Q_1^+(iE\Lambda\varepsilon\varphi) + \langle D\rangle^{-2s}[\langle D\rangle^{2s} Q_1^+, i\Lambda] = 0 \tag{5.103}$$

and $Q_1^- \in \mathbf{M}_1^{-1}$ *such that*

$$A_1^- + Q_1^-(iE\Lambda\varphi) - i\langle D\rangle^{-2s}(\langle D\rangle^{2s} Q_1^- \Lambda + \Lambda \langle D\rangle^{2s} Q_1^-) = 0, \tag{5.104}$$

where A_1^+, A_1^- are the operators introduced in (5.88) *and Λ is the Fourier multiplier in* (2.5). *Moreover $Q_1^\sigma = \overline{Q_1^\sigma}$, $\sigma \in \{\pm\}$.*

Proof We start by writing the equation (5.103) for the Fourier coefficients. We have, fixing $j, k \in S^c$, $p \in S$ and $\sigma \in \{\pm\}$ such that $\sigma p = j - k$

$$\Big((Q_1^+)^\sigma(j, p, k)\,\sigma i\,\Lambda(p) + i(\Lambda(k) - \Lambda(j))(Q_1^+)^\sigma(j, p, k) + (A_1^+)^\sigma(j, p, k)\Big)\widehat{\varphi}^\sigma(p) = 0.$$

Then we set

$$(Q_1^+)^\sigma(j, p, k) = \frac{1}{i\big(\Lambda(j) - \Lambda(k) - \sigma\Lambda(p)\big)} (A_1^+)^\sigma(j, p, k). \tag{5.105}$$

The coefficients $(Q_1^+)^\sigma(j, p, k)$ vanish as $j = k$ because by Lemma 5.13 we have $(A_1^+)^\sigma(j, 0, j) = 0$. Recalling (5.89), Definitions 2.3, 2.5, (2.20), we have

$$(A_1^+)^\sigma(j, p, k) := (a_1)^\sigma(\sigma p, (j + k)/2)\chi\Big(\frac{2|p|}{\langle j + k\rangle}\Big) + A_1^\sigma(j, p, k),$$

where

$$|A_1^\sigma(j, p, k)| \lesssim \langle p\rangle^\mu (\max\{\langle p\rangle, \langle k\rangle\})^{-2} \tag{5.106}$$

for some $\mu > 0$. We observe that, by definition of S in (4.1), the modulus of a site in S^c is always greater than the modulus of a site in S, then we must have $jk > 0$ because otherwise the momentum condition implies $|p| = |j|+|k|$, which is not possible since $p \in S$, $j, k \in S^c$. Now we claim that we can expand the denominator in (5.105)

$$\frac{1}{\Lambda(j) - \Lambda(k) - \sigma\Lambda(p)} = \begin{cases} \dfrac{1}{\sigma p - \sigma\Lambda(p)} + r_{-2}(p; j, k) & \text{if } j, k > 0 \\[4mm] -\dfrac{1}{\sigma p + \sigma\Lambda(p)} + \tilde{r}_{-2}(p; j, k) & \text{if } j, k < 0 \end{cases}$$

where

$$|r_{-2}(p; j, k)|, |\tilde{r}_{-2}(p; j, k)| \lesssim \langle p \rangle^\mu \, (\min\{\langle j \rangle, \langle k \rangle\})^{-2} \, , \tag{5.107}$$

for some $\mu > 0$. Let us prove the claim in the case $j, k > 0$. The other case is similar. We use that

$$\Lambda(j) = |j| + \frac{m}{2|j|} + \mathcal{O}(|j|^{-3}).$$

We have that

$$\frac{1}{\Lambda(j) - \Lambda(k) - \sigma \Lambda(p)} - \frac{1}{\sigma p - \sigma \Lambda(p)} = \frac{\Lambda(k) - \Lambda(j) + \sigma p}{\sigma p - \sigma \Lambda(p)}$$

$$\frac{1}{\Lambda(j) - \Lambda(k) - \sigma \Lambda(p)} =: r_{-2}(p; j, k) \, .$$

By momentum condition $j - k = \sigma p$ we have that

$$\sigma p + \Lambda(k) - \Lambda(j) = \frac{m}{2}(k^{-1} - j^{-1}) + \mathcal{O}((\min\{\langle j \rangle, \langle k \rangle\})^{-3})$$

$$= \frac{m}{2} \frac{\sigma p}{j k} + \mathcal{O}((\min\{\langle j \rangle, \langle k \rangle\})^{-3}) = \mathcal{O}((\min\{\langle j \rangle, \langle k \rangle\})^{-2}) \, .$$

Therefore, using bound (3.8), we notice that r_{-2} satisfies (5.107). This concludes the proof of the claim.

Therefore (recall that $j, k \in S^c$ then $j, k \neq 0$) we write

$$(Q_1^+)^\sigma (j, p, k) = \tilde{\mathfrak{q}}_1^\sigma (j, p, k) \chi \left(\frac{2|p|}{\langle j + k \rangle} \right) + \tilde{Q}_1^\sigma (j, p, k) \tag{5.108}$$

where

$$\tilde{\mathfrak{q}}_1^\sigma (j, p, k) := \frac{1}{i(\tilde{\chi}(j + k) \, \sigma p - \sigma \Lambda(p))} \, (\mathtt{a}_1)^\sigma (\sigma p, (j + k)/2) \tag{5.109}$$

and (recall $j + k \neq 0$ because $j k > 0$)

$$\tilde{\chi}(x) = \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x < 0 \end{cases}$$

and $\tilde{Q}_1^\sigma (j, p, k)$ is defined by difference. Using the bound (5.107) and Lemma 3.8 (see bound (3.8)) we deduce the estimate

$$|\tilde{Q}_1^\sigma (j, p, k)| \lesssim \langle p \rangle^\mu \, (\max\{\langle p \rangle, \langle k \rangle\})^{-2} \, ,$$

for some $\mu > 0$ (possibly larger than the one appearing in (5.106)). Then the operator $\tilde{Q}_1$ of the form (2.30)–(2.31) (with $p = 1$) with coefficients $\tilde{Q}_1^\sigma (j, p, k)$ is a remainder in the class $\mathbf{M}_1^{-2}$. We shall consider a slight modification of $\tilde{\mathfrak{q}}_1^\sigma$. Let us now introduce the C^∞ cut-off function $\eta : \mathbb{R} \to \mathbb{R}$ defined as

$$\eta(y) := \begin{cases} 1 & y \geq 1/2, \\ 0 & y \leq -1/2. \end{cases} \tag{5.110}$$

Let us define the symbol

$$\mathfrak{q}_1^\sigma (\varepsilon \varphi; x, \xi) := \eta(\xi) \tilde{\mathfrak{q}}_1^\sigma (L_+ \varepsilon \varphi; x, \xi) + (1 - \eta(\xi)) \tilde{\mathfrak{q}}_1^\sigma (L_- \varepsilon \varphi; x, \xi)$$

where $L_\pm$ are the operator defined by linearity as

$$L_\pm e^{i\sigma p x} = \frac{1}{\pm \sigma p - \sigma \Lambda(p)} e^{i\sigma p x}, \quad p \in S,$$

and 0 otherwise. It is easy to check that $q_1^\sigma(\varepsilon\varphi; x, \xi) \in \mathbf{SM}_1^0$ and $q_1^\sigma - \widetilde{q}_1^\sigma$ is infinitely smoothing. By the discussion above we have that the operator Q_1^+ of the form (2.30) with coefficients as in (5.108) can be written as

$$Q_1^+(\varepsilon\varphi) := Op^{\mathrm{BW}}(q_1(\varepsilon\varphi; x, \xi)) + Q_1(\varepsilon\varphi), \qquad q_1 \in \mathbf{SM}_1^0, \quad Q_1 \in \mathbf{M}_1^{-2},$$

where $Q_1 := Op^{\mathrm{BW}}(q_1 - \widetilde{q}_1) + \widetilde{Q}_1$. We now study the equation (5.104). In Fourier we have, fixing $j, k \in S^c$, $p \in S$ and $\sigma \in \{\pm\}$ such that $\sigma p = j - k$

$$(Q_1^-)^\sigma(j, p, k)\, \sigma\, i\, \Lambda(p)\, \widehat{\varphi}^\sigma(p) - i(\Lambda(j) + \Lambda(k))(Q_1^-)^\sigma(j, p, k)\, \widehat{\varphi}^\sigma(p)$$
$$+ (A_1^-)^\sigma(j, p, k)\, \widehat{\varphi}^\sigma(p) = 0.$$

We set

$$(Q_1^-)^\sigma(j, p, k) = \frac{1}{i\big(\Lambda(j) + \Lambda(k) - \sigma \Lambda(p)\big)} (A_1^-)^\sigma(j, p, k). \qquad (5.111)$$

First we note that, by Lemma 5.13, $A_1^- \in \mathbf{M}_1^0$ and then

$$|(A_1^-)^\sigma(j, p, k)| \lesssim \langle p \rangle^\mu$$

for some $\mu > 0$. Moreover $|\Lambda(j) + \Lambda(k) - \sigma \Lambda(p)| \geq \Lambda(j)$, since $\min\{\langle j \rangle, \langle k \rangle\} \geq |p|$. Therefore the operator Q_1^- with coefficients in (5.111) belongs to $\mathbf{M}_1^{-1}$. This concludes the proof. □

5.2.2 Second order Homological Equation

In this section we prove the following result.

Lemma 5.16 (Homological equation 2) *Recall* (5.17) *and Lemma* 5.13. *Consider the operators* Q_1^σ, $\sigma \in \{\pm\}$, *given by Lemma* 5.15. *There exist operators* $Q_2^+ \in \mathbf{M}_2^0$ *such that*

$$Q_2^+(iE\Lambda\varepsilon\varphi, \varepsilon\varphi) + Q_2^+(\varepsilon\varphi, iE\Lambda\varepsilon\varphi) + \langle D \rangle^{-2s}[\langle D \rangle^{2s} Q_2^+, i\Lambda] =$$
$$= -A_2^+ - Q_1^+(M_1(\varepsilon\varphi)\varepsilon\varphi) \qquad (5.112)$$
$$- \langle D \rangle^{-2s}\Big([\langle D \rangle^{2s} Q_1^+, i(\mathcal{M}_1^+ + \mathcal{Z}_1^+)] - i(\langle D \rangle^{2s} Q_1^- \mathcal{Z}_1^- + \mathcal{Z}_1^- \langle D \rangle^{2s} Q_1^-)\Big).$$

and $Q_2^- \in \mathbf{M}_2^{-1}$ *such that*

$$Q_2^-(iE\Lambda\varepsilon\varphi, \varepsilon\varphi) + Q_2^-(\varepsilon\varphi, iE\Lambda\varepsilon\varphi) + i\Lambda Q_2^- + iQ_2^-\Lambda =$$
$$= -A_2^- - Q_1^-(M_1(\varepsilon\varphi)\varepsilon\varphi) - i[Q_1^+, \mathcal{Z}_1^-] - \langle D \rangle^{-2s}[\langle D \rangle^{2s} Q_1^+, i\mathcal{Z}_1^-]$$
$$+ \langle D \rangle^{-2s}\Big(i(\langle D \rangle^{2s} Q_1^-(\mathcal{M}_1^+ + \mathcal{Z}_1^+) + (\mathcal{M}_1^+ + \mathcal{Z}_1^+)\langle D \rangle^{2s} Q_1^-)\Big).$$
$$(5.113)$$

Moreover the operator Q_2^+ *has the form* (5.100) *and satisfies* (5.102). *Finally* $Q_2^\sigma = \overline{Q_2^\sigma}$, $\sigma \in \{\pm\}$.

Proof We start by considering the equation (5.112). We claim that the right hand side of (5.112) can be rewritten as

$$F(\varepsilon\varphi) = Op^{\text{BW}}(\mathtt{f}(\varepsilon\varphi; x, \xi)) + \mathrm{F}(\varepsilon\varphi)$$

where $\mathtt{f}(\varepsilon\varphi; x, \xi) \in \mathbf{SM}_2^0$ and $\mathrm{F}(\varepsilon\varphi) \in \mathbf{M}_2^{-1}$. By Lemma 5.13–(5.90) A_2^+ is the sum of a pseudo differential operator of order 0 and a 2-smoothing operator. By Lemma 5.15 the same holds for Q_1^+. The most critical term is the one involving $\mathcal{M}_1^+ + \mathcal{Z}_1^+$, because it is unbounded. Indeed, we recall that $\mathcal{M}_1^+ + \mathcal{Z}_1^+$ in (5.80) is the sum of a pseudo differential operator of order 1 and an operator $\mathbf{M}_1^{-2}$, that for notation convenience we call respectively $\mathcal{P}$, $\mathcal{B}$. Then the commutator $[\langle D\rangle^{2s}Q_1^+, i(\mathcal{M}_1^+ + \mathcal{Z}_1^+)]$ looses just $2s$ derivatives, indeed

$$[\langle D\rangle^{2s}Q_1^+, i(\mathcal{M}_1^+ + \mathcal{Z}_1^+)] = [\langle D\rangle^{2s}Op^{\text{BW}}(\mathtt{q}_1), i\mathcal{P}] + [\langle D\rangle^{2s}\mathcal{Q}_1, i\mathcal{P}] + [\langle D\rangle^{2s}Q_1^+, i\mathcal{B}].$$

The first term is the commutator between two pseudo differential operators, then it is a pseudo differential operator and it gains one derivative, hence it has a symbol in $\mathbf{SM}_2^{2s}$; the second term is, by composition, a linear operator in $\mathbf{M}_2^{2s-1}$ because $\mathcal{Q}_1 \in \mathbf{M}_1^{-2}$; the third term is, by composition, in $\mathbf{M}_2^{2s-2}$. Therefore

$$\langle D\rangle^{-2s}[\langle D\rangle^{2s}Q_1^+, i(\mathcal{M}_1^+ + \mathcal{Z}_1^+)]$$

is the sum of a pseudo differential operator with symbol in $\mathbf{SM}_2^0$ and an operator in $\mathbf{M}_2^{-1}$. By recalling that $\mathcal{Z}_1^-$ is bounded and $Q_1^- \in \mathbf{M}_1^{-1}$, the remaining terms in the right hand side of (5.112) are $\mathbf{M}_2^{-1}$. This proves the claim.

Recalling Definitions 2.3 2.5 and Lemma 2.6 we have the following bounds on the coefficients of $\mathtt{f}(\varepsilon\varphi; x, \xi)$ and $\mathrm{F}(\varepsilon\varphi)$:

$$|\mathtt{f}^{\sigma_1\sigma_2}(p_1, p_2, j)| \lesssim_s \max_1\{\langle p_1\rangle, \langle p_2\rangle\}^\mu ,$$

$$|\mathrm{F}^{\sigma_1\sigma_2}(j, p_1, p_2, k)| \lesssim_s \frac{\max_1\{\langle p_1\rangle, \langle p_2\rangle\}^\mu \max_2\{\langle p_1\rangle, \langle p_2\rangle, \langle k\rangle\}^{\mu+1}}{\max_1\{\langle p_1\rangle, \langle p_2\rangle, \langle k\rangle\}} , \tag{5.114}$$

for any $j, k \in S^c$, $p_1, p_2 \in S$, $\sigma_1, \sigma_2 \in \{\pm 1\}$ and for some $\mu > 0$.

Then we define

$$\left(Q_2^+\right)^{\sigma_1\sigma_2}(j, p_1, p_2, k) = \frac{\mathtt{f}^{\sigma_1\sigma_2}(p_1, p_2, (j+k)/2)\chi\left(\frac{2|p|}{\langle j+k\rangle}\right) + \mathrm{F}^{\sigma_1\sigma_2}(j, p_1, p_2, k)}{i\left(\Lambda(j) - \Lambda(k) - \sigma_1\Lambda(p_1) - \sigma_2\Lambda(p_2)\right)},$$

$$\text{if} \quad \Lambda(j) - \Lambda(k) - \sigma_1\Lambda(p_1) - \sigma_2\Lambda(p_2) \neq 0 \quad \text{and} \quad \sigma_1 p_1 + \sigma_2 p_2 = j - k , \tag{5.115}$$

and

$$\left(Q_2^+\right)^{\sigma_1\sigma_2}(j, p_1, p_2, k) = 0$$

$$\text{if} \quad \Lambda(j) - \Lambda(k) - \sigma_1\Lambda(p_1) - \sigma_2\Lambda(p_2) = 0 \quad \text{and} \quad \sigma_1 p_1 + \sigma_2 p_2 = j - k . \tag{5.116}$$

We have to show that the operator defined by (5.115)–(5.116) solves the equation (5.112). We know that the only possible resonances at order four are (j, k, p_1, p_2) such that (up to permutations)

$$|j| = |k|, \quad |p_1| = |p_2| \quad \text{or} \quad |j| = |p_1|, \quad |k| = |p_2|.$$

Since S is symmetric and $j, k \notin S$, $p_1, p_2 \in S$ then the only possible case is $|j| = |k|$, $|p_1| = |p_2|$ with $\sigma_1 \neq \sigma_2$. By the Lemma 3.8 we have that

$$\Lambda(j) - \Lambda(k) - \sigma_1\Lambda(p_1) - \sigma_2\Lambda(p_2) = 0, \qquad \sigma_1 p_1 + \sigma_2 p_2 = j - k,$$

only if

$$j = k, \quad p_1 = p_2, \quad \sigma_1 = -\sigma_2. \tag{5.117}$$

Therefore the operator Q_2^+ defined by (5.115)–(5.116) is a solution of the equation (5.112) if and only if the coefficients of the operator in the right hand side of (5.112) are zero when (5.117) holds true. In order to prove this we study each summand in the r.h.s. of (5.112). By the previous step in Lemma 5.15, in particular (5.105) and the fact that $(A_1^+)^\sigma(j, 0, j) = 0$, the coefficients of the operator $Q_1^+(iEM_1(\varepsilon\varphi)\varepsilon\varphi)$ are zero as $j = k$. By Lemma 5.13 the same holds for the coefficients of A_2^+.

Consider the linear operator $[\langle D \rangle^{2s} Q_1^+, i(\mathcal{M}_1^+ + \mathcal{Z}_1^+)]$. Recalling the expansion (2.32), and the (5.105), (5.88), for indexes satisfying

$$\sigma_1 p_1 + \sigma_2 p_2 = j - k,$$

we have that

$$
\begin{aligned}
[\langle D \rangle^{2s} Q_1^+, &\, i(\mathcal{M}_1^+ + \mathcal{Z}_1^+)]^{\sigma_1, \sigma_2}(j, p_1, p_2, k) \\
&= (\mathcal{M}_1^+ + \mathcal{Z}_1^+)^{\sigma_1}(j, p_1, k')(\mathcal{M}_1^+ + \mathcal{Z}_1^+)^{\sigma_2}(k', p_2, k) \times \\
&\quad \times \left(\frac{\langle j \rangle^{2s} - \langle k' \rangle^{2s}}{i(\Lambda(j) - \Lambda(k') - \sigma_1 \Lambda(p_2))} - \frac{\langle k' \rangle^{2s} - \langle k \rangle^{2s}}{i(\Lambda(k') - \Lambda(k) - \sigma_2 \Lambda(p_2))} \right)
\end{aligned}
$$

where

$$k' = j - \sigma_1 p_1 = k + \sigma_2 p_2.$$

If (5.117) holds one can check that $[\langle D \rangle^{2s} Q_1^+, i(\mathcal{M}_1^+ + \mathcal{Z}_1^+)]^{\sigma_1, -\sigma_1}(j, p_1, -p_1, j) = 0$. Concerning the term

$$\langle D \rangle^{2s} Q_1^- \mathcal{Z}_1^- + \mathcal{Z}_1^- \langle D \rangle^{2s} Q_1^-$$

one can reason as above using equation (5.111). This proves the claim.

It remains to prove the (5.100) for the operator Q_2^+. It is easy to check (using (5.114) and (3.9) in Lemma 3.8) that the coefficients

$$\frac{\mathrm{F}^{\sigma_1 \sigma_2}(j, p_1, p_2, k)}{\Lambda(j) - \Lambda(k) - \sigma_1 \Lambda(p_1) - \sigma_2 \Lambda(p_2)}$$

appearing in (5.115) contribute to a smoothing remainder satisfying (5.102). To estimate the contribution of the first summand in the r.h.s. of (5.115) we reason as done for the operator Q_1^+ in Lemma 5.15. Notice that, by the momentum condition $\sigma_1 p_1 + \sigma_2 p_2 = j - k$ we deduce that, if $jk < 0$, then

$$\max\{|j|, |k|\} \lesssim \max\{|p_1|, |p_2|\}.$$

Therefore the coefficients

$$\frac{\mathfrak{f}^{\sigma_1 \sigma_2}(p_1, p_2, (j+k)/2)}{\Lambda(j) - \Lambda(k) - \sigma_1 \Lambda(p_1) - \sigma_2 \Lambda(p_2)}, \qquad jk < 0$$

satisfy the (5.102), which means that they contribute to a smoothing remainder in $\mathbf{M}_2^{-1}$. Hence we only study the case $jk > 0$. We Taylor expand the denominator

$$
\frac{1}{\Lambda(j) - \Lambda(k) - \sigma_1 \Lambda(p_1) - \sigma_2 \Lambda(p_2)} =
\begin{cases}
\dfrac{1}{\sigma_1 p_1 + \sigma_2 p_2 + \sigma_1 \Lambda(p_1) + \sigma_2 \Lambda(p_2)} + r_1, & j, k > 0 \\[3ex]
\dfrac{1}{-\sigma_1 p_1 - \sigma_2 p_2 + \sigma_1 \Lambda(p_1) + \sigma_2 \Lambda(p_2)} + \tilde{r}_1, & j, k < 0,
\end{cases}
$$

where $r_1 := r_1(j, p_1, p_2, k)$, $\tilde{r}_1 := \tilde{r}_1(j, p_1, p_2, k)$ satisfy

$$|r_1(j, p_1, p_2, k)|, |\tilde{r}_1(j, p_1, p_2, k)| \leq \max\{\langle p_1 \rangle, \langle p_2 \rangle\}^\mu \min\{\langle j \rangle, \langle k \rangle\}^{-1},$$

and $p_1, p_2 \in S$. Recalling (5.110) we set

$$q_2(\varepsilon\varphi; x, \xi) := \sum_{p_1, p_2 \in \mathbb{Z}, \sigma \in \{\pm\}} q_2^{\sigma\sigma}(p_1, p_2, \xi) \varphi_{p_1}^\sigma \varphi_{p_2}^\sigma e^{i\sigma(p_1+p_2)x}$$

$$+ \sum_{p_1, p_2 \in \mathbb{Z}} q_2^{+-}(p_1, p_2, \xi) \varphi_{p_1} \overline{\varphi_{p_2}} e^{i(p_1-p_2)x}$$

with

$$q_2^{\sigma_1\sigma_2}(p_1, p_2, \xi) := \frac{i\eta(\xi) f^{\sigma_1\sigma_2}(p_1, p_2, \xi)}{\sigma_1 p_1 + \sigma_2 p_2 - \sigma_1 \Lambda(p_1) - \sigma_2 \Lambda(p_2)} + \frac{i(1 - \eta(\xi)) f^{\sigma_1\sigma_2}(p_1, p_2, \xi)}{-\sigma_1 p_1 - \sigma_2 p_2 - \sigma_1 \Lambda(p_1) - \sigma_2 \Lambda(p_2)}.$$

By (5.114) and (3.9) one can check that $q_2(\varphi; x, \xi)$ is a symbol in the class $\mathbf{SM}_2^0$ with coefficients satisfying (5.101). Therefore, the discussion above implies that the operator Q_2^+ with coefficients in (5.115) can be written as

$$Q_2^+ := Op^{\mathrm{BW}}(q_2(\varepsilon\varphi; x, \xi)) + \mathcal{Q}_2(\varepsilon\varphi)$$

for some $\mathcal{Q}_2(\varepsilon\varphi) \in \mathbf{M}_2^{-1}$. Then formula (5.100) follows.

Consider now the equation (5.113). By Lemma 5.13 and Lemma 5.15, using Lemmata 2.2, 2.10 2.8 and 2.11, we note that the right hand side of (5.113) is a linear operator $\mathrm{R}(\varphi)$ in the class $\mathbf{M}_2^0$ with coefficients satisfying

$$|\mathrm{R}^{\sigma_1\sigma_2}(j, p_1, p_2, k)| \lesssim_2 \max\{\langle p_1 \rangle, \langle p_2 \rangle, \langle k \rangle\}^\mu, \tag{5.118}$$

for any $j, k, p_1, p_2 \in \mathbb{Z}$, $\sigma_1, \sigma_2 \in \{\pm\}$, for some $\mu > 0$. Then we define

$$(Q_2^-)^{\sigma_1\sigma_2}(j, p_1, p_2, k) = \frac{i\mathrm{R}^{\sigma_1\sigma_2}(j, p_1, p_2, k)}{\Lambda(j) + \Lambda(k) - \sigma_1 \Lambda(p_1) - \sigma_2 \Lambda(p_2)}, \tag{5.119}$$

for any $j, k, p_1, p_2 \in \mathbb{Z}$, $\sigma_1, \sigma_2 \in \{\pm\}$ with $\sigma_1 p_1 + \sigma_2 p_2 = j + k$. If $jk > 0$ then

$$\max\{|j|, |k|\} \lesssim \max\{|p_1|, |p_2|\}.$$

Hence, without loss of generality we can assume that $j > 0$, $k < 0$, otherwise Q_2^- is infinitely smoothing because j, k are equivalent to the inner frequencies p_1, p_2. The momentum condition reads as $|j| - |k| = \sigma_1 p_1 + \sigma_2 p_2$. If $p_1 = p_2 = 0$ the bound on the coefficients (5.119) is trivial. Hence we assume that $|p_1| + |p_2| \geq 1$ (they cannot be both zero). Assume that k is such that

$$|k| > 4(|p_1| + |p_2|). \tag{5.120}$$

Then we have that $\Lambda(k) > \Lambda(p_1) + \Lambda(p_2)$, indeed

$$\Lambda(k)^2 > 16(|p_1| + |p_2|)^2 + m > (|p_1| + |p_2|)^2 + 2m^2 + m(|p_1| + |p_2|)$$
$$+ m > (\Lambda(p_1) + \Lambda(p_2))^2,$$

where we used that $m \leq 2$ and $|p_1| + |p_2| \geq 1$. Hence

$$\Lambda(j) + \Lambda(k) - \sigma_1 \Lambda(p_1) - \sigma_2 \Lambda(p_2) \geq \Lambda(j).$$

Therefore the operator with coefficients in (5.119) with k such that (5.120) holds belongs to $\mathbf{M}_1^{-1}$. If k is such that $|k| \leq 4(|p_1| + |p_2|)$, then by the momentum condition $|j| \leq$

$5(|p_1| + |p_2|)$. Therefore we can argue as before, because j, k are equivalent to the inner frequencies p_1, p_2.

By estimates (5.118) and the above discussion we deduce that the coefficients in (5.119) satisfy (5.102). Hence the thesis follows.

5.2.3 Proof of Proposition 5.14

Recalling (5.96) we have

$$\partial_t \mathcal{E}_s(W^\perp) = \partial_t \mathcal{N}_s(W^\perp) + \frac{1}{2}(\mathrm{i}\, E \mathcal{D}^{2s} \partial_t(\mathcal{Q}(\varepsilon\varphi)) W^\perp, W^\perp)_{L^2}$$
$$+ (\mathrm{i}\, E \mathcal{D}^{2s} \mathcal{Q}(\varepsilon\varphi) \dot{W}^\perp, W^\perp)_{L^2} + (\mathrm{i}\, E \mathcal{D}^{2s} \mathcal{Q}(\varepsilon\varphi) W^\perp, \dot{W}^\perp)_{L^2}.$$

Therefore, using Lemma 5.13 and the equation (5.83) (recall also Remark 5.12), one gets

$$\partial_t \mathcal{E}_s(W^\perp) = \tfrac{1}{2}(\mathrm{i}E\mathcal{D}^{2s}\mathcal{L}(\varepsilon\varphi)W^\perp, W^\perp)_{L^2} + \mathcal{B}(\varepsilon\varphi, V)$$
$$+ \frac{1}{2}(\mathrm{i}E[\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi), \mathrm{i}E(\mathcal{M}_> + \mathcal{Z}_>)]W^\perp, W^\perp)_{L^2} \qquad (5.121)$$
$$+ \tfrac{1}{2}(\mathrm{i}E\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi)\widetilde{\mathcal{Z}}, W^\perp)_{L^2} + \tfrac{1}{2}(\mathrm{i}E\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi)W^\perp, \widetilde{\mathcal{Z}})_{L^2}$$

where $\mathcal{M}_>, \mathcal{Z}_>$ are in (5.81), $\widetilde{\mathcal{Z}}$ is in (5.82) and where

$$\mathcal{L}(\varepsilon\varphi) := \mathcal{A}(\varepsilon\varphi) + \partial_t(\mathcal{Q}(\varepsilon\varphi)) + \mathcal{D}^{-2s}[\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi), \mathrm{i}E(\mathcal{M} + \mathcal{Z})],$$

with $\mathcal{M}, \mathcal{Z}$ in (5.80). By (5.95), (5.97), (5.80) we can write

$$\mathcal{L}(\varepsilon\varphi) = \begin{pmatrix} \mathcal{L}^{(\mathrm{diag})} & \mathcal{L}^{(\mathrm{off})} \\ \hline \mathcal{L}^{(\mathrm{off})} & \mathcal{L}^{(\mathrm{diag})} \end{pmatrix},$$

$$\mathcal{L}^{(\mathrm{diag})} := A^+ + \partial_t(Q^+(\varepsilon\varphi))$$
$$+ \langle D \rangle^{-2s}\Big([\langle D \rangle^{2s} Q^+, \mathrm{i}(\mathcal{M}^+ + \mathcal{Z}^+)] - \mathrm{i}(\langle D \rangle^{2s} Q^- \mathcal{Z}^- + \mathcal{Z}^- \langle D \rangle^{2s} Q^-)\Big),$$
$$(5.122)$$

$$\mathcal{L}^{(\mathrm{off})} := A^- + \partial_t(Q^-(\varepsilon\varphi))$$
$$+ \langle D \rangle^{-2s}\Big([\mathcal{D}^{2s} Q^+, \mathrm{i}\mathcal{Z}^-] - \mathrm{i}(\langle D \rangle^{2s} Q^-(\mathcal{M}^+ + \mathcal{Z}^+) + (\mathcal{M}^+ + \mathcal{Z}^+)\langle D \rangle^{2s} Q^-)\Big).$$
$$(5.123)$$

Notice that by (5.17) we have, for $\sigma \in \{\pm\}$,

$$\partial_t(Q^\sigma(\varepsilon\varphi)) = Q_1^\sigma(\varepsilon\varphi_t) + Q_2^\sigma(\varepsilon\varphi_t, \varepsilon\varphi) + Q_2^\sigma(\varepsilon\varphi, \varepsilon\varphi_t)$$
$$= Q_1^\sigma(\mathrm{i}E\Lambda\varphi) + Q_1^\sigma(M_1(\varepsilon\varphi)\varepsilon\varphi) + Q_2^\sigma(\mathrm{i}E\Lambda\varepsilon\varphi, \varepsilon\varphi) \qquad (5.124)$$
$$+ Q_2^\sigma(\varepsilon\varphi, \mathrm{i}E\Lambda\varepsilon\varphi) + \mathbf{R}_{\geq 3}^\sigma$$

where

$$\mathbf{R}_{\geq 3}^\sigma := Q_1^\sigma(M_2(\varepsilon\varphi)\varepsilon\varphi + M_{\geq 3}(\varepsilon\varphi)\varepsilon\varphi)$$
$$+ Q_2^\sigma\big(\varepsilon\varphi, M_1(\varepsilon\varphi)\varepsilon\varphi + M_2(\varepsilon\varphi)\varepsilon\varphi + M_{\geq 3}(\varepsilon\varphi)\varepsilon\varphi\big)$$
$$+ Q_2^\sigma(M_1(\varepsilon\varphi)\varepsilon\varphi + M_2(\varepsilon\varphi)\varepsilon\varphi + M_{\geq 3}(\varepsilon\varphi)\varepsilon\varphi, \varepsilon\varphi) \qquad (5.125)$$
$$+ Q_1^\sigma(-\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)) + Q_2^\sigma(\varepsilon\varphi, -\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi)) + Q_2^\sigma(-\mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi), \varepsilon\varphi).$$

Then, using (5.87), (5.97), (5.80), (5.124), we Taylor expand $\mathcal{L}^{(\mathrm{diag})}$ and $\mathcal{L}^{(\mathrm{off})}$ in (5.122)–(5.123) as

$$\mathcal{L}^{(\mathrm{diag})} = \mathcal{L}_1^{(\mathrm{diag})} + \mathcal{L}_2^{(\mathrm{diag})} + \mathcal{L}_{\geq 3}^{(\mathrm{diag})}, \qquad \mathcal{L}^{(\mathrm{off})} = \mathcal{L}_1^{(\mathrm{off})} + \mathcal{L}_2^{(\mathrm{off})} + \mathcal{L}_{\geq 3}^{(\mathrm{off})}$$

where

$$\mathcal{L}_1^{(\mathrm{diag})} = A_1^+ + Q_1^+(\mathrm{i} E\Lambda\varepsilon\varphi) + \langle D\rangle^{-2s}[\langle D\rangle^{2s} Q_1^+, \mathrm{i}\Lambda],$$

$$\mathcal{L}_2^{(\mathrm{diag})} = A_2^+ + Q_1^+(M_1(\varepsilon\varphi)\varepsilon\varphi) + Q_2^+(\mathrm{i} E\Lambda\varepsilon\varphi, \varepsilon\varphi) + Q_2^+(\varepsilon\varphi, \mathrm{i} E\Lambda\varepsilon\varphi),$$
$$\qquad + \langle D\rangle^{-2s}\Big([\langle D\rangle^{2s} Q_2^+, \mathrm{i}\Lambda] + [\langle D\rangle^{2s} Q_1^+, \mathrm{i}(\mathcal{M}_1^+ + \mathcal{Z}_1^+)] - \mathrm{i}(\langle D\rangle^{2s} Q_1^- \mathcal{Z}_1^-$$
$$\qquad + \mathcal{Z}_1^- \langle D\rangle^{2s} Q_1^-)\Big),$$

$$\mathcal{L}_{\geq 3}^{(\mathrm{diag})} = \mathbf{R}_{\geq 3}^+ + \langle D\rangle^{-2s}\Big([\langle D\rangle^{2s} Q_2^+, \mathrm{i}(\mathcal{M}_1^+ + \mathcal{Z}_1^+)]\Big)$$
$$\qquad + \langle D\rangle^{-2s}\Big([\langle D\rangle^{2s} Q_1^+, \mathrm{i}(\mathcal{M}_2^+ + \mathcal{Z}_2^+)] + [\langle D\rangle^{2s} Q_2^+, \mathrm{i}(\mathcal{M}_2^+ + \mathcal{Z}_2^+)]\Big),$$
$$\qquad - \mathrm{i}\langle D\rangle^{-2s}\Big(\langle D\rangle^{2s} Q_1^- \mathcal{Z}_2^- + \mathcal{Z}_2^- \langle D\rangle^{2s} Q_1^- + \langle D\rangle^{2s} Q_2^- \mathcal{Z}_1^- + \mathcal{Z}_1^- \langle D\rangle^{2s} Q_2^-\Big)$$
$$\qquad - \mathrm{i}\langle D\rangle^{-2s}\Big(\langle D\rangle^{2s} Q_2^- \mathcal{Z}_2^- + \mathcal{Z}_2^- \langle D\rangle^{2s} Q_2^-\Big),$$

$$\tag{5.126}$$

and

$$\mathcal{L}_1^{(\mathrm{off})} = A_1^- + Q_1^-(\mathrm{i} E\Lambda\varphi) - \mathrm{i}\langle D\rangle^{-2s}(\langle D\rangle^{2s} Q_1^- \Lambda + \Lambda \langle D\rangle^{2s} Q_1^-),$$

$$\mathcal{L}_2^{(\mathrm{off})} = A_2^- + Q_1^-(M_1(\varepsilon\varphi)\varepsilon\varphi) + Q_2^-(\mathrm{i} E\Lambda\varphi, \varepsilon\varphi) + Q_2^-(\varepsilon\varphi, \mathrm{i} E\Lambda\varepsilon\varphi)$$
$$\qquad + \langle D\rangle^{-2s}\Big([\langle D\rangle^{2s} Q_1^+, \mathrm{i}\mathcal{Z}_1^-]\Big)$$
$$\qquad - \langle D\rangle^{-2s}\Big(\mathrm{i}(\langle D\rangle^{2s} Q_1^-(\mathcal{M}_1^+ + \mathcal{Z}_1^+) + (\mathcal{M}_1^+ + \mathcal{Z}_1^+)\langle D\rangle^{2s} Q_1^-)\Big)$$
$$\qquad - \mathrm{i}\langle D\rangle\Big(\langle D\rangle^{2s} Q_2^- \Lambda + \Lambda\langle D\rangle^{2s} Q_2^-\Big),$$

$$\mathcal{L}_{\geq 3}^{(\mathrm{off})} = \mathbf{R}_{\geq 3}^- + \langle D\rangle^{-2s}\Big([\langle D\rangle^{2s} Q_2^+, \mathrm{i}\mathcal{Z}_1^-] + [\langle D\rangle^{2s} Q_1^+, \mathrm{i}\mathcal{Z}_2^-] + [\langle D\rangle^{2s} Q_2^+, \mathrm{i}\mathcal{Z}_2^-]\Big)$$
$$\qquad - \mathrm{i}\langle D\rangle^{-2s}\Big(\langle D\rangle^{2s} Q_2^-(\mathcal{M}_1^+ + \mathcal{Z}_1^+) + (\mathcal{M}_1^+ + \mathcal{Z}_1^+)\langle D\rangle^{2s} Q_2^-\Big)$$
$$\qquad - \mathrm{i}\langle D\rangle^{-2s}\Big(D\rangle^{2s} Q_1^- \mathcal{Z}_2^+ + \mathcal{Z}_2^+ \langle D\rangle^{2s} Q_1^-\Big)$$
$$\qquad - \mathrm{i}\langle D\rangle^{-2s}\Big(\langle D\rangle^{2s} Q_2^- \mathcal{Z}_2^+ + \mathcal{Z}_2^+ \langle D\rangle^{2s} Q_2^-\Big).$$

$$\tag{5.127}$$

By Lemmata 5.15, 5.16 we have constructed $Q^{(1)}$, $Q^{(2)}$ such that

$$\mathcal{L}_1^{(\mathrm{diag})} = \mathcal{L}_2^{(\mathrm{diag})} = \mathcal{L}_1^{(\mathrm{off})} = \mathcal{L}_2^{(\mathrm{off})} \equiv 0.$$

Then, by (5.121), we deduce

$$\partial_t \mathcal{E}_s(W^\perp) = \tfrac{1}{2}(\mathrm{i} E\mathcal{D}^{2s}\mathcal{L}_{\geq 3}(\varepsilon\varphi)W^\perp, W^\perp)_{L^2} + \mathcal{B}(\varepsilon\varphi, V)$$
$$\qquad + \frac{1}{2}(\mathrm{i} E[\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi), \mathrm{i} E(\mathcal{M}_{>} + \mathcal{Z}_{>})]W^\perp, W^\perp)_{L^2}$$
$$\qquad + \tfrac{1}{2}(\mathrm{i} E\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi)\widetilde{\mathcal{Z}}, W^\perp)_{L^2} + \tfrac{1}{2}(\mathrm{i} E\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi)W^\perp, \widetilde{\mathcal{Z}})_{L^2}$$

where

$$\mathcal{L}_{\geq 3}(\varepsilon\varphi) := \begin{pmatrix} \mathcal{L}_{\geq 3}^{(\mathrm{diag})} & \mathcal{L}_{\geq 3}^{(\mathrm{off})} \\ \overline{\mathcal{L}_{\geq 3}^{(\mathrm{off})}} & \overline{\mathcal{L}_{\geq 3}^{(\mathrm{diag})}} \end{pmatrix}.$$

We claim that the operators $\mathcal{L}_{\geq 3}^{(\mathrm{diag})}$, $\mathcal{L}_{\geq 3}^{(\mathrm{off})}$ belong to $\mathbf{NH}_3^0[r]$ and satisfy for $s \geq s_0$ (s_0 large enough)

$$\|\mathcal{L}_{\geq 3}^{(\mathrm{diag})}h\|_{H^s} + \|\mathcal{L}_{\geq 3}^{(\mathrm{off})}\overline{h}\|_{H^s} \lesssim_s L^2 \|\varphi\|_{H^s(\mathbb{T})}^3 \|h\|_{H^s}, \tag{5.128}$$

for all $h \in H^s$.

We start considering the term $\mathcal{L}_{\geq 3}^{(\mathrm{diag})}$ in (5.126). First of all we recall that, by (5.97), (5.99), (5.100), we have

$$Q^+(\varphi) = Op^{\mathrm{W}}(\mathfrak{q}_1 + \mathfrak{q}_2) + \mathcal{Q}_1(\varphi) + \mathcal{Q}_2(\varphi), \qquad \mathfrak{q}_1 \in \mathbf{SM}_1^0, \quad \mathfrak{q}_2 \in \mathbf{SM}_2^0,$$
$$\mathcal{Q}_1 \in \mathbf{M}_1^{-2}, \quad \mathcal{Q}_2 \in \mathbf{M}_2^{-1}.$$

We also recall that

$$Q^-(\varphi) = Q_1^-(\varphi) + Q_2^-(\varphi), \qquad Q_1^- \in \mathbf{M}_1^{-1}, \quad Q_2^- \in \mathbf{M}_2^{-1}. \tag{5.129}$$

We now analyse the cubic contributions coming from each summand in (5.126) and we show that they are all bounded operators. We point out that $\mathcal{M}^+$ is an unbounded operator of order one, while $\mathcal{Z}^+$ is bounded. By the properties of Q_j^σ, $\sigma = \pm$, $j = 1, 2$ discussed above we have that $\mathbf{R}_{\geq 3}^{(j)}$ in (5.125) are bounded.

In the second line of (5.126) the terms are bounded because $Q_j^+ = Op^{\mathrm{W}}(\mathfrak{q}_j) + \mathcal{Q}_j$ and

(i) the commutator between the pseudo differential operators

$$Op^{\mathrm{W}}(\mathfrak{q}_j), \quad \mathcal{M}_k^+ + \mathcal{Z}_k^+, \quad (j, k) \in \{(1, 2), (2, 1), (2, 2)\},$$

gains one derivative;

(ii) the composition between $\mathcal{Q}_j$ and $\mathcal{M}_k^+ + \mathcal{Z}_k^+$, $(j, k) \in \{(1, 2), (2, 1), (2, 2)\}$ is bounded by Lemma 2.10 and because $\mathcal{Q}_j$ are 1-smoothing.

The third and fourth lines in (5.126) are bounded because they are compositions of bounded operators.

Now we make a similar analysis for $\mathcal{L}_{\geq 3}^{(\mathrm{off})}$ in (5.127). As before the only critical terms are the ones involving $\mathcal{M}^+$, namely the terms of the third and fourth lines in (5.127). By (5.129) and Lemma 2.10 the composition of Q_j^- with $\mathcal{M}_k^+$, $(j, k) \in \{(1, 2), (2, 1), (2, 2)\}$, is bounded. The estimate (5.128) is consequence of (5.102) and (5.76), (5.77), (5.64) for the estimates on $\mathcal{M}_k^+$, $\mathcal{Z}_k^\pm$, $k = 1, 2$.

Recall $\widetilde{\mathcal{Z}}$ in (5.82). By (5.66), (5.59) and (4.7) we have

$$\|\widetilde{\mathcal{Z}}\|_{H^s} \lesssim_s \|\varepsilon\varphi\|_{H^s}^3 \|\varepsilon^\beta V\|_{H^s} + \|\varepsilon^\beta V\|_{H^s}^2 + \varepsilon^5. \tag{5.130}$$

We observe that $\left[\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi), \mathrm{i}E(\mathcal{M}_> + \mathcal{Z}_>) \right]$ looses only $2s$ derivatives thanks to the commutator structure since $\mathcal{M}_>$ is a diagonal matrix of para-differential operator of order one, while $\mathcal{Q}$, $\mathcal{Z}_>$ are (uo to smoothing remainders) para-differential of order zero.

By Cauchy–Schwarz inequality, (5.85), (5.128), (5.78), (5.79) and (5.130) we obtain (5.98).

5.3 Proof of Theorem 5.1

We are now in position to conclude the proof of Theorem 5.1. Consider the function $\varepsilon^\beta V$ which is a solution of (4.9). Our aim is to prove an *a priori* estimate on $\varepsilon^\beta \Pi_S^\perp V$ which solves the equation (5.1). Recall that, by Lemma 5.2, the (5.1) can be written as (5.4). By Proposition 5.11 (see also Lemma 5.9) we have that the function $W^\perp$ in (5.60) solves the equation (5.63). Now consider the modified energy $\mathcal{E}_s(W^\perp)$ in (5.95)–(5.96) (recall also (5.84)) where the operator $\mathcal{Q}(\varepsilon\varphi)$ is given by Proposition 5.14. We claim that there is $C_s > 0$ such that, for $\varepsilon > 0$ small enough,

$$\frac{1}{1 + \varepsilon C_s} \|W^\perp\|_{H^s}^2 \leq |\mathcal{E}_s(W^\perp)| \leq (1 + \varepsilon C_s)\|W^\perp\|_{H^s}^2 . \tag{5.131}$$

Indeed, using (5.99)–(5.102), recalling Definitions 2.3, 2.5, Lemma 2.2 and the Cauchy–Schwarz inequality, one gets

$$|\mathcal{E}_s(W^\perp)| \overset{(5.96)}{\leq} \|W^\perp\|_{H^s}^2 + |\tfrac{1}{2}(\mathrm{i}\, E\mathcal{D}^{2s}\mathcal{Q}(\varepsilon\varphi)W^\perp, W^\perp)_{\mathbf{H}^0}|$$
$$\leq \|W^\perp\|_{H^s}^2 + c_s\|\mathcal{Q}(\varepsilon\varphi)W^\perp\|_{H^s}\|W^\perp\|_{H^s}$$
$$\leq \|W^\perp\|_{H^s}^2 (1 + c_s\|\varepsilon\varphi\|_{H^s})$$

which implies the second in (5.131). The first inequality follows similarly. Using (5.131), estimate (5.98) and integrating in t one obtains

$$\|W^\perp(t)\|_{H^s}^2 \leq \|W^\perp(0)\|_{H^s}^2 + C_s T \sup_{t \in [0,T]} \|\varphi\|_{H^s}^3 \sup_{t \in [0,T]} \|V\|_{H^s}^2$$
$$+ C_s T \sup_{t \in [0,T]} \|V\|_{H^s}^3 + C_s T\varepsilon^5 \sup_{t \in [0,T]} \|V\|_{H^s} ,$$

for $t \in [0, T]$. Therefore, using the equivalence (5.61), one gets the bound (5.2). This concludes the proof.

6 The Estimates on the Remainder and Proof of the Main Result

Consider the function V in (4.8) which solves the problem (4.9) and set

$$0 < \sigma < 1/4 , \quad \beta \in (2 + 2\sigma, 3 - 2\sigma) . \tag{6.1}$$

The main result of this section is the following.

Proposition 6.1 (Main bootstrap) *Let σ, β as in (6.1). There exists $s_0 \gg 1$ such that for all $s \geq s_0$ there exist constants $c_0 = c_0(s) > 0$, $C_s > 0$ such that the following holds. Let $\varepsilon > 0$ satisfying (4.5) and let $\varepsilon\varphi$ be the approximate solution in (4.3) and satisfying (4.4). Let $V(t, x)$ be a solution of (4.9) with initial condition $V_0 \in H^s(\mathbb{T})$ defined for times $t \in [0, T_0]$ for some $T_0 > 0$. If*

$$\varepsilon^\beta \|V_0\|_{H^s} \leq \varepsilon^\beta , \quad \varepsilon^\beta \|\Pi_S V_0\|_{H^1} \leq \varepsilon^{\beta+1} , \tag{6.2}$$

then $V(t, x)$ extends over an interval $[0, \widetilde{T}]$ with

$$\widetilde{T} := c_0 \varepsilon^{-2-\sigma} , \tag{6.3}$$

and we have the following bound

$$\sup_{t \in [0,\widetilde{T}]} \varepsilon^\beta \|V(t)\|_{H^s} \leq 2\varepsilon^\beta . \tag{6.4}$$

The proof of Proposition 6.1 is based on a bootstrap argument. We define

$$T_* := \sup\left\{ T \geq 0 : \sup_{t \in [0,T]} \|\varepsilon^\beta V\|_{H^s} \leq 2\varepsilon^\beta \right\}. \tag{6.5}$$

By assumption of Proposition 6.1 we have that $T_* > 0$. We shall prove, by contradiction, that the supremum T_* cannot be smaller than $\widetilde{T}$ in (6.3).

Remark 6.2 The bootstrap assumption implies (4.11).

We split the equation (4.9) in the following way

$$\begin{cases} \varepsilon^\beta \partial_t \Pi_S V = \varepsilon^\beta \Pi_S dX_{\mathcal{H}}(\varepsilon\varphi)[V] + \varepsilon^{2\beta} \Pi_S \mathcal{Q}(\varepsilon\varphi)[V, V] + \Pi_S \mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi), \\ \varepsilon^\beta \partial_t \Pi_S^\perp V = \varepsilon^\beta \Pi_S^\perp dX_{\mathcal{H}}(\varepsilon\varphi)[V] + \varepsilon^{2\beta} \Pi_S^\perp \mathcal{Q}(\varepsilon\varphi)[V, V] + \Pi_S^\perp \mathrm{Res}_{\mathcal{H}}(\varepsilon\varphi). \end{cases} \tag{6.6}$$

The rest of the section is organized as follows: in Sect. 6.1 we show the improved bound

$$\varepsilon^\beta \sup_{t \in [0,T]} \|\Pi_S V\|_{H^s(\mathbb{T})} \leq \frac{3}{2}\varepsilon^\beta \tag{6.7}$$

as long as $T \leq \widetilde{T}$ in (6.3). Similarly in Sect. 6.2 we prove

$$\varepsilon^\beta \sup_{t \in [0,T]} \|\Pi_S^\perp V\|_{H^s(\mathbb{T})} \leq \frac{3}{2}\varepsilon^\beta \tag{6.8}$$

as long as $T \leq \widetilde{T}$. By continuity the bounds (6.7)–(6.8) imply that the supremum T_* in (6.5) should be larger than $\widetilde{T}$ in (6.3). Then we will conclude the proof of Proposition 6.1. From now on we consider (by contradiction) $T \leq T_* \leq \widetilde{T}$ in (6.5)–(6.3).

6.1 The Equation for Low Frequencies

We study the time evolution of the Sobolev norms

$$\frac{d}{dt}\|\varepsilon^\beta \Pi_S V\|_{H^s}^2 = \varepsilon^{2\beta}(\mathcal{D}^{2s}\partial_t \Pi_S V, \Pi_S V)_{L^2} + \varepsilon^{2\beta}(\mathcal{D}^{2s}\Pi_S V, \partial_t \Pi_S V)_{L^2}.$$

We use the first equation in (6.6) and we provide estimates for each term. The most delicate term is the one involving $dX_{\mathcal{H}}(\varepsilon\varphi)$. Recall the definition of $\mathcal{H}_{res}$ in (3.31).

Lemma 6.3 *We have*

$$\Pi_S dX_{\mathcal{H}}(\varepsilon\varphi)[V] = dX_{\mathcal{H}_{res}}(\varepsilon\varphi)[\Pi_S V] + \Pi_S dX_{\mathfrak{R}^{(\geq 6)}}(\varepsilon\varphi)[V].$$

Proof Notice that, since $\varepsilon\varphi$ is Fourier supported on S, we have

$$\Pi_S dX_{\mathcal{H}^{(n, \geq 2)}}(\varepsilon\varphi)[V] = 0, \qquad n = 3, 4, 5.$$

Hence the thesis follows.

Lemma 6.4 *Recall T_* in (6.5). We have for $0 < T \leq T_*$*

$$\varepsilon^{2\beta}\left| \int_0^T (\mathcal{D}^{2s}\Pi_S dX_{\mathfrak{R}^{(\geq 6)}}(\varepsilon\varphi)[V], \Pi_S V)_{L^2}\, dt \right| \lesssim_s \varepsilon^{4+2\beta}T.$$

Proof The estimate holds by item (iii) of Proposition 3.9, in particular the first bound in (3.19), (4.4) and using the Cauchy–Schwarz inequality.

In the following Lemma we give energy estimate for the term $\Pi_S dX_{\mathcal{H}_{res}}(\varepsilon\varphi)[V]$.

Lemma 6.5 *Recall T_* in (6.5) and let V be the solution of (6.6). We have, for $0 < T \le T_*$,*

$$\varepsilon^{2\beta}\left|\int_0^T \left((\mathcal{D}^{2s}\Pi_S dX_{\mathcal{H}_{res}}(\varepsilon\varphi) + [dX_{\mathcal{H}_{res}}(\varepsilon\varphi)]^* \Pi_S \mathcal{D}^{2s})[\Pi_S V], \Pi_S V\right)_{L^2} dt\right| \le \frac{1}{4}\varepsilon^{2\beta}.$$
(6.9)

Proof By (3.16) we have

$$\mathcal{H}_{res}(W) = H^{(2)}(W) + \mathcal{H}_{res}^{(4,0)}(W) = \sum_{j\in\mathbb{Z}} \Lambda(j)|w_j|^2 + \sum_{j,k\in S} C_{jk}|w_j|^2|w_k|^2,$$
(6.10)

with $C_{jk} = C_{kj} \in \mathbb{R}$, and hence $X_{\mathcal{H}_{res}}(W) = (X_{\mathcal{H}_{res}}^+(W), \overline{X_{\mathcal{H}_{res}}^+(W)})^T$ with

$$\left(X_{\mathcal{H}_{res}}^+(W)\right)_j = -i\partial_{\overline{w_j}}\mathcal{H}_{res}(W) = -i\Lambda(j)w_j - i\left(\sum_{k\in S} C_{jk}|w_k|^2\right)w_j, \qquad j \in S.$$

We define $B := \Pi_S dX_{\mathcal{H}_{res}}(\varepsilon\varphi)$. Then, for $j \in S$, we have

$$(B[\Pi_S V])_j = \sum_{k\in S} \begin{pmatrix} B_{j,+}^{k,+}v_k + B_{j,+}^{k,-}\overline{v_k} \\ B_{j,-}^{k,+}v_k + B_{j,-}^{k,-}\overline{v_k} \end{pmatrix}$$

$$= \begin{pmatrix} -i\Lambda(j)v_j - i\left(\sum_{k\in S} C_{jk}|\varepsilon\varphi_k|^2\right)v_j - 2i\varepsilon\left(\sum_{k\in S} C_{jk}\mathrm{Re}\,(v_k\overline{\varphi_k})\right)\varepsilon\varphi_j \\ i\Lambda(j)\overline{v_j} + i\left(\sum_{k\in S} C_{jk}|\varepsilon\varphi_k|^2\right)\overline{v_j} + 2i\varepsilon\left(\sum_{k\in S} C_{jk}\mathrm{Re}\,(v_k\overline{\varphi_k})\right)\varepsilon\overline{\varphi_j} \end{pmatrix}.$$
(6.11)

Therefore we have

$$B_{j,+}^{k,+} := \begin{cases} -i(\Lambda(j) + 2\varepsilon^2 C_{jj}|\varphi_j|^2) & j = k, \\ -2i\varepsilon^2 C_{jk}\overline{\varphi_k}\varphi_j & j \ne k, \end{cases} \qquad B_{j,+}^{k,-} := \begin{cases} -2i\varepsilon^2 C_{jj}\varphi_j^2 & j = k, \\ -2i\varepsilon^2 C_{jk}\varphi_k\varphi_j & j \ne k, \end{cases}$$

$$B_{j,-}^{k,-} := \begin{cases} i(\Lambda(j) + 2\varepsilon^2 C_{jj}|\varphi_j|^2) & j = k, \\ 2i\varepsilon^2 C_{jk}\varphi_k\overline{\varphi_j} & j \ne k, \end{cases} \qquad B_{j,-}^{k,+} := \begin{cases} 2i\varepsilon^2 C_{jj}\overline{\varphi_j}^2 & j = k, \\ 2i\varepsilon^2 C_{jk}\overline{\varphi_k}\overline{\varphi_j} & j \ne k. \end{cases}$$

We define

$$A := \mathcal{D}^{2s}B + B^*\mathcal{D}^{2s} = \mathcal{D}^{2s}\Pi_S dX_{\mathcal{H}_{res}}(\varepsilon\varphi) + [dX_{\mathcal{H}_{res}}(\varepsilon\varphi)]^* \Pi_S \mathcal{D}^{2s}.$$

By using the following formulas

$$(B^*)_{j,\sigma}^{k,\sigma} = \overline{B_{j,-\sigma}^{k,-\sigma}} = B_{j,\sigma}^{k,\sigma}, \qquad (B^*)_{j,\sigma}^{k,\sigma'} = \overline{B_{j,\sigma'}^{k,\sigma}} = B_{j,-\sigma'}^{k,-\sigma},$$

we have

$$(A[\Pi_S V])_j = \sum_{k\in S} \begin{pmatrix} (\langle j\rangle^{2s}B_{j,+}^{k,+} + \overline{B_{j,+}^{k,-}}\langle k\rangle^{2s})v_k + (\langle j\rangle^{2s}B_{j,+}^{k,-} + \overline{B_{j,-}^{k,+}}\langle k\rangle^{2s})\overline{v_k} \\ (\langle j\rangle^{2s}B_{j,-}^{k,+} + \overline{B_{j,+}^{k,-}}\langle k\rangle^{2s})v_k + (\langle j\rangle^{2s}B_{j,-}^{k,-} + \overline{B_{j,+}^{k,+}}\langle k\rangle^{2s})\overline{v_k} \end{pmatrix}.$$

By using that

$$\overline{B_{j,-}^{k,-}} = B_{j,+}^{k,+}, \qquad \overline{B_{j,-}^{k,+}} = B_{j,+}^{k,-},$$

we get

$$(\mathbb{A}[\Pi_S V], \Pi_S V)_{L^2} = \sum_{j,k \in S} (\langle j \rangle^{2s} + \langle k \rangle^{2s}) \left(\mathrm{B}_{j,+}^{k,+} v_k \overline{v_j} + \mathrm{B}_{j,+}^{k,-} \overline{v_k v_j} + \mathrm{B}_{j,-}^{k,+} v_k v_j + \mathrm{B}_{j,-}^{k,-} \overline{v_k} v_j \right).$$

By using that $C_{jk} = C_{kj} \in \mathbb{R}$ and that the set S is symmetric we have

$$\mathrm{B}_{j,\sigma}^{k,\sigma'} = \overline{\mathrm{B}_{j,-\sigma}^{k,-\sigma'}}, \qquad \mathrm{B}_{\sigma,j}^{\sigma,k} = -\mathrm{B}_{-\sigma,k}^{-\sigma,j}, \qquad \forall j, k \in S.$$

Therefore

$$\begin{aligned}
(\mathbb{A}[\Pi_S V], \Pi_S V)_{L^2} &= \sum_{j,k \in S} (\langle j \rangle^{2s} + \langle k \rangle^{2s}) \left(\mathrm{B}_{j,+}^{k,-} \overline{v_k v_j} + \overline{\mathrm{B}_{j,+}^{k,-}} v_k v_j \right) \\
&= 2i\varepsilon^2 \sum_{j,k \in S} (\langle j \rangle^{2s} + \langle k \rangle^{2s})(-C_{jk} \varphi_k \varphi_j \overline{v_k v_j} + C_{jk} \overline{\varphi_k \varphi_j} v_k v_j) \\
&= -8\varepsilon^2 \sum_{j,k \in S} (\langle j \rangle^{2s} + \langle k \rangle^{2s}) C_{jk} \mathrm{Re}\, (\overline{\varphi_k} v_k)\, \mathrm{Im}\, (\overline{\varphi_j}\, v_j).
\end{aligned}$$

We have

$$(\mathbb{A}[\Pi_S V], \Pi_S V)_{L^2} = 16 \sum_{j \in S} \left(\sum_{k \in S} C_{jk} \langle k \rangle^{2s} \mathrm{Re}\, (v_k \varepsilon \overline{\varphi_k}) \right) \mathrm{Im}\, (\varepsilon \varphi_j \overline{v_j}). \tag{6.12}$$

We now study the derivative of $\mathrm{Re}\, (v_p \varepsilon \overline{\varphi_p})$ with $p \in S$. One has

$$\frac{d}{dt} \mathrm{Re}\, (v_p \overline{\varphi}_p) = \dot{v}_p \varepsilon \overline{\varphi_p} + v_p \varepsilon \overline{\dot{\varphi}_p} + \overline{\dot{v}_p} \varepsilon \varphi_p + \overline{v_p} \varepsilon \dot{\varphi}_p. \tag{6.13}$$

Recall that in Sect. 3.3 we constructed $\varepsilon \varphi$ in such a way (see (3.30), (3.31))

$$\varepsilon \partial_t \varphi = X_{\mathcal{N}}(\varepsilon \varphi) \equiv X_{\mathcal{H}_{res}}(\varepsilon \varphi), \tag{6.14}$$

which in Fourier reads as

$$\dot{\varphi}_p = i\omega_p(\xi)\, \varphi_p, \qquad p \in S.$$

The first equation in (6.6) reads in Fourier, for $p \in S$, as[3]

$$\begin{aligned}
\dot{v}_p &= (\Pi_S dX_{\mathcal{H}}(\varepsilon \varphi)[V])_p^+ + \varepsilon^\beta (\Pi_S \mathcal{Q}(\varepsilon \varphi)[V, V])_p^+ + \varepsilon^{-\beta} (\Pi_S \mathrm{Res}_{\mathcal{H}}(\varepsilon \varphi))_p^+ \\
&= (\Pi_S dX_{\mathcal{H}_{res}}(\varepsilon \varphi)[V])_p^+ \\
&\quad + \underbrace{(\Pi_S dX_{\mathfrak{R}^{(\geq 6)}}(\varepsilon \varphi)[V])_p^+}_{I} + \varepsilon^\beta \underbrace{(\Pi_S \mathcal{Q}(\varepsilon \varphi)[V, V])_p^+}_{II} + \underbrace{\varepsilon^{-\beta} (\Pi_S \mathrm{Res}_{\mathcal{H}}(\varepsilon \varphi))_p^+}_{III}.
\end{aligned} \tag{6.15}$$

Notice that, by estimate (3.19) on $dX_{\mathfrak{R}^{(\geq 6)}}$, (4.4) on $\varepsilon \varphi$, and the bootstrap assumption on $\varepsilon^\beta V$, one gets $|I| \lesssim_s \varepsilon^4$. Recalling (4.30)–(4.31) and that we are projecting on a finite dimensional subspace of H^s, we get $|II| \lesssim_s \varepsilon^\beta$. Finally, by Lemma 4.1 on the residual, one gets $|III| \lesssim_s \varepsilon^{5-\beta}$. All the latter estimates are uniform in $t \in [0, T]$. By substituting (6.14),

---

[3] Here by $X^\pm$ we are denoting the first component of a vector field $X = (X^+, X^-)^T$ (see (2.11)), and recall that $\overline{X^-} = X^+$.

(6.15) into (6.13) we obtain

$$\frac{d}{dt}\mathrm{Re}\,(v_p\overline{\varphi}_p) = \dot{v}_p\varepsilon\overline{\varphi_p} + v_p\varepsilon\dot{\overline{\varphi_p}} + \overline{\dot{v}_p}\varepsilon\varphi_p + \overline{v_p}\varepsilon\dot{\varphi}_p$$

$$= 2\mathrm{Re}\big(\varepsilon\overline{\varphi_p}(\Pi_S dX_{\mathcal{H}_{res}}(\varepsilon\varphi)[V])_p^+ + \overline{v_p}\varepsilon\mathrm{i}\omega_p(\xi)\varphi_p\big) \qquad (6.16)$$

$$+ 2\mathrm{Re}\big(\varepsilon\overline{\varphi_p}(I + II + III)\big). \qquad (6.17)$$

Using the explicit form of $\Pi_S dX_{\mathcal{H}_{res}}(\varepsilon\varphi)[V]$ in (6.11) and recalling (3.34), one can check that (6.16)$\equiv 0$. On the other hand, we also have $|(6.17)| \lesssim_s \varepsilon(\varepsilon^4 + \varepsilon^\beta + \varepsilon^{5-\beta})$. Let us define

$$\mathrm{P}_p := \mathrm{Re}\,(v_p\varepsilon\overline{\varphi_p})(0), \qquad \mathrm{L}_p(t) := \mathrm{Re}\,(v_p\varepsilon\overline{\varphi_p})(t) - \mathrm{Re}\,(v_p\varepsilon\overline{\varphi_p})(0).$$

By the discussion above we deduce

$$\sup_{\tau\in[0,t]} |L_p(\tau)| \lesssim \varepsilon\big(\varepsilon^4 + \varepsilon^\beta + \varepsilon^{5-\beta}\big)t, \qquad \forall t \in [0, T].$$

Let us now write (6.12) as

$$(\mathbb{A}[\Pi_S V(t)], \Pi_S V(t))_{L^2} = 16\sum_{j\in S}\left(\sum_{k\in S} C_{jk}\langle k\rangle^{2s}(\mathrm{P}_k + L_k(t))\right)\mathrm{Im}\,(\varepsilon\varphi_j(t)\overline{v_j}(t)).$$

By the bootstrap assumption (6.5) and the bound (4.4) we get

$$\sup_{t\in[0,T]} |\mathrm{Im}\,(\varepsilon\varphi_j\overline{v_j})| \lesssim \varepsilon.$$

Setting $\delta := \max_{k\in S} \mathrm{P}_k$, one can note that

$$\delta \leq 2\|\Pi_S V_0\|_{H^1(\mathbb{T})}\varepsilon \overset{(6.2)}{\leq} 2\varepsilon^2.$$

Thus

$$\left|\int_0^T (\mathbb{A}[\Pi_S V(t)], \Pi_S V(t))_{L^2}\,dt\right| \lesssim_s \delta\varepsilon T + \varepsilon^2\big(\varepsilon^4 + \varepsilon^\beta + \varepsilon^{5-\beta}\big)T^2 \lesssim_s c_0, \qquad (6.18)$$

provided that

$$\delta\varepsilon T \lesssim \varepsilon^3 T \lesssim 1, \qquad \varepsilon^6 T^2 \lesssim 1, \qquad \varepsilon^{\beta+2}T^2 \lesssim 1, \qquad \varepsilon^{7-\beta}T^2 \lesssim 1.$$

Since we set $T \lesssim \varepsilon^{-2-\sigma}$ (see (6.3)), we need that

$$\varepsilon^{6-4-2\sigma} = \varepsilon^{2-2\sigma} \lesssim 1, \qquad \varepsilon^{\beta+2-4-2\sigma} = \varepsilon^{\beta-2-2\sigma} \lesssim 1, \qquad \varepsilon^{7-\beta-4-2\sigma} = \varepsilon^{3-\beta-2\sigma} \lesssim 1, \qquad (6.19)$$

which follows by taking σ and β as in (6.1). This implies the thesis.

Lemma 6.6 *Recall T_* in (6.5). We have, for $0 < T \leq T_*$,*

$$\varepsilon^{3\beta}\left|\int_0^T (\mathcal{D}^{2s}\Pi_S\mathcal{Q}(\varepsilon\varphi)[V], \Pi_S V)_{L^2}\,dt\right| \lesssim_s \varepsilon^{3\beta}T.$$

Proof It follows by Cauchy–Schwarz, recalling (4.30)–(4.31) and that we are projecting on a finite dimensional subspace of H^s.

Lemma 6.7 *Recall T_* in (6.5). We have, for $0 < T \leq T_*$,*

$$\varepsilon^\beta\left|\int_0^T (\mathcal{D}^{2s}\Pi_S Res_{\mathcal{H}}(\varepsilon\varphi), V)_{L^2}\,dt\right| \lesssim_s \varepsilon^{\beta+5}T.$$

Proof It follows by Lemma 4.1.

By collecting Lemmata 6.4, 6.5, 6.6, 6.7, we get

$$\varepsilon^{2\beta} \sup_{t\in[0,T]} \|\Pi_S V(t)\|_{H^s}^2 \leq \varepsilon^{2\beta} \|\Pi_S V_0\|_{H^s}^2$$
$$+ C(s)\varepsilon^{2\beta}\left(\varepsilon^4 T + 1 + \varepsilon^\beta T + \varepsilon^{5-\beta}T\right).$$

The bound (6.7) holds for $T \lesssim \varepsilon^{-2-\sigma}$ by (6.2) and (6.1).

6.2 The Equation for High Frequencies

Consider the second equation in (6.6), which is the (5.1), and take an initial condition V_0 satisfying the assumption of Proposition 6.1, and take the solution $V(t,x)$ evolving form V_0 and defined on a time interval $[0,T]$ with $T \leq T_*$ (see (6.5)). By Remark 6.2 and the assumptions of Proposition 6.1 we can apply Theorem 5.1. Then, by the bootstrap assumption on V and the bound (4.4) on $\varepsilon\varphi$, we have that estimate (5.2) implies

$$\varepsilon^{2\beta} \sup_{t\in[0,T]} \|\Pi_S^{\perp} V(t)\|_{H^s}^2 \leq (1 + C_s\varepsilon)\varepsilon^{2\beta} \|\Pi_S^{\perp} V(0)\|_{H^s}^2$$
$$+ C_s\varepsilon^{2\beta}\left(\varepsilon^3 T + \varepsilon^\beta T + T\,\varepsilon^{5-\beta}\right),$$

for some $C_s > 0$. If we take $T = c_0\varepsilon^{-2-\sigma}$, $0 < \sigma < 1/4$, then taking ε small enough and (6.1) we obtain the (6.8), provided that c_0 is small enough.

Proof of Proposition 6.1 By the discussion of Sects. 6.1, 6.2 we have that, if the bound (6.5) holds true, then one obtains the improved estimates (6.7) and (6.8). Then the time $T_* \leq c_0\varepsilon^{-2-\sigma}$ is not the maximal time for which the bound (6.4) holds. This is in contradiction with the definition (6.5). Then the thesis follows.

6.3 Proof of the Main Results

Proof of Theorem 1.1 Fix $N \in \mathbb{N}$ and consider a *symmetric* subset $S \subset \mathbb{Z}$ of cardinality N as in (4.1). Fix the mass m in the full measure set $\mathcal{M}$ given by Lemma 3.8, so that the normal form Proposition 3.9 applies. By the arguments in Sect. 4.1, for $(\xi, \theta) \in [1,2]^N \times \mathbb{T}^N$,[4] we construct an oscillating function $\varepsilon\varphi(t,x)$ of the form (4.3) supported on S, with frequency of oscillation $\omega_j(\xi)$, $j \in S$, in (3.34), where C_{jk} are the coefficients computed in (3.16). This also defines the linear operator $\mathsf{C} : \mathbb{R}^N \to \mathbb{R}^N$.

Now take any $V_0 \in H^s$ such that (6.2) holds with $\beta := 5/2$ and set

$$Z_0 = Z(0) = \Phi_B^{-1}\left(\varepsilon\varphi(0) + \varepsilon^\beta V_0\right). \tag{6.20}$$

We observe that $\varepsilon\varphi(0)$ is determined by the choice of (ξ, θ). Consider the solution $Z(t)$ of the Klein–Gordon equation (1.4) with initial condition $Z(0) = Z_0$ (which exists by local theory), and notice that, by the continuity of the map Φ_B^{-1}, Z_0 can be chosen in an open set of H^s.

---

[4] Here for simplicity we fix $\mathcal{O} = [1,2]$.

We shall prove that the bound (1.7) holds over a time interval $[0, T]$ with $T = \varepsilon^{-\frac{9}{4}+\delta}$ for any $\delta > 0$. We write

$$T = \frac{1}{\varepsilon^{9/4-\delta}} = \varepsilon^{-2-\sigma}\varepsilon^{\delta/2}, \quad \sigma := \frac{1}{4} - \frac{\delta}{2}.$$

By this choice of σ, we have

$$2 + 2\sigma = \frac{5}{2} - \delta, \qquad 3 - 2\sigma = \frac{5}{2} + \delta.$$

Then the choice $\beta = 5/2$ is so that the condition (6.1) is fulfilled and Proposition 6.1 applies. It guarantees that if V is a solution of (4.9) with initial condition $V(0) = V_0$ satisfying (6.2) then the bound (6.4) holds over a time interval $[0, \widetilde{T}]$ with $\widetilde{T} = c_0\varepsilon^{-2-\sigma}$ for some $c_0 = c_0(s)$ small. Notice that, taking ε small enough w.r.t. s, we also have that

$$\widetilde{T} = c_0\varepsilon^{-2-\sigma} \geq T = \varepsilon^{\delta/2}\varepsilon^{-2-\sigma} = \frac{1}{\varepsilon^{9/4-\delta}}.$$

The bound (6.4) on $\varepsilon^\beta V$, and estimate (4.4) imply that

$$\sup_{[0,T]} \|W\|_{H^s} = \sup_{[0,T]} \|\varepsilon\varphi + \varepsilon^\beta V\|_{H^s} \lesssim_s \varepsilon,$$

which, together with the estimates on the map Φ_B, guarantees that 4.2 holds true, so Proposition 4.2 applies. Estimate (4.12) implies (1.7). This concludes the proof.

Proof of Proposition 1.3 We define the map $\mathcal{F} := \Phi_B^{-1}$ where Φ_B is the Birkhoff map constructed in the proof of Theorem 1.1 (see Prop. 3.9).

By definition (recall (1.8), (1.9)) a function $u \in \mathcal{A}_{\varepsilon^{7/2}}^N$ can be written as

$$u = u_1 + u_2, \quad u_1 = \mathcal{F}(\varepsilon\varphi_{\xi,\theta}), \quad \|u_2\|_{H^s} \leq \frac{\varepsilon^{\beta+1}}{2}, \quad \beta := 5/2, \tag{6.21}$$

for some $(\xi, \theta) \in \mathcal{O}^N \times \mathbb{T}^N$. By the proof of Theorem 1.1 we know that $u \in \mathcal{U}_{\xi,\theta}^N$ if it is the image, under the map $\mathcal{F}$ (see (6.20)), of a function $\varepsilon\varphi_{\xi,\theta} + \varepsilon^\beta V_0$ for some V_0 satisfying (6.2). So to prove the inclusion (1.9) it is sufficient to show that there is some V_0 (satisfying (6.2)) such that

$$\mathcal{F}^{-1}(u_1 + u_2) = \mathcal{F}^{-1}(u) = \varepsilon\varphi_{\xi,\theta} + \varepsilon^\beta V_0 \quad \Leftrightarrow \quad u_1 + u_2 = u = \mathcal{F}(\varepsilon\varphi_{\xi,\theta} + \varepsilon^\beta V_0).$$

By (6.21) we have

$$\begin{aligned}
\mathcal{F}^{-1}(u) &= \mathcal{F}^{-1}\big(\mathcal{F}(\varepsilon\varphi_{\xi,\theta}) + u_2\big) \\
&= \varepsilon\varphi_{\xi,\theta} + u_2 + \mathcal{F}^{-1}\big(\mathcal{F}(\varepsilon\varphi_{\xi,\theta} + u_2) + R\big) - \mathcal{F}^{-1}\big(\mathcal{F}(\varepsilon\varphi_{\xi,\theta} + u_2)\big)
\end{aligned} \tag{6.22}$$

where

$$R := \mathcal{F}(\varepsilon\varphi_{\xi,\theta}) + u_2 - \mathcal{F}(\varepsilon\varphi_{\xi,\theta} + u_2) = u_2 - d\mathcal{F}(\varepsilon\varphi_{\xi,\theta} + \sigma u_2)[u_2],$$

for some $\sigma \in [0, 1]$. By estimates (3.13), (3.14) we get

$$\|R\|_{H^s} \leq C_s\varepsilon \|u_2\|_{H^s}, \tag{6.23}$$

taking $\varepsilon > 0$ small enough. Similarly, by (6.22), we can write

$$\mathcal{F}^{-1}(u) = \varepsilon\varphi_{\xi,\theta} + f, \quad f := u_2 + d\mathcal{F}^{-1}\big(\mathcal{F}(\varepsilon\varphi_{\xi,\theta} + u_2) + \tau R\big)[R],$$

for some $\tau \in [0, 1]$. Hence, again by (3.13), (3.14), we deduce that

$$\|f\|_{H^s} \leq \|u_2\|_{H^s} + C_2\varepsilon\|R\|_{H^s} \overset{(6.21),(6.23)}{<} \varepsilon^{\beta+1} , \qquad (6.24)$$

taking $\varepsilon > 0$ small enough. Notice that the function f depends only on the known functions $\varepsilon\varphi_{\xi,\theta}, u_2$. Setting $\varepsilon^{\beta}V_0 := f$ we have that condition (6.2) is fulfilled thanks to estimate (6.24). This implies the thesis.

Declarations

Ethical approval Not applicable.

Conflicts of interest The authors have no conflicts of interest to declare.

References

1. P. Baldi, M. Berti, R. Montalto, KAM for autonomous quasi-linear perturbations of KdV, Ann. I. H. Poincaré (C) Anal. Non Linéaire, 33(6):1589–1638 (2015)
2. Bambusi, D., Grébert, B.: Birkhoff normal form for partial differential equations with tame modulus. Duke Math. J. **135**(3), 507–567 (2006)
3. Bambusi, D., Delort, J.M., Grébert, B., Szeftel, J.: Almost global existence for Hamiltonian semi-linear Klein–Gordon equations with small Cauchy data on Zoll manifolds. Commun. Pure Appl. Math. **60**, 1665–1690 (2007)
4. Bernier, J., Grébert, B.: Long time dynamics for generalized Korteweg–de Vries and Benjamin–Ono equations. Arch. Rational Mech. Anal. **241**, 1139–1241 (2021)
5. Bernier, J., Faou, E., Grébert, B.: Rational normal forms and stability of small solutions to nonlinear Schrödinger equations. Ann. PDE **6**, 14 (2020)
6. Bernier, J., Feola, R., Grébert, B., Iandoli, F.: Long-time existence for semi-linear beam equations on irrational tori. J. Dyn. Differ. Equ. **33**(3), 1363–1398 (2021)
7. Berti, M., Biasco, L., Procesi, M.: KAM theory for the Hamiltonian derivative wave equation. Ann. Sci. Éc. Norm. Supér. (4), **46**(2):301–373 (2013)
8. Berti, M., Bolle, P.: Quasi-Periodic Solutions of Nonlinear Wave Equations on the d-Dimensional Torus, p. 2020. EMS Monographs in Mathematics. EMS Publishing House, Berlin (2020)

9. Berti, M., Delort, J.M.: Almost global solutions of capillary-gravity water waves equations on the circle. UMI Lecture Notes (2017)

10. M. Berti, A. Maspero, F. Murgante. Hamiltonian Birkhoff normal form for gravity-capillary water waves with constant vorticity: almost global existence. Preprint arXiv:2212.12255

11. Berti, M., Bolle, P.: Sobolev periodic solutions of nonlinear wave equations in higher spatial dimensions. Arch. Rational Mech. Anal. **195**, 609–642 (2010)

12. Berti, M., Biasco, L., Procesi, M.: KAM for reversible derivative wave equations. Arch. Ration. Mech. Anal. **212**(3), 905–955 (2014)

13. 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)

14. Berti, M., Feola, R., Pusateri, F.: Birkhoff normal form and long time existence for periodic gravity water waves. Commun. Pure Appl. Math. **76**(7), 1416–1494 (2023)

15. Biasco, L., Massetti, J.E., Procesi, M.: An abstract Birkhoff normal form theorem and exponential type stability of the 1d NLS. Commun. Math. Phys. **375**(3), 2089–2153 (2020)

16. Bobenko, A., Kuksin, S.: The nonlinear Klein–Gordon equation on an interval as a perturbed sine-Gordon equation. Comment. Math. Helv. **70**(1), 63–112 (1995)

17. Bourgain, J.: Construction of periodic solutions of nonlinear wave equations in higher dimension. GAFA **5**(4), 629–639 (1995)

18. A. Castro, D. Cordoba, F. Zheng. Stability of traveling waves for the Burgers–Hilbert equation. Analysis and PDE, Volume 16, Issue 9, p. 2109 (2023)

19. Chierchia, L., You, J.: KAM tori for 1D nonlinear wave equations with periodic boundary conditions. Commun. Math. Phys. **211**, 497–525 (2000)

20. Craig, W., Wayne, C.E.: Newton's method and periodic solutions of nonlinear wave equation. Commun. Pure Appl. Math. **46**, 1409–1498 (1993)

21. Delort, J.M., Szeftel, J.: Long-time existence for semi-linear Klein–Gordon equations with small Cauchy data on Zoll manifolds. Am. J. Math. **128** (2006)

22. Delort, J.M., Szeftel, J.: Long-time existence for small data nonlinear Klein–Gordon equations on tori and spheres. Int. Math. Res. Notices, **37** (2004)

23. Delort, J.-M.: A quasi-linear Birkhoff normal forms method. Application to the quasi-linear Klein–Gordon equation on $\mathbb{S}^1$. Astérisque, no. 341 (2012), 119 pp

24. Delort, J.-M.: Global existence and asymptotic behavior for the quasilinear Klein–Gordon equation with small data in dimension1. Ann. Sci. Ecole Norm. Sup., (4)34, no.1,1 – 61 (2001)

25. Delort, J.-M.: Quasi-Linear Perturbations of Hamiltonian Klein–Gordon Equations on Spheres. Memoirs of AMS, no. 234, 80 pp (2014)

26. Delort, J.M.: On long time existence for small solutions of semi-linear Klein–Gordon equations on the torus. J. d'Analyse Math. **107**, 161–194 (2009)

27. Düll, W.-P.: Justification of the nonlinear Schrödinger approximation for a quasilinear Klein–Gordon equation. Commun. Math. Phys. **355**, 1189–1207 (2017)

28. Düll, W.-P.: Validity of the nonlinear Schrödinger approximation for the two-dimensional water wave problem with and without surface tension in the arc length formulation. Arch. Ration. Mech. Anal. **239**, 1–84 (2021)

29. Düll, W.-P., Schneider, G., Wayne, E.: Justification of the nonlinear Schrödinger equation for the evolution of gravity driven 2D surface water waves in a canal of finite depth. Arch. Ration. Mech. Anal. **220**, 543–602 (2016)

30. Faou, E., Grébert, B.: Quasi invariant modified Sobolev norms for semi linear reversible PDEs. Nonlinearity **23**, 429–443 (2010)

31. Faou, E., Gauckler, L., Lubich, C.: Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus. Commun. Partial Differ. Equ. **38**(7), 1123–1140 (2013)

32. Feola, R., Giuliani, F.: Quasi-periodic traveling waves on an infinitely deep perfect fluid under gravity, ph accepted on Memoirs of AMS (2021)

33. Feola, R., Grebért, B., Iandoli, F.: Long time solutions for quasi-linear Hamiltonian perturbations of Schrödinger and Klein–Gordon equations on tori. ph Anal. PDEs, **16**(5):1133–1203 (2023)

34. Feola, R., Iandoli, F.: Long time existence for fully nonlinear NLS with small Cauchy data on the circle. Ann. Sci. Norm. Super. Pisa Cl. Sci., XXII(5):109–182 (2021)

35. Feola, R., Iandoli, F.: Local well-posedness for quasi-linear NLS with large Cauchy data on the circle. Ann. de l'Inst. Henri Poincaré C Anal. Non Linéaire **36**, 119–164 (2019)

36. Feola, R., Iandoli, F.: Local well-posedness for the Hamiltonian quasi-linear Schrödinger equation on tori. J. Math. Pures Appl. **157**, 243–281 (2022)

37. Feola, R., Montalto, R.: Quadratic lifespan and growth of Sobolev norms for derivative Schrödinger equations on generic tori. J. Differ. Equ. **312**, 276–316 (2022)
38. Feola, R., Giuliani, F., Procesi, M.: Reducibile KAM tori for the Degasperis Procesi equation. ph Commun. Math. Phys. **377**, 1681–1759 (2020)
39. Gentile, G., Mastropietro, V., Procesi, M.: Periodic solutions for completely resonant nonlinear wave equations with Dirichlet boundary conditions. Commun. Math. Phys. **256**(2), 437–490 (2005)
40. Gómez-Serrano, J., Ionescu, A.D., Park, J.: Quasiperiodic solutions of the generalized SQG equation, preprint arXiv:2303.03992 (2023)
41. Grébert, B., Paturel, E.: KAM for the Klein–Gordon equation on S^d. Bollettino dell'Unione Matematica Italiana (BUMI) **9**(2), 237–288 (2016)
42. Ifrim, M., Tataru, D.: The lifespan of small data solutions in two dimensional capillary water waves. Arch. Rational Mech. Anal. **225**(3), 1279–1346 (2017)
43. Ionescu, A.D., Pausader, B.: Global solutions of quasi linear systems of Klein–Gordon equations in 3D. J. Eur. Math. Soc. **16**(11), 2355–2431 (2014)
44. Ionescu, A.D., Pusateri, F.: Long-time existence for multi-dimensional periodic water waves. Geom. Funct. Anal. **29**(3), 811–870 (2019)
45. Kappeler, T., Montalto, R.: On the stability of periodic multi-solitons of the KdV equation. Commun. Math. Phys. **385**(3), 1871 (2021)
46. Klainerman, S.: Global existence of small amplitude solutions to nonlinear Klein–Gordon equations in four space-time dimensions. Commun. Pure Appl. Math. **38**(5), 631–641 (1985)
47. Maspero, A., Procesi, M.: Long time stability of small finite gap solutions of the cubic nonlinear Schrödinger equation on $\mathbb{T}^2$. J. Differ. Equ. **265**(7), 3212 (2018)
48. Métivier, G.: Para-Differential Calculus and Applications to the Cauchy Problem for Nonlinear Systems, vol. 5. Edizioni della Normale (2008)
49. Pöschel, J.: Quasi-periodic solutions for a nonlinear wave equation. Commentarii Math. Helvetici **71**, 269–296 (1996)
50. Schneider, G.: Approximation of the Korteweg–de Vries equation by the nonlinear Schrödinger equation. J. Differ. Equ. **147**(2), 333–354 (1998)
51. Schneider, G., Wayne, C.E.: Justification of the NLS approximation for a quasilinear water wave model. J. Differ. Equ. **251**, 238–269 (2011)
52. A. Stingo, Global existence and asymptotics for quasi-linear one-dimensional Klein–Gordon equations with mildly decaying Cauchy data. Bull. Soc. Math. France **146**(1), 155–213 (2018)
53. Taylor, M.: Tools for PDE. Am. Math. Soc. (2007)
54. Wang, W.M.: Quasi-periodic solutions for nonlinear wave equations. Comptes Rendus Math. **353**(7), 601–604 (2015)
55. Wayne, C.E.: Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory. Commun. Math. Phys. **127**, 479–528 (1990)
56. Wilson, B.: Sobolev stability of plane wave solutions to the nonlinear Schrödinger equation. Commun. Partial Differ. Equ. **40**(8), 1521–1542 (2015)