Rigorous derivation of the Whitham equations from the water waves equations in the shallow water regime

Louis Emerald

Rigorous derivation of the Whitham equations from the water waves equations in the shallow water regime

Louis Emerald

January 7, 2021

Abstract

We derive the Whitham equations from the water waves equations in the shallow water regime using two different methods, thus obtaining a direct and rigorous link between these two models. The first one is based on the construction of approximate Riemann invariants for a Whitham-Boussinesq system and is adapted to unidirectional waves. The second one is based on a generalisation of Birkhoff's normal form algorithm for almost smooth Hamiltonians and is adapted to bidirectional propagation. In both cases we clarify the improved accuracy on the fully dispersive Whitham model with respect to the long wave Korteweg-de Vries approximation.

1 Introduction

1.1 Motivations

In this paper we work on a unidirectionnal model for surface waves in coastal oceanography, the Whitham equation, introduced by Whitham in [20] and [21]. This equation, having the same dispersion relation as the general water waves model, is a full dispersion modification of the Korteweg-de Vries equation (KdV). As such one expects an improved accuracy and range of validity as a model for surface waves compared to the long wave KdV approximation. More precisely, we anticipate that in the presence of weak nonlinearities, the Whitham equation stays close to the water waves equations, even if the dispersion effects are not small, as obtained in [10] for the bidirectional full dispersion systems.

The Whitham model has weaker dispersive effects compared to the KdV model and thus allows for wavebreaking and Stokes waves of maximal amplitude to occur, as one would expect of a model for surface waves in coastal oceanography. This was the historical reason for which Whitham introduced it. A rigorous proof of the existence of the wavebreaking phenomenon has been established by Hur [11] and Saut & Wang [17]. Ehrnström & Wahlén [8] and Truong *et al.* [19] proved the existence of Stokes waves of maximal amplitude.

Besides [11, 17, 8, 19] we emphasize some other works on the Whitham equation. Klein *et al.* [12] compared rigorously the solutions of the Whitham equation with those of the KdV equation and exhibited three different regimes of behaviour: scattering for small initial data, finite time blow-up and a KdV long wave regime. Ehrnström & Wang [9] proved an enhanced existence time for solutions of the Whitham equation associated with small initial data. Ehrnström & Kalisch [7] proved the existence of periodic traveling-waves solutions. Sanford *et al.* [16] established that large-amplitude periodic traveling-wave solutions are unstable, while those of small-amplitude are stable if their wavelength is long enough. Denoting by μ the shallow water parameter and by ε the nonlinearity parameter, Klein *et al.* [12] proved the validity of the Whitham equation in the KdV regime $\{\mu = \varepsilon\}$ by obtaining that the Whitham equation approximate the KdV equation at a precision of order ε^2. Moldabayev *et al.* [15] derived the Whitham equation from the Hamiltonian of the water waves equations in the unusual regime $\{\varepsilon = e^{-\frac{\kappa}{\mu^{\nu/2}}}\}$, where κ and ν are parameters numerically inferred from the solitary waves of the Whitham model. They also assumed the initial data of the water waves equations prepared to generate unidirectional waves. If we transpose their derivation in the general shallow water regime $\{0 \leq \mu \leq \mu_{\max}, 0 \leq \varepsilon \leq 1\}$, where $\mu_{\max} > 0$ is fixed, we obtain a precision of order $\mu\varepsilon + \varepsilon^2$. To our point of view, the two previous results are not satisfying as the authors did not characterize the range of validity of the Whitham equation as a model for water waves with enough accuracy. In this work we propose to improve these results using two different methods.

The first method is based on an adaptation of the one used to derive the inviscid Burgers equations from the Nonlinear Shallow Water system using the Riemann invariants of the latter and requires, as in [15], that initial data are prepared to generate unidirectional waves. We show from it that the Whitham equation can be derived from the water waves equations at the order of precision $\mu\varepsilon$, which is the same order as for the Whitham-Boussinesq equations when considering general initial data (see [10] for a rigorous derivation of the latter equations in the shallow water regime). From this derivation, we prove that solutions of the water waves equations, associated with well-prepared initial data, can be approximated in the shallow water regime, to within order $\mu\varepsilon t$ in a time scale of order ε^{-1}, by a right or left propagating wave solving a Whitham-type equation.

The second method does not need well-prepared initial data. It is based on a generalization of Birkhoff's normal form algorithm for almost smooth functions (see Definition 3.14) developed by Bambusi in [1] to show that the solutions of the water waves equations can be approximated in the KdV regime, to within order ε in a time scale of order ε^{-1}, by two counter-propagating waves solving KdV-type equations, thus improving both the results obtained by Schneider and Wayne in [18], and those obtained by Bona, Colin and Lannes in [2]. If we transpose Bambusi's result in the shallow water regime, we get a precision of order $\varepsilon^2 + (\mu^2 + \mu\varepsilon + \varepsilon^2)t$ in a time scale of order $(\max(\mu, \varepsilon))^{-1}$, which tells us that the two KdV-type equations are not appropriate to approximate the water waves equations when μ is not small. Bambusi's method uses the local structure of the KdV equation. We adapt it to the nonlocal Whitham equation and prove that the solutions of the water waves equations can be approximated in the shallow water regime, to within order $\varepsilon^2 + (\mu\varepsilon + \varepsilon^2)t$ in a time scale of order $(\max(\mu, \varepsilon))^{-1}$, by two counter-propagating waves solving Whitham-type equations. In addition, we express explicitly the loss of regularity needed to derive the two uncoupled Whitham-type equations solved by the counter-propagating waves.

We now compare the results from the first method and Bambusi's method for the KdV and the Whitham equations. In the presence of strong nonlinearities ($\varepsilon \geq \mu$), we get from the first method a precision order μ in a time scale of order ε^{-1}, whereas Bambusi's method for the KdV equations and for the Whitham equations gives a precision of order $\mu + \varepsilon$ in the same time scale. The ε in the latter order of precision comes from the coupling effects between the two counter-propagating waves. Hence, if μ is small, we have a good approximation from the Whitham equation, and not from the KdV equation, of the solutions of the water waves equations associated with well-prepared initial data, even if ε is not small. In the presence of weak nonlinearities ($\varepsilon \leq \mu$), we get from the first method and Bambusi's method for the Whitham equations a precision of order ε in a time scale of order μ^{-1}, whereas Bambusi's method for the KdV equations gives a precision of order $\mu + \varepsilon$ in the same time scale. Hence, if ε is small, we have a good approximation from the Whitham equation, and not from the KdV equation, of the solutions of the water waves equations, even if μ is not small.

1.2 Main results

The starting point of our study is the water waves equations:

$$\begin{cases} \partial_t \zeta - \frac{1}{\mu}\mathcal{G}^\mu[\varepsilon\zeta]\psi = 0, \\ \partial_t \psi + \zeta + \frac{\varepsilon}{2}(\partial_x \psi)^2 - \frac{\mu\varepsilon}{2}\frac{(\frac{1}{\mu}\mathcal{G}^\mu[\varepsilon\zeta]\psi + \varepsilon\partial_x\zeta\cdot\partial_x\psi)^2}{1+\varepsilon^2\mu(\partial_x\zeta)^2} = 0. \end{cases} \tag{1.1}$$

Here

- The free surface elevation is the graph of ζ, which is a function of time t and horizontal space $x \in \mathbb{R}$.
- $\psi(t,x)$ is the trace at the surface of the velocity potential.
- $\mathcal{G}^\mu$ is the Dirichlet-Neumann operator defined later in Definition 1.4.

Every variable and function in (1.2) is compared with physical characteristic parameters of the same dimension. Among those are the characteristic water depth H_0, the characteristic wave amplitude a_{surf} and the characteristic wavelength L. From these comparisons appear two adimensional parameters of main importance:

- $\mu := \frac{H_0^2}{L^2}$: the shallow water parameter,
- $\varepsilon := \frac{a_{\text{surf}}}{H_0}$: the nonlinearity parameter.

We refer to [13] for details on the derivation of these equations.

In [22], Zakharov proved that the water waves system (1.1) enjoys an Hamiltonian formulation. Let H_{WW} be defined by

$$H_{\text{WW}} := \frac{1}{2}\int_{\mathbb{R}} \zeta^2 \,\mathrm{dx} + \frac{1}{2}\int_{\mathbb{R}} \psi \frac{1}{\mu}\mathcal{G}^\mu[\varepsilon\zeta]\psi \,\mathrm{dx}\,. \tag{1.2}$$

Then (1.1) is equivalent to the Hamilton equations

$$\begin{cases} \partial_t \zeta = \delta_\psi H_{\mathrm{WW}}, \\ \partial_t \psi = -\delta_\zeta H_{\mathrm{WW}}, \end{cases}$$

where δ_ζ and δ_ψ are functional derivatives.

Before giving the main definitions of this section, here is one assumption maintained throughout this paper.

Hypothesis 1.1. *A fundamental hypothesis is the lower boundedness by a positive constant of the water depth (non-cavitation assumption):*

$$\exists h_{\min} > 0, \forall x \in \mathbb{R}, \quad h := 1 + \varepsilon\zeta(t,x) \geq h_{\min}. \tag{1.3}$$

Moreover, we will always work in the shallow water regime which we define here.

Definition 1.2. *Let $\mu_{\max} > 0$, we define the shallow water regime $\mathcal{A} := \{(\mu, \varepsilon), \quad 0 \leq \mu \leq \mu_{\max}, \quad 0 \leq \varepsilon \leq 1\}$.*

In what follows we need some notations on the functional framework of this paper.

Notations 1.3.

- *For any $\alpha \geq 0$ we will respectively denote $H^\alpha(\mathbb{R}), W^{\alpha,1}(\mathbb{R})$ and $W^{\alpha,\infty}(\mathbb{R})$ the Sobolev spaces of order α in respectively $L^2(\mathbb{R}), L^1(\mathbb{R})$ and $L^\infty(\mathbb{R})$. We will denote their associated norms by $|\cdot|_{H^\alpha}, |\cdot|_{W^{\alpha,1}}$ and $|\cdot|_{W^{\alpha,\infty}}$.*
- *For any $\alpha \geq 0$ we will denote $\dot{H}^{\alpha+1}(\mathbb{R}) := \{f \in L^2_{\mathrm{loc}}(\mathbb{R}), \quad \partial_x f \in H^\alpha(\mathbb{R})\}$ and $\dot{W}^{\alpha+1,1}(\mathbb{R}) := \{f \in L^1_{\mathrm{loc}}(\mathbb{R}), \quad \partial_x f \in W^{\alpha,1}(\mathbb{R})\}$ the Beppo-Levi spaces of order α. Their associated seminorms are respectively $|\cdot|_{\dot{H}^{\alpha+1}} := |\partial_x(\cdot)|_{H^\alpha}$ and $|\cdot|_{\dot{W}^{\alpha+1,1}} := |\partial_x(\cdot)|_{W^{\alpha,1}}$.*

Now we define the Dirichlet-Neumann operator.

Definition 1.4. *For all ζ and ψ sufficiently smooth, let ϕ be the velocity potential solving the elliptic problem*

$$\begin{cases} (\mu\partial_x^2 + \partial_z^2)\phi = 0, \\ \phi|_{z=\varepsilon\zeta} = \psi, \quad \partial_z\phi|_{z=-1} = 0. \end{cases}$$

where $z \in (-1, \varepsilon\zeta)$ is the vertical space variable.

We define the Dirichlet-Neumann operator by the formula

$$\mathcal{G}^\mu[\varepsilon\zeta]\psi = \sqrt{1 + \varepsilon^2(\partial_x\zeta)^2}\,\partial_{\boldsymbol{n}^\mu}\phi|_{\varepsilon\zeta},$$

where $\boldsymbol{n}^\mu$ is the outward normal vector of the free surface $\varepsilon\zeta$. It depends on time and space.

$\mathcal{G}^\mu$ is linear in ψ and nonlinear in ζ. See [13] for a thorough study of this operator.

We also recall the definition of a Fourier multiplier.

Definition 1.5. *Let $u : \mathbb{R} \to \mathbb{R}$ be a tempered distribution, let $\widehat{u}$ be its Fourier transform. Let $F \in \mathcal{C}^\infty(\mathbb{R})$ be a smooth function such that there exists $m \in \mathbb{Z}$ for which $\forall \beta \geq 0$, $|\partial_\xi^\beta F(\xi)| \leq \widetilde{C}_\beta(1+|\xi|)^{m-|\beta|}$, where $\widetilde{C}_\beta > 0$ is a constant depending on β. Then the Fourier multiplier associated with F is denoted $\mathrm{F}(D)$ (denoted F when no confusion is possible) and defined by the formula:*

$$\widehat{\mathrm{F}(D)u}(\xi) = F(\xi)\widehat{u}(\xi).$$

The first method to rigorously justify the Whitham equations from the water waves equations gives the two following results.

Proposition 1.6. *Let $\mu_{\max} > 0$. Let also $\mathrm{F}_\mu = \sqrt{\frac{\tanh(\sqrt{\mu}|D|)}{\sqrt{\mu}|D|}}$ be a Fourier multiplier. There exists $n \in \mathbb{N}^*$ such that for any $(\mu, \varepsilon) \in \mathcal{A}$ and $\zeta \in L_t^\infty H_x^{\alpha+n}(\mathbb{R})$ and $\psi \in L_t^\infty \dot{H}_x^{\alpha+n}(\mathbb{R})$ solutions of the water waves equations* (1.1) *satisfying* (1.3)*, there exists $R_1, R_2 \in H^\alpha(\mathbb{R})$ uniformly bounded in (μ, ε) such that the quantities*

$$u^+ = \frac{\sqrt{h}-1}{\varepsilon} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v], \quad u^- = -\frac{\sqrt{h}-1}{\varepsilon} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v], \tag{1.4}$$

where $v = \mathrm{F}_\mu^2[\partial_x\psi]$, satisfy the equations

$$\begin{cases} \partial_t u^+ + (\varepsilon\frac{3u^+ + u^-}{2} + 1)\mathrm{F}_\mu\partial_x u^+ = \mu\varepsilon R_1, \\ \partial_t u^- + (\varepsilon\frac{u^+ + 3u^-}{2} - 1)\mathrm{F}_\mu\partial_x u^- = \mu\varepsilon R_2. \end{cases}$$

Remark 1.7. *The quantities u^+ and u^- can almost be seen as Riemann invariants of the Whitham-Boussinesq equations*

$$\begin{cases} \partial_t\zeta + \partial_x^2\psi + \varepsilon\partial_x(\zeta\partial_x\psi) + \varepsilon\zeta\partial_x^2\psi = 0, \\ \partial_t\mathrm{F}_\mu^2[\partial_x\psi] + \mathrm{F}_\mu^2[\partial_x\zeta] + \varepsilon\partial_x\psi\partial_x^2\psi = 0, \end{cases} \tag{1.5}$$

for which we know the rigorous derivation from the water waves equations (see [10]). This is the main idea to prove this result.

Using Theorem 4.16 and Theorem 4.18 of [13] (the Rayleigh-Taylor condition is always satisfied when the bottom is flat, see Subsection 4.3.5 in [13]) we get the following corollary.

Corollary 1.8. *For any $\alpha \geq 0$, there exists $n \in \mathbb{N}^*$ such that for any $(\mu, \varepsilon) \in \mathcal{A}$, we have what follows. Consider the Cauchy problem for the water wave problem* (1.1) *with initial conditions $(\zeta_0, \psi_0) \in H^{\alpha+n}(\mathbb{R}) \times \dot{H}^{\alpha+n+1}(\mathbb{R})$ satisfying the non-cavitation assumption* (1.3)*, and denote by (ζ, ψ) the corresponding solution. Let $(u_{e,0}^+, u_{e,0}^-)$ be defined by the formulas* (1.4) *applied to (ζ_0, ψ_0). There exists a unique solution (u_e^+, u_e^-) of the exact diagonalized system*

$$\begin{cases} \partial_t u_e^+ + (\varepsilon\frac{3u_e^+ + u_e^-}{2} + 1)\mathrm{F}_\mu\partial_x u_e^+ = 0, \\ \partial_t u_e^- + (\varepsilon\frac{u_e^+ + 3u_e^-}{2} - 1)\mathrm{F}_\mu\partial_x u_e^- = 0, \end{cases} \tag{1.6}$$

with initial conditions $(u_{e,0}^+, u_{e,0}^-)$, *which satisfy the following property: denote by* $(\zeta_{\rm c}, \psi_{\rm c})$ *the following quantities*

$$\zeta_{\rm c} = \frac{1}{\varepsilon}\Big[(\frac{\varepsilon}{2}(u_e^+ - u_e^-) + 1)^2 - 1\Big] \quad \text{and} \quad \psi_{\rm c} = \int_0^x {\rm F}_\mu^{-1}[u_e^+ + u_e^-]\,{\rm dx}, \tag{1.7}$$

then for all times $t \in [0, \frac{T}{\varepsilon}]$, *one has*

$$|(\zeta - \zeta_{\rm c}, \psi - \psi_{\rm c})|_{H^\alpha \times \dot{H}^{\alpha+1}} \leq \mu\varepsilon C t,$$

where $C, T^{-1} = C(\frac{1}{h_{\min}}, \mu_{\max}, |\zeta_0|_{H^{\alpha+n}}, |\psi_0|_{\dot{H}^{\alpha+n+1}})$.

Remark 1.9. *The existence and uniqueness of the solution* (u_e^+, u_e^-) *of the exact diagonalized system* (1.6) *with initial conditions* $(u_{e,0}^+, u_{e,0}^-)$ *on the required time scale is given by Proposition 2.9.*

To write the next result we need another notation.

Notation 1.10. *Let* $k \in \mathbb{N}$ *and* $l \in \mathbb{N}$. *A function* R *is said to be of order* $O(\mu^k \varepsilon^l)$, *denoted* $R = O(\mu^k\varepsilon^l)$, *if divided by* $\mu^k\varepsilon^l$ *this function is uniformly bounded with respect to* $(\mu, \varepsilon) \in \mathcal{A}$ *in the Sobolev norms* $|\cdot|_{H^\alpha}$, $\alpha \geq 0$.

Proposition 1.11. *With the same hypotheses and notations as in Proposition 1.6, if* $u^-(0) = O(\mu)$, *then there exists* $T_1 > 0$ *such that for all times* $t \in [0, \frac{T_1}{\epsilon}]$, $u^-(t) = O(\mu)$. *Moreover for these times,* u^+ *satisfies the Whitham equation up to a remainder term of order* $O(\mu\varepsilon)$, *i.e.*

$$\partial_t u^+ + {\rm F}_\mu[\partial_x u^+] + \frac{3\varepsilon}{2} u^+ \partial_x u^+ = O(\mu\epsilon).$$

If instead $u^+(0) = O(\mu)$, *then there exists* $T_2 > 0$ *such that for all times* $t \in [0, \frac{T_2}{\epsilon}]$, $u^+(t) = O(\mu)$. *Moreover for these times,* u^- *satisfies the counter propagating Whitham equation up to remainder term of order* $O(\mu\varepsilon)$, *i.e.*

$$\partial_t u^- - {\rm F}_\mu[\partial_x u^-] + \frac{3\varepsilon}{2} u^- \partial_x u^- = O(\mu\varepsilon). \tag{1.8}$$

Proposition 1.12. *In Corollary 1.8, if* $u_e^-(0) = O(\mu)$, *then one can replace* $\zeta_{\rm c}$ *and* $\psi_{\rm c}$ *by*

$$\zeta_{\rm Wh,+} = \frac{1}{\varepsilon}\Big[(\frac{\varepsilon}{2}u^+ + 1)^2 - 1\Big] \quad \text{and} \quad \psi_{\rm Wh,+} = \int_0^x {\rm F}_\mu^{-1}[u^+]\,{\rm dx},$$

where $u^+ \in \mathcal{C}([0, \frac{T}{\varepsilon}], H^\alpha(\mathbb{R}))$ *solves the exact Whitham equation*

$$\partial_t u^+ + {\rm F}_\mu[\partial_x u^+] + \frac{3\varepsilon}{2} u^+ \partial_x u^+ = 0,$$

and get for all times $t \in [0, \frac{T}{\varepsilon}]$,

$$|(\zeta - \zeta_{\rm Wh,+}, \psi - \psi_{\rm Wh,+})|_{H^\alpha \times \dot{H}^{\alpha+1}} \leq C(|u_e^-(0)|_{H^{\alpha+1}} + \mu\varepsilon t),$$

where $C, T^{-1} = C(\frac{1}{h_{\min}}, \mu_{\max}, |\zeta_0|_{H^{\alpha+n}}, |\psi_0|_{\dot{H}^{\alpha+n+1}})$.

If instead $u_e^+(0) = O(\mu)$, then one can replace $\zeta_{\rm c}$ and $\psi_{\rm c}$ by

$$\zeta_{\rm Wh,-} = \frac{1}{\varepsilon}\Big[(\frac{\varepsilon}{2}u^- - 1)^2 - 1\Big] \quad \text{and} \quad \psi_{\rm Wh,-} = \int_0^x {\rm F}_\mu^{-1}[u^-]\,{\rm dx},$$

where $u^- \in \mathcal{C}([0, \frac{T}{\varepsilon}], H^\alpha(\mathbb{R}))$ solves the exact counter-propagating Whitham equation

$$\partial_t u^- - {\rm F}_\mu[\partial_x u^-] + \frac{3\varepsilon}{2}u^-\partial_x u^- = 0,$$

and get for all times $t \in [0, \frac{T}{\varepsilon}]$,

$$|(\zeta - \zeta_{\rm Wh,-}, \psi - \psi_{\rm Wh,-})|_{H^\alpha \times \dot{H}^{\alpha+1}} \leq C(|u_e^+(0)|_{H^{\alpha+1}} + \mu\varepsilon t),$$

where $C, T^{-1} = C(\frac{1}{h_{\min}}, \mu_{\max}, |\zeta_0|_{H^{\alpha+n}}, |\psi_0|_{\dot{H}^{\alpha+n+1}})$.

The second method is based on a generalisation of Birkhoff's normal form algorithm developed by Bambusi in [1]. Its application to our case leads us to define some operator and transformations which we need to write our main results.

Definition 1.13.

- *Let $\alpha \geq 0$. We define $\mathcal{T}_{\rm I} : H^\alpha(\mathbb{R}) \times H^\alpha(\mathbb{R}) \to H^\alpha(\mathbb{R}) \times \dot{H}^{\alpha+1}(\mathbb{R})$ by the formula*

$$\mathcal{T}_{\rm I}(\zeta, v) = \begin{pmatrix} \zeta \\ \int_0^x v(y)\,{\rm dy} \end{pmatrix}.$$

- *Let $\alpha \geq 0$. We define $\mathcal{T}_{\rm D} : H^{\alpha+1}(\mathbb{R}) \times H^{\alpha+1}(\mathbb{R}) \to H^\alpha(\mathbb{R}) \times H^\alpha(\mathbb{R})$ by the formula*

$$\mathcal{T}_{\rm D}(r, s) = \begin{pmatrix} r + s \\ {\rm F}_\mu^{-1}[r - s] \end{pmatrix}.$$

- *We define $\partial^{-1} : L^1(\mathbb{R}) \to L^\infty(\mathbb{R})$ by the formula*

$$\partial^{-1}u(y) = \frac{1}{2}\int_{\mathbb{R}} {\rm sgn}(y - y_1)u(y_1)\,{\rm dy}_1\,.$$

- *Let $\alpha \geq 0$ be a positive integer. We define $\mathcal{T}_{\rm B} : W^{\alpha+1,1}(\mathbb{R}) \times W^{\alpha+1,1}(\mathbb{R}) \to W^{\alpha,1}(\mathbb{R}) \times W^{\alpha,1}(\mathbb{R})$ by the formula*

$$\mathcal{T}_{\rm B}(r, s) = \begin{pmatrix} r + \frac{\varepsilon}{4}\partial_x(r)\partial^{-1}(s) + \frac{\varepsilon}{4}rs + \frac{\varepsilon}{8}s^2 \\ s + \frac{\varepsilon}{4}\partial_x(s)\partial^{-1}(r) + \frac{\varepsilon}{4}rs + \frac{\varepsilon}{8}r^2 \end{pmatrix}.$$

Theorem 1.14. *Let $\alpha \geq 0$ be a positive integer and $(\mu, \varepsilon) \in \mathcal{A}$. Let $r, s \in \mathcal{C}^1([0, \frac{T}{\varepsilon}], W^{\alpha+7,1}(\mathbb{R}))$ be solutions of the Whitham equations*

$$\begin{cases} \partial_t r + {\rm F}_\mu[\partial_x r] + \frac{3\varepsilon}{2}r\partial_x r = 0, \\ \partial_t s - {\rm F}_\mu[\partial_x s] - \frac{3\varepsilon}{2}s\partial_x s = 0. \end{cases} \tag{1.9}$$

Let also H_{WW} *be the Hamiltonian of the water waves equations* (1.2) *and* $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ *the Poisson tensor associated with. Then the quantities*

$$\begin{pmatrix} \zeta_{\mathrm{Wh}} \\ \psi_{\mathrm{Wh}} \end{pmatrix} := \mathcal{T}_{\mathrm{I}}(\mathcal{T}_{\mathrm{D}}(\mathcal{T}_{\mathrm{B}}(r,s)), \tag{1.10}$$

satisfy

$$\partial_t \begin{pmatrix} \zeta_{\mathrm{Wh}} \\ \psi_{\mathrm{Wh}} \end{pmatrix} = J\nabla(H_{\mathrm{WW}})(\zeta_{\mathrm{Wh}}, \psi_{\mathrm{Wh}}) + (\mu\varepsilon + \varepsilon^2)R, \quad \forall t \in [0, \frac{T}{\varepsilon}],$$

where $|R|_{H^\alpha \times \dot{H}^{\alpha+1}} \leq C(\mu_{\max}, |r|_{W^{\alpha+7,1}}, |s|_{W^{\alpha+7,1}})$.

Remark 1.15.

- *The existence of* $r, s \in \mathcal{C}^1([0, \frac{T}{\varepsilon}], W^{\alpha+6,1}(\mathbb{R}))$ *solutions of system* (1.9) *is given by Lemma 3.23.*

- *The remainder* R *is uniformly bounded in* (μ, ε) *for times* $t \in [0, \frac{T}{\max(\mu,\varepsilon)}]$ *in* $H^\alpha(\mathbb{R}) \times \dot{H}^{\alpha+1}(\mathbb{R})$*. The particular time interval* $[0, \frac{T}{\max(\mu,\varepsilon)}]$ *comes from the estimates of* r *and* s *in* $W^{\alpha+7,1}(\mathbb{R})$ *given by Lemma 3.23. Indeed, due to the dispersive effects one has a loss of decrease of order* $O((1+\mu t)^\theta)$ *for any* $\theta > 1/2$*, in the Sobolev spaces* $W^{\beta,1}(\mathbb{R})$ *where* $\beta \geq 0$.

- *The second equation of* (1.9) *is different from* (1.8)*. One has to replace the definition of* u^- *in Proposition 1.6 by* $u^- = \frac{\sqrt{h}-1}{\varepsilon} - \frac{1}{2}\mathrm{F}_\mu^{-1}[v]$ *to get*

$$\partial_t u^- - \mathrm{F}_\mu[\partial_x u^-] - \frac{3\varepsilon}{2} u^- \partial_x u^- = \mu\varepsilon R,$$

  when setting $u^+(0) = O(\mu)$.

- *In Theorem 1.14 we express the loss of regularity needed to pass rigorously from the Whitham equations* (1.9) *to the water waves equations* (1.1).

Using again Theorem 4.16 and Theorem 4.18 of [13], we get the following corollary.

Corollary 1.16. *There exists* $n \in \mathbb{N}^*$ *such that for any* $\alpha \geq 0$ *and any* $(\mu, \varepsilon) \in \mathcal{A}$*, we have what follows. Consider the Cauchy problem for the water wave problem* (1.1) *with initial conditions* $(\zeta_0, \psi_0) \in W^{\alpha+n,1}(\mathbb{R}) \times \dot{W}^{\alpha+n+1,1}(\mathbb{R})$ *satisfying the non-cavitation assumption* (1.3)*, and denote by* (ζ, ψ) *the corresponding solution. There exists a solution* (r, s) *of the Hamiltonian system* (1.9) *which satisfies the following property: denote by* $(\zeta_{\mathrm{Wh}}, \psi_{\mathrm{Wh}})$ *the solution defined by* (1.10)*, then for all times* $t \in [0, \frac{T}{\max(\mu,\varepsilon)}]$*, one has*

$$|(\zeta - \zeta_{\mathrm{Wh}}, \psi - \psi_{\mathrm{Wh}})|_{H^\alpha \times \dot{H}^{\alpha+1}} \leq C(\varepsilon^2 + (\mu\varepsilon + \varepsilon^2)t),$$

where $C, T^{-1} = C(\frac{1}{h_{\min}}, \mu_{\max}, |\zeta_0|_{W^{\alpha+n}}, |\psi_0|_{\dot{W}^{\alpha+n+1}})$*. See* (3.23) *for the initial conditions* (r_0, s_0) *associated with* (r, s).

1.3 Outline

In section 2 we prove Proposition 1.6 and Proposition 1.11. In subsection 2.1 we focus on the proof of Proposition 1.6 using symbolic calculus to diagonalize system (1.5) (see remark 1.7). In subsection 2.2 we prove Proposition 1.11 using a local well-posedness result (see Proposition 2.9) and a stability result (see Proposition 2.13) on the diagonalized system (1.6). Then we prove Proposition 1.12.

The section 3 is dedicated to the proof of Theorem 1.14. In subsection 3.1 we do a first approximation of the water waves' Hamiltonian by the one of a specific Whitham-Boussinesq similar to (1.5) (see Proposition 3.2). Then we do a simple change of unknowns separating the waves into a right and left front (see Proposition 3.5). In subsection 3.2 we apply the generalised Birkhoff's algorithm for almost smooth functions to the Hamiltonian obtained in the previous subsection. It gives an explicit transformation (3.13) allowing to pass from the latter Hamiltonian to the one of a system composed of two decoupled Whitham equations (see Property 3.20 and remark 3.22) at the desired order of precision. In subsection 3.3 we prove the existence of solutions of the two decoupled Whitham equations in the suitable Sobolev spaces $W^{\alpha,1}(\mathbb{R})$ for every integer $\alpha \geq 0$ (see lemma 3.23). Then we prove an intermediate theorem (see Theorem 3.24) which describes the action of the transformation resulting from the generalised Birkhoff's algorithm on the two decoupled Whitham equations. At the end, we prove Theorem 1.14 and corollary 1.16.

2 Derivation of the Whitham equations from the Riemann invariants of a Whitham-Boussinesq system

The goal of this section is to prove the Propositions 1.6 and 1.11.

We consider the following Whitham-Boussinesq equations.

$$\begin{cases} \partial_t \zeta + \partial_x v + \varepsilon \partial_x(\zeta) v + \varepsilon \zeta \partial_x v = 0, \\ \partial_t v + \mathrm{F}_\mu^2[\partial_x \zeta] + \varepsilon v \partial_x v = 0, \end{cases} \tag{2.1}$$

where $v = \mathrm{F}_\mu^2[\partial_x \psi] = \frac{\tanh(\sqrt{\mu}|D|)}{\sqrt{\mu}|D|}[\partial_x \psi]$. Using Proposition 1.15 of [10] we easily get the following proposition.

Proposition 2.1. *Let $\mu_{\max} > 0$. There exists $n \in \mathbb{N}^*$ and $T > 0$ such that for all $\alpha \geq 0$ and $p = (\mu, \varepsilon) \in \mathcal{A}$ (see Definition 1.2), and for every solution $(\zeta, \psi) \in C([0, \frac{T}{\varepsilon}]; H^{\alpha+n}(\mathbb{R}) \times \dot{H}^{\alpha+n+1}(\mathbb{R}))$ to the water waves equations* (1.1) *one has*

$$\begin{cases} \partial_t \zeta + \partial_x v + \varepsilon \partial_x(\zeta v) = \mu \varepsilon R_1, \\ \partial_t v + \mathrm{F}_\mu^2[\partial_x \zeta] + \varepsilon v \partial_x v = \mu \varepsilon R_2, \end{cases} \tag{2.2}$$

with $|R_1|_{H^\alpha}, |R_2|_{H^\alpha} \leq C(\frac{1}{h_{\min}}, \mu_{\max}, |\zeta|_{H^{\alpha+n}}, |\psi|_{\dot{H}^{\alpha+n+1}})$.

We say that the water waves equations are consistent with the Whitham-Boussinesq equations (2.1) *at precision order $O(\mu\varepsilon)$.*

Remark 2.2. *We only chose a Whitham-Boussinesq system fitted for the computations. There exists a whole class of these systems. For example, one can add a regularizing Fourier multiplier such as* F_μ^2 *on the nonlinear terms of system* (2.1) *and still be precise at order* $O(\mu\varepsilon)$*. The system thus obtained is*

$$\begin{cases} \partial_t\zeta + \partial_x v + \varepsilon \mathrm{F}_\mu^2 \partial_x[\zeta v] = 0, \\ \partial_t v + \mathrm{F}_\mu^2[\partial_x\zeta] + \frac{\varepsilon}{2}\mathrm{F}_\mu^2[v\partial_x v] = 0. \end{cases} \tag{2.3}$$

In [5] the authors proved the local well-posedness of (2.3)*. The reasoning of this section is the same for any Whitham-Boussinesq system such as the latter.*

The system (2.1) is the starting point of our reasoning. We will adapt the method used to derive the inviscid Burgers equations from the Nonlinear Shallow Water equations which use the Riemann invariants of the latter.

Throughout this section we will use Notation 1.10.

2.1 Diagonalization of the Whitham-Boussinesq equations

In this subsection we formally diagonalize the system (2.1) to get Proposition 1.6. Because of the consistency of the water waves equations with the system (2.1), we can take ζ and v solutions of the latter instead of taking solutions of the water waves equations. For our purpose, we just need to prove the following proposition.

Proposition 2.3. *Let* $\mu_{\max} > 0$*, there exists* $n \in \mathbb{N}$ *such that for any* $\alpha \geq 0$ *and* $(\zeta, v) \in \mathcal{C}([0, \frac{T}{\varepsilon}]; H^{\alpha+n}(\mathbb{R}) \times H^{\alpha+n}(\mathbb{R}))$ *solutions of system* (2.1) *satisfying the non-cavitation hypothesis* (1.3)*, the quantities*

$$u^+ = 2\frac{\sqrt{h}-1}{\varepsilon} + \mathrm{F}_\mu^{-1}[v], \quad u^- = -2\frac{\sqrt{h}-1}{\varepsilon} + \mathrm{F}_\mu^{-1}[v], \tag{2.4}$$

where $h = 1 + \varepsilon\zeta$*, satisfy*

$$\begin{cases} \partial_t u^+ + (\varepsilon\frac{3u^+ + u^-}{4} + 1)\mathrm{F}_\mu \partial_x u^+ = O(\mu\varepsilon), \\ \partial_t u^- + (\varepsilon\frac{u^+ + 3u^-}{4} - 1)\mathrm{F}_\mu \partial_x u^- = O(\mu\varepsilon). \end{cases} \tag{2.5}$$

where $\mathrm{F}_\mu = \sqrt{\frac{\tanh(\sqrt{\mu}|D|)}{\sqrt{\mu}|D|}}$.

Before proving this result, we recall some definition and classical results we need.

Definition 2.4. *Let* $f(x) \in \mathcal{C}^n(\mathbb{R})$ *with* $n \in \mathbb{N}^*$ *be such that for any* $0 \leq \beta \leq n$, $|\partial_x^\beta f(x)|_{L^\infty} \leq C_\beta$*, where* $C_\beta > 0$ *is a constant depending on* β*. Let also* $u(\xi) \in \mathcal{C}^\infty(\mathbb{R})$ *be a smooth function such that there exists* $m \in \mathbb{Z}$ *for which* $\forall\beta \geq 0$, $|\partial_\xi^\beta u(\xi)| \leq \widetilde{C}_\beta(1+|\xi|)^{m-|\beta|}$*, where* $\widetilde{C}_\beta > 0$ *is a constant depending on* β*.*

Then we define the operator $Op(fu)$ *as follow: for any sufficiently regular functions* v *we have*

$$Op(fu)v(x) := f(x)\mathrm{u}(D)[v](x),$$

where $\mathrm{u}(\mathrm{D})$ *is the Fourier multiplier associated with the function* u*.*

Remark 2.5. *The function* $F_\mu(\xi) = \sqrt{\frac{\tanh(\sqrt{\mu}|\xi|)}{\sqrt{\mu}|\xi|}}$ *defining the Fourier multiplier* F_μ *satisfies the condition in the previous Definition 2.4. Idem for its inverse or its square.*

Lemma 2.6. *Let* $\mu_{\max} > 0$. *Let* $(\varepsilon, \mu) \in \mathcal{A} = \{(\varepsilon, \mu), \quad 0 \le \varepsilon \le 1, \quad 0 \le \mu \le \mu_{\max}\}$. *Let* G_μ *be a Fourier multiplier for which there exists* $n \in \mathbb{N}^*$ *such that for any* $\beta \ge 0$ *and* $u \in H^{\beta+n}(\mathbb{R})$, $|(\mathrm{G}_\mu - 1)[u]|_{H^\beta} \lesssim \mu |u|_{H^{\beta+n}}$.

- *Let* $\alpha \ge 0$. *Let also* f *be a function in* $H^{\alpha+n}(\mathbb{R})$. *Then for any* $u \in H^{\alpha+n}(\mathbb{R})$, *the commutator* $[\mathrm{G}_\mu, f]u := \mathrm{G}_\mu[fu] - f\mathrm{G}_\mu[u]$ *satisfies*
$$|[\mathrm{G}_\mu, f]u|_{H^\alpha} = |[\mathrm{G}_\mu - 1, f]u|_{H^\alpha} \lesssim \mu |f|_{H^{\alpha+n}} |u|_{H^{\alpha+n}}.$$
- *Let* $\alpha \ge 0$. *Let also* g *be a function such that* $\frac{g-1}{\varepsilon}$ *is uniformly bounded in* ε *in* $H^{\alpha+n}(\mathbb{R})$. *Then for any function* $u \in H^{\alpha+n}(\mathbb{R})$,
$$|[\mathrm{G}_\mu, g]u|_{H^\alpha} = |[\mathrm{G}_\mu - 1, g - 1]u|_{H^\alpha} \lesssim \mu\varepsilon |\frac{g-1}{\varepsilon}|_{H^{\alpha+n}} |u|_{H^{\alpha+n}}.$$

Proof. Use product estimates A.1. □

Remark 2.7. *The Fourier multiplier* F_μ *satisfies the assumption of Lemma 2.6. Idem for its inverse or its square. The functions* h, $\sqrt{h}$ *and its inverse satisfy the assumption on* g.

Proof. The point on the Fourier multipliers F_μ, its inverse and its square is obvious. Idem for the function h. To prove the point on $\sqrt{h}$, we just need to use composition estimates A.2 with $G(x) = \sqrt{1+x} - 1$.

We now deal with $\frac{1}{\sqrt{h}}$. We remark that $\frac{1}{\sqrt{h}} - 1 = \frac{1-\sqrt{h}}{1+\sqrt{h}-1}$. Then we use quotient estimates A.3 ($\sqrt{h} \ge \sqrt{h_{\min}}$ because ζ satisfies the non-cavitation hypothesis (1.3)). We get
$$|\frac{1}{\sqrt{h}} - 1|_{H^\alpha} \le C(\frac{1}{\sqrt{h_{\min}}}, |\sqrt{h} - 1|_{H^{\max(t_0,\alpha)}}) |\sqrt{h} - 1|_{H^\alpha}.$$
□

We now prove Proposition 2.3.

Proof. We can write system (2.1) under the form
$$\partial_t U + A(U)\partial_x U = 0,$$
where $U := \begin{pmatrix} \zeta \\ v \end{pmatrix}$ and $A(U) := \begin{pmatrix} \varepsilon v & h \\ \mathrm{F}_\mu^2[\circ] & \varepsilon v \end{pmatrix}$, with $h := 1 + \varepsilon\zeta$.

The symbol of operator $A(U)$ is $A(U, \xi) = \begin{pmatrix} \varepsilon v & h \\ F_\mu^2(\xi) & \varepsilon v \end{pmatrix}$, $\xi \in \mathbb{R}$ being the frequency variable. So that for any smooth functions $W = (w_1, w_2)^T$, for all $x \in \mathbb{R}$ we have
$$A(U)W(t,x) = \mathrm{Op}(A(U,\xi))W(t,x) = \begin{pmatrix} \varepsilon v(t,x) w_1(t,x) + h(t,x) w_2(t,x) \\ \mathrm{F}_\mu^2[w_1](t,x) + \varepsilon v(t,x) w_2(t,x) \end{pmatrix}.$$

Now, we diagonalize $A(U,\xi)$.

$$\det(A(U,\xi)-\lambda I_2)=(\varepsilon v-\lambda)^2-F_\mu^2(\xi)h,$$

so that

$$A(U,\xi)=PDP^{-1},$$

with

$$P=\begin{pmatrix}\frac{\sqrt{h}}{F_\mu(\xi)} & -\frac{\sqrt{h}}{F_\mu(\xi)}\\ 1 & 1\end{pmatrix},\quad P^{-1}=\begin{pmatrix}\frac{F_\mu(\xi)}{2\sqrt{h}} & \frac{1}{2}\\ -\frac{F_\mu(\xi)}{2\sqrt{h}} & \frac{1}{2}\end{pmatrix}\quad\text{and}\quad D=\begin{pmatrix}\varepsilon v+F_\mu(\xi)\sqrt{h} & 0\\ 0 & \varepsilon v-F_\mu(\xi)\sqrt{h}\end{pmatrix}.$$

Based on the symbolic calculus we apply $\mathrm{F}_\mu^{-1}\operatorname{Op}(P^{-1})$ to the system:

$$\begin{aligned}&\mathrm{F}_\mu^{-1}\operatorname{Op}(P^{-1})\partial_t U+\mathrm{F}_\mu^{-1}\operatorname{Op}(P^{-1})\operatorname{Op}(A(U,\xi))\operatorname{Op}(PP^{-1})\partial_x U=0,\\ &\qquad\iff \mathrm{F}_\mu^{-1}\operatorname{Op}(P^{-1})\partial_t U+\mathrm{F}_\mu^{-1}\operatorname{Op}(D)\mathrm{F}_\mu\mathrm{F}_\mu^{-1}\operatorname{Op}(P^{-1})\partial_x U=R_1,\end{aligned}\tag{2.6}$$

where

$$\begin{aligned}R_1=&\mathrm{F}_\mu^{-1}(\operatorname{Op}(P^{-1}A(U,\xi))-\operatorname{Op}(P^{-1})\operatorname{Op}(A(U,\xi)))\partial_x U\\ &+\mathrm{F}_\mu^{-1}\operatorname{Op}(P^{-1}A(U,\xi))(\operatorname{Op}(P)\operatorname{Op}(P^{-1})-\operatorname{Op}(PP^{-1}))\partial_x U\\ &+\mathrm{F}_\mu^{-1}(\operatorname{Op}(D)-\operatorname{Op}(P^{-1}A(U,\xi))\operatorname{Op}(P))\operatorname{Op}(P^{-1})\partial_x U.\end{aligned}$$

We remark that for each operators F_μ^{-1}, ∂_x, $\operatorname{Op}(P^{-1}A(U,\xi))$ and $\operatorname{Op}(P^{-1})$ there exists $n\in\mathbb{N}$ such that for any $\beta\geq 0$, the operator is bounded from $H^{\beta+n}(\mathbb{R})$ to $H^\beta(\mathbb{R})$.

Lemma 2.8. *R_1 is of order $O(\mu\varepsilon)$.*

Proof. There are three terms in R_1. We deal with each of those separately.

- First term: for any $W=(w_1,w_2)\in H^\beta(\mathbb{R})\times H^\beta(\mathbb{R})$ with $\beta\geq 0$,

$$(\operatorname{Op}(P^{-1}A(U,\xi))-\operatorname{Op}(P^{-1})\operatorname{Op}(A(U,\xi)))W=\frac{1}{2\sqrt{h}}\begin{pmatrix}-[\mathrm{F}_\mu,\varepsilon v]w_1-[\mathrm{F}_\mu,h]w_2\\ [\mathrm{F}_\mu,\varepsilon v]w_1+[\mathrm{F}_\mu,h]w_2\end{pmatrix}.$$

  Using Lemma 2.6 and $h\geq h_{\min}$, we have

$$\frac{1}{2\sqrt{h}}[\mathrm{F}_\mu,\varepsilon v]w_1=O(\mu\varepsilon)\quad\text{and}\quad\frac{1}{2\sqrt{h}}[\mathrm{F}_\mu,h]w_2=O(\mu\varepsilon).$$

- Second term: for any $W=(w_1,w_2)\in H^\beta(\mathbb{R})\times H^\beta(\mathbb{R})$ with $\beta\geq 0$,

$$(\operatorname{Op}(P)\operatorname{Op}(P^{-1})-I_d)W=\begin{pmatrix}\sqrt{h}[\mathrm{F}_\mu^{-1},\frac{1}{\sqrt{h}}]\mathrm{F}_\mu[w_1]\\ w_2\end{pmatrix}.$$

  Using Lemma 2.6 and the boundedness of F_μ and $\sqrt{h}$ we have

$$\sqrt{h}[\mathrm{F}_\mu^{-1},\frac{1}{\sqrt{h}}]\mathrm{F}_\mu[w_1]=O(\mu\varepsilon).$$

- Third term: for any $W = (w_1, w_2) \in H^\beta(\mathbb{R}) \times H^\beta(\mathbb{R})$ with $\beta \geq 0$,

$$(\mathrm{Op}(D) - \mathrm{Op}(P^{-1}A(U,\xi))\,\mathrm{Op}(P))W = \begin{pmatrix} -\frac{\varepsilon v}{2\sqrt{h}}[\mathrm{F}_\mu - 1, \sqrt{h}]\mathrm{F}_\mu^{-1}[w_1 - w_2] - \frac{1}{2}[\mathrm{F}_\mu^2, \sqrt{h}]\mathrm{F}_\mu^{-1}[w_1 - w_2] \\ \frac{\varepsilon v}{2\sqrt{h}}[\mathrm{F}_\mu - 1, \sqrt{h}]\mathrm{F}_\mu^{-1}[w_1 - w_2] - \frac{1}{2}[\mathrm{F}_\mu^2, \sqrt{h}]\mathrm{F}_\mu^{-1}[w_1 - w_2] \end{pmatrix}$$

Using Lemma 2.6, $h \geq h_{\min}$ and the boundedness of F_μ^{-1}, we have

$$\begin{cases} \frac{\varepsilon v}{2\sqrt{h}}[\mathrm{F}_\mu - 1, \sqrt{h}]\mathrm{F}_\mu^{-1}[w_1 - w_2] = O(\mu\varepsilon), \\ \frac{1}{2}[\mathrm{F}_\mu^2, \sqrt{h}]\mathrm{F}_\mu^{-1}[w_1 - w_2] = O(\mu\varepsilon). \end{cases}$$

□

We continue the proof of Proposition 2.3. We compute $\mathrm{F}_\mu^{-1}\,\mathrm{Op}(P^{-1})\partial_t U$:

$$\mathrm{F}_\mu^{-1}\,\mathrm{Op}(P^{-1})\partial_t U = \mathrm{F}_\mu^{-1}\begin{pmatrix} \frac{1}{2\sqrt{h}}\mathrm{F}_\mu[\partial_t\zeta] + \frac{\partial_t v}{2} \\ -\frac{1}{2\sqrt{h}}\mathrm{F}_\mu[\partial_t\zeta] + \frac{\partial_t v}{2} \end{pmatrix} = \begin{pmatrix} \partial_t(\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) \\ \partial_t(-\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) \end{pmatrix} + R_2,$$

where $R_2 = \mathrm{F}_\mu^{-1}\begin{pmatrix} -[\mathrm{F}_\mu, \frac{1}{2\sqrt{h}}]\partial_t\zeta \\ [\mathrm{F}_\mu, \frac{1}{2\sqrt{h}}]\partial_t\zeta \end{pmatrix}$. Here, to prove that R_2 is of order $O(\mu\varepsilon)$, in addition of Lemma 2.6 we need a control of $|\partial_t\zeta|_{H^\beta}$ for any $\beta \geq 0$. We get it using the first equation of (2.1) and product estimates A.1.

The same computations for $\mathrm{F}_\mu^{-1}\,\mathrm{Op}(P^{-1})\partial_x U$ gives

$$\mathrm{F}_\mu^{-1}\,\mathrm{Op}(P^{-1})\partial_x U = \begin{pmatrix} \partial_x(\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) \\ \partial_x(-\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) \end{pmatrix} + R_3,$$

where $R_3 = \mathrm{F}_\mu^{-1}\begin{pmatrix} -[\mathrm{F}_\mu, \frac{1}{2\sqrt{h}}]\partial_x\zeta \\ [\mathrm{F}_\mu, \frac{1}{2\sqrt{h}}]\partial_x\zeta \end{pmatrix}$ is of order $O(\mu\varepsilon)$.

So that

$$\begin{aligned}\mathrm{F}_\mu^{-1}\,\mathrm{Op}(D)\mathrm{F}_\mu\mathrm{F}_\mu^{-1}\,\mathrm{Op}(P^{-1})\partial_x U &= \begin{pmatrix} \varepsilon\mathrm{F}_\mu^{-1}[v\mathrm{F}_\mu[\circ]] + \mathrm{F}_\mu^{-1}[\sqrt{h}\mathrm{F}_\mu^2[\circ]] & 0 \\ 0 & \varepsilon\mathrm{F}_\mu^{-1}[v\mathrm{F}_\mu[\circ]] - \mathrm{F}_\mu^{-1}[\sqrt{h}\mathrm{F}_\mu^2[\circ]] \end{pmatrix} \\ &\quad\times \begin{pmatrix} \partial_x(\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) \\ \partial_x(-\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) \end{pmatrix} + R_4,\end{aligned}$$

where $R_4 = \mathrm{F}_\mu^{-1}\,\mathrm{Op}(D)\mathrm{F}_\mu[R_3] = O(\mu\varepsilon)$ because of the boundedness of the operators $\mathrm{F}_\mu^{-1}, \mathrm{Op}(D)$ and F_μ.

Hence, we get

$$\begin{cases} \partial_t(\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) + \big(\varepsilon\mathrm{F}_\mu^{-1}[v\circ] + \mathrm{F}_\mu^{-1}[\sqrt{h}\mathrm{F}_\mu[\circ]]\big)\mathrm{F}_\mu\partial_x(\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) = R_5, \\ \partial_t(-\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) + \big(\varepsilon\mathrm{F}_\mu^{-1}[v\circ] - \mathrm{F}_\mu^{-1}[\sqrt{h}\mathrm{F}_\mu[\circ]]\big)\mathrm{F}_\mu\partial_x(-\frac{1}{\varepsilon}\sqrt{h} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) = R_6, \end{cases}$$

where R_5 and R_6 are combinations of R_1, R_2, R_3, R_4. They are of order $O(\mu\varepsilon)$. See also that $\frac{1}{\epsilon}\partial_x\sqrt{h}$ and $\frac{1}{\epsilon}\partial_x\sqrt{h}$ are of order $O(1)$

But by Lemma 2.6, we know that for any $w \in H^\beta(\mathbb{R})$ with $\beta \geq 0$, we have

$$[\mathrm{F}_\mu, \sqrt{h}]w = O(\mu\varepsilon) \quad \text{and} \quad [\mathrm{F}_\mu^{-1}, v]w = O(\mu).$$

So

$$\begin{cases} \partial_t(\frac{\sqrt{h}-1}{\varepsilon} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) + \big(\varepsilon\mathrm{F}_\mu^{-1}[v] + \sqrt{h}\big)\mathrm{F}_\mu\partial_x(\frac{\sqrt{h}-1}{\varepsilon} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) = R_7, \\ \partial_t(-\frac{\sqrt{h}-1}{\varepsilon} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) + \big(\varepsilon\mathrm{F}_\mu^{-1}[v] - \sqrt{h}\big)\mathrm{F}_\mu\partial_x(-\frac{\sqrt{h}-1}{\varepsilon} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v]) = R_8, \end{cases}$$

where R_7 and R_8 are of order $O(\mu\varepsilon)$.

We now denote

$$u^+ = \frac{\sqrt{h}-1}{\varepsilon} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v], \quad u^- = -\frac{\sqrt{h}-1}{\varepsilon} + \frac{1}{2}\mathrm{F}_\mu^{-1}[v].$$

We get from (2.6) and the above:

$$\begin{cases} \partial_t u^+ + (\varepsilon\frac{3u^+ + u^-}{2} + 1)\mathrm{F}_\mu\partial_x u^+ = O(\mu\varepsilon), \\ \partial_t u^- + (\varepsilon\frac{u^+ + 3u^-}{2} - 1)\mathrm{F}_\mu\partial_x u^- = O(\mu\varepsilon). \end{cases}$$

□

2.2 From the diagonalized Whitham-Boussinesq equations to the Whitham equations

In this subsection we prove Proposition 1.11.

To prove the first part of the proposition, i.e. there exists a time $T > 0$ such that for all times $t \in [0, \frac{T}{\epsilon}]$, $u^-(t,\cdot)$ and $u^+(t,\cdot)$ are of order $O(\mu)$, we need local well-posedness and stability results on the exact diagonalized system

$$\begin{cases} \partial_t u_e^+ + (\varepsilon\frac{3u_e^+ + u_e^-}{2} + 1)\mathrm{F}_\mu\partial_x u_e^+ = 0, \\ \partial_t u_e^- + (\varepsilon\frac{u_e^+ + 3u_e^-}{2} - 1)\mathrm{F}_\mu\partial_x u_e^- = 0. \end{cases} \tag{2.7}$$

We begin by proving the local well-posedness of system (2.7).

Proposition 2.9. *(local existence) Let $(\mu,\epsilon) \in \mathcal{A}$ (see Definition 1.2). Let $1/2 < t_0 \leq 1$, $\alpha \geq t_0 + 1$ and $u_{0,e}^+, u_{0,e}^- \in H^\alpha(\mathbb{R})$. Then there exists a time $T > 0$ such that system (2.7) admits a unique solution $(u_e^+, u_e^-) \in \mathcal{C}([0, \frac{T}{\varepsilon}]; H^\alpha(\mathbb{R})^2)$ with initial conditions $(u_{0,e}^+, u_{0,e}^-)$.*

Moreover $\frac{1}{T}$ and $\sup_{0\leq t\leq\frac{T}{\varepsilon}} |u_e^+|_{H^\alpha} + |u_e^-|_{H^\alpha}$ can be estimated by $|u_{0,e}^+|_{H^\alpha}$, $|u_{0,e}^-|_{H^\alpha}$, t_0, $\mu_{\max}$ (and are independent of the parameters $(\mu,\varepsilon) \in \mathcal{A}$).

Proof. System (2.7) is very similar to symmetric quasilinear hyperbolic systems. The only difference is the presence of the non-local operator F_μ. The well-posedness of such systems relies on the energy estimates (a priori estimates). We will establish them here. For the rest of the proof, one can follow the one in [14] where the author uses a method of regularisation suitable for our system .

First, remark that if we define $U_e(t) := \begin{pmatrix} u_e^+(t) \\ u_e^-(t) \end{pmatrix}$ and

$$A(U_e) := \begin{pmatrix} (\varepsilon\frac{3u_e^+ + u_e^-}{2} + 1) & 0 \\ 0 & (\varepsilon\frac{u_e^+ + 3u_e^-}{2} - 1) \end{pmatrix}. \tag{2.8}$$

Then system (2.7) can be written under the form

$$\partial_t U_e + A(U_e)\mathrm{F}_\mu \partial_x U_e = 0. \tag{2.9}$$

It justifies the following framework for energy estimates. Let $\beta \geq 0$, $R \in L^\infty([0, \frac{T}{\varepsilon}]; H^\beta(\mathbb{R})^2)$ and $\underline{U} \in L^\infty([0, \frac{T}{\varepsilon}]; H^{\max(\beta, t_0+1)}(\mathbb{R})^2)$. Let $U \in \mathcal{C}^0([0, \frac{T}{\varepsilon}]; H^\beta(\mathbb{R})^2) \cap \mathcal{C}^1([O, \frac{T}{\varepsilon}]; H^{\beta+1}(\mathbb{R})^2)$ solves the equation

$$\partial_t U + A(\underline{U})\mathrm{F}_\mu \partial_x U = \varepsilon R. \tag{2.10}$$

Lemma 2.10. *(Energy estimates of order 0) For all times $t \in [0, \frac{T}{\varepsilon}]$,*

$$|U(t)|_2 \leq e^{\varepsilon\delta_0 t}|U(0)|_2 + \varepsilon\gamma_0 \int_0^t |R(t')|_2 \,\mathrm{d}t',$$

where $\delta_0, \gamma_0 = C(T, |\underline{U}|_{L^\infty([0,\frac{T}{\varepsilon}];H^{t_0+1})})$

Proof. In what follows we will write $(\cdot,\cdot)_2$ the scalar product in $L^2(\mathbb{R})^2$. Using the fact that U solves equation (2.10) we get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}|U|_2^2 = (\partial_t U, U)_2 = -(A(\underline{U})\mathrm{F}_\mu \partial_x U, U)_2 + \varepsilon(R, U)_2.$$

But $A(\underline{U})$ is symmetric, so

$$\begin{aligned}(A(\underline{U})\mathrm{F}_\mu\partial_x U, U)_2 &= (\partial_x \mathrm{F}_\mu U, A(\underline{U})U)_2 = -(\mathrm{F}_\mu U, (\partial_x A(\underline{U}))U)_2 - (U, \mathrm{F}_\mu[A(\underline{U})\partial_x U])_2 \\ &= (\mathrm{F}_\mu U, (\partial_x A(\underline{U}))U)_2 - (A(\underline{U})\mathrm{F}_\mu\partial_x U, U)_2 - ([\mathrm{F}_\mu, A(\underline{U})]\partial_x U, U)_2,\end{aligned}$$

where $\partial_x A(\underline{U}) = \begin{pmatrix} \varepsilon\frac{3\partial_x \underline{u_1} + \partial_x \underline{u_2}}{2} & 0 \\ 0 & \varepsilon\frac{\partial_x \underline{u_1} + 3\partial_x \underline{u_2}}{2} \end{pmatrix}$ and

$$[\mathrm{F}_\mu, A(\underline{U})] = \begin{pmatrix} [\mathrm{F}_\mu, \varepsilon\frac{3\underline{u_1}+\underline{u_2}}{2} + 1] & 0 \\ 0 & [\mathrm{F}_\mu, \varepsilon\frac{\underline{u_1}+3\underline{u_2}}{2} - 1] \end{pmatrix} = \varepsilon \begin{pmatrix} [\mathrm{F}_\mu, \frac{3\underline{u_1}+\underline{u_2}}{2}] & 0 \\ 0 & [\mathrm{F}_\mu, \frac{\underline{u_1}+3\underline{u_2}}{2}] \end{pmatrix}.$$

Hence

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}|U|_2^2 = \frac{1}{2}(\mathrm{F}_\mu U, (\partial_x A(\underline{U}))U)_2 + \frac{1}{2}([\mathrm{F}_\mu, A(\underline{U})]\partial_x U, U)_2 + \varepsilon(R, U)_2. \tag{2.11}$$

Then using Cauchy-Schwarz inequality we get

$$\begin{aligned} &|U|_2\frac{\mathrm{d}}{\mathrm{dt}}|U|_2 \leq \frac{1}{2}|\partial_x(A(\underline{U}))U|_2|U|_2 + \frac{1}{2}|[\mathrm{F}_\mu, A(\underline{U})]\partial_x U|_2|U|_2 + \varepsilon|R|_2|U|_2 \\ \iff &\frac{\mathrm{d}}{\mathrm{dt}}|U|_2 \leq \frac{1}{2}|\partial_x(A(\underline{U}))U|_2 + \frac{1}{2}|[\mathrm{F}_\mu, A(\underline{U})]\partial_x U|_2 + \varepsilon|R|_2.\end{aligned}$$

But by product estimates A.1 we have

$$|\partial_x(A(\underline{U}))U|_2 \leq \varepsilon C(|\underline{U}|_{L_t^\infty H_x^{t_0+1}})|U|_2. \tag{2.12}$$

Moreover by commutator estimates A.5 we also have

$$|[\mathrm{F}_\mu, A(\underline{U})]\partial_x U|_2 \leq \varepsilon C(|\underline{U}|_{L_t^\infty H_x^{t_0+1}})|U|_2. \tag{2.13}$$

So that

$$\frac{\mathrm{d}}{\mathrm{dt}}|U|_2 \leq \varepsilon C(|\underline{U}|_{L_t^\infty H_x^{t_0+1}})|U|_2 + \varepsilon|R|_2 \leq \varepsilon\delta_0|U|_2 + \varepsilon|R|_2,$$

where $\delta_0 = C(|\underline{U}|_{L_t^\infty H_x^{t_0+1}})$.

We integrate this inequality in time

$$\begin{aligned} |U|_2 &\leq e^{\varepsilon\delta_0 t}|U(0)|_2 + \varepsilon\int_0^t e^{\varepsilon\delta_0(t-t')}|R|_2\,\mathrm{d}t' \\ &\leq e^{\varepsilon\delta_0 t}|U(0)|_2 + \varepsilon e^{\delta_0 T}\int_0^t |R|_2\,\mathrm{d}t'. \end{aligned}$$

It is the thesis with $\gamma_0 = e^{\delta_0 T}$.

□

Using this lemma, we get the energy estimates of order β.

Lemma 2.11. *(Energy estimates of order β) For all times $t \in [0, \frac{T}{\varepsilon}]$,*

$$|U(t)|_{H^\beta} \leq e^{\varepsilon\delta_\beta t}|U(0)|_{H^\beta} + \varepsilon\gamma_\beta\int_0^t |R(t')|_{H^\beta}\,\mathrm{d}t',$$

where $\delta_\beta, \gamma_\beta = C(T, |\underline{U}|_{L^\infty([0,\frac{T}{\varepsilon}];H^{\max(\beta,t_0+1)})})$.

Proof. Let $\Lambda^\beta := (1-\partial_x^2)^{\beta/2}$. Then $|\Lambda^\beta U|_2$ is equivalent to $|U|_{H^\beta}$. Applying Λ^β to (2.10) we have

$$\begin{aligned} &\Lambda^\beta\partial_t U + \Lambda^\beta[A(\underline{U})\mathrm{F}_\mu\partial_x U] = \varepsilon\Lambda^\beta R \\ \Longleftrightarrow\ &\partial_t\Lambda^\beta U + A(\underline{U})\mathrm{F}_\mu\partial_x\Lambda^\beta U = \varepsilon\Lambda^\beta R + \varepsilon B_\beta(\underline{U})\mathrm{F}_\mu\partial_x U, \end{aligned}$$

where

$$B_\beta(\underline{U}) := \frac{1}{\varepsilon}\begin{pmatrix} [\Lambda^\beta, \varepsilon\frac{3\underline{u}_1+\underline{u}_2}{2}+1] & 0 \\ 0 & [\Lambda^\beta, \varepsilon\frac{\underline{u}_1+3\underline{u}_2}{2}-1] \end{pmatrix} = \begin{pmatrix} [\Lambda^\beta, \frac{3\underline{u}_1+\underline{u}_2}{2}] & 0 \\ 0 & [\Lambda^\beta, \frac{\underline{u}_1+3\underline{u}_2}{2}] \end{pmatrix}.$$

Using commutator estimates A.5 and the boundedness of F_μ we get

$$|B_\beta(\underline{U})\mathrm{F}_\mu\partial_x U|_2 \lesssim |\underline{U}|_{L_t^\infty H_x^{\max(t_0+1,\beta)}}|U|_{H^\beta}.$$

Then we use the energy estimates of order 0 (see Lemma 2.10) to get

$$\begin{aligned} |U|_{H^\beta} &\leq e^{\varepsilon\delta_0 t}|U(0)|_{H^\beta} + \varepsilon\gamma_0\int_0^t |R|_{H^\beta} + \varepsilon\gamma_0\int_0^t |B_\beta(\underline{U})\partial_x U|_2\,\mathrm{d}t' \\ &\lesssim e^{\varepsilon\delta_0 t}|U(0)|_{H^\beta} + \varepsilon C(T, |\underline{U}|_{L_t^\infty H_x^{\max(t_0+1,\beta)}})\int_0^t |U|_{H^\beta}\,\mathrm{d}t' + \varepsilon\gamma_0\int_0^t |R|_{H^\beta}\,\mathrm{d}t'. \end{aligned}$$

It only remains to use Grönwall's lemma to get the result. □

This concluses the proof of Proposition 2.9. $\square$

Corollary 2.12. *Let $t_0 > 1/2$ and $\alpha \geq t_0 + 1$. Let $U_e = \begin{pmatrix} u_e^+ \\ u_e^- \end{pmatrix} \in \mathcal{C}([0, \frac{T}{\varepsilon}], H^\alpha(\mathbb{R})^2)$ solves the exact diagonalized system* (2.7) *with $u_e^-(0) = O(\mu)$, then for all times $[0, \frac{T}{\varepsilon}]$, $u_e^- = O(\mu)$. If instead we take $u_e^+(0) = O(\mu)$ then for all times $[0, \frac{T}{\varepsilon}]$, $u_e^+ = O(\mu)$.*

Proof. Using energy estimates of Lemma 2.11 on the equation on u_e^- of system (2.7), we get

$$|u_e^-|_{H^\alpha} \leq C(T, |U_e|_{L_t^\infty H_x^\alpha})|u_e^-(0)|_{H^\alpha}.$$

It gives the result on u_e^-. We do the same for u_e^+ with the other equation of (2.7). $\square$

Proposition 2.13. *(Stability) Let $t_0 > 1/2$ and $\alpha \geq t_0 + 1$. Let $U_e = \begin{pmatrix} u_e^+ \\ u_e^- \end{pmatrix} \in \mathcal{C}([0, \frac{T}{\varepsilon}], H^\alpha(\mathbb{R})^2)$ solve the exact diagonalized system* (2.7) *and $U = \begin{pmatrix} u^+ \\ u^- \end{pmatrix} \in \mathcal{C}([0, \frac{T}{\varepsilon}], H^\alpha(\mathbb{R})^2)$, defined from sufficiently regular solutions of the water waves equations satisfying the non-cavitation hypothesis* (1.3) *through the formulas* (2.4)*. By Proposition 2.3, we have*

$$\begin{cases} \partial_t U_e + A(U_e)\mathrm{F}_\mu \partial_x U_e = 0, \\ \partial_t U + A(U)\mathrm{F}_\mu \partial_x U = \mu\varepsilon R, \end{cases} \tag{2.14}$$

where A is defined by (2.8) *and $R \in L^\infty([0, \frac{T}{\varepsilon}], H^\alpha(\mathbb{R})^2)$ is uniformly bounded in $(\mu, \varepsilon) \in \mathcal{A}$.*

Then for all times $t \in [0, \frac{T}{\varepsilon}]$, we have

$$|U_e - U|_{H^\alpha} \leq c_\alpha(|U_e(0) - U(0)|_{H^\alpha} + \mu\varepsilon t|R|_{L_t^\infty H_x^\alpha}), \tag{2.15}$$

where $c_\alpha = C(T, |U_e|_{L_t^\infty H_x^\alpha}, |U|_{L_t^\infty H_x^{\alpha+1}})$.

In particular, if we take $U_e(0) = U(0)$ and $u^-(0) = O(\mu)$, then for all times $t \in [0, \frac{T}{\epsilon}]$,

$$u^- = O(\mu).$$

If instead, we take $u^+(0) = O(\mu)$, then for all times $t \in [0, \frac{T}{\epsilon}]$,

$$u^+ = O(\mu).$$

Proof. First we subtract the two equations of (2.14), we get

$$\partial_t(U_e - U) + A(U_e)\mathrm{F}_\mu \partial_x(U_e - U) = -\varepsilon(\mu R + \frac{1}{\epsilon}(A(U_e) - A(U))\mathrm{F}_\mu \partial_x U).$$

We use the estimates of order α from Lemma 2.11

$$\begin{aligned} |U_e - U|_{H^\alpha} \leq & e^{\varepsilon\delta_\alpha t}|U_e(0) - U(0)|_{H^\alpha} + \varepsilon\gamma_\alpha \int_0^t |\mu R(t')|_{H^\alpha}\, \mathrm{d}t' \\ & + \varepsilon\gamma_\alpha \int_0^t \frac{1}{\epsilon}|(A(U_e) - A(U))\mathrm{F}_\mu \partial_x U|_{H^\alpha}\, \mathrm{d}t'. \end{aligned}$$

But using product estimates A.1, the algebra properties of $H^\alpha(\mathbb{R})$ and the boundedness of F_μ, we have

$$|(A(U_e)-A(U))\mathrm{F}_\mu\partial_x U|_{H^\alpha} \lesssim \epsilon|U_e-U|_{H^\alpha}|U|_{H^{\alpha+1}}.$$

So that using Grönwall's lemma we get

$$\begin{aligned}|U_e-U|_{H^\alpha} &\leq \Big[e^{\varepsilon\delta_\alpha t}|U_e(0)-U(0)|_{H^\alpha}+\varepsilon\gamma_\alpha\int_0^t|\mu R(t')|_{H^\alpha}\,\mathrm{dt}'\Big]e^{\varepsilon\gamma_\alpha\int_0^t|U|_{H^{\alpha+1}}\,\mathrm{dt}'}\\ &\leq C(T,|U_e|_{L_t^\infty H_x^\alpha})(|U_e(0)-U(0)|_{H^\alpha}+\mu\varepsilon t|R|_{L_t^\infty H_x^\alpha})e^{\varepsilon\gamma_\alpha t|U|_{L_t^\infty H_x^{\alpha+1}}},\end{aligned}$$

which is the first part of the result.

Now if we suppose $U_e(0)=U(0)$ and $u^-(0)=O(\mu)$, then for all times $t\in[0,\frac{T}{\epsilon}]$, we have

$$|u^-|_{H^\alpha}\leq|u_e^- - u^-|_{H^\alpha}+|u_e^-|_{H^\alpha}\leq\mu c_\alpha T|R|_{L_t^\infty H_x^\alpha}+|u_e^-|_{H^\alpha}.$$

Using Corollary 2.12 we get the result. We do the same if $u^+(0)=O(\mu)$. □

We now have all the elements to prove Proposition 1.11.

Proof. Supposing $u^-(0)=O(\mu)$, we have for all times $t\in[0,\frac{T}{\varepsilon}]$, $\varepsilon u^-(t)=O(\mu\varepsilon)$ (see Proposition 2.13). But u^- and u^+ solve the first equation of (2.5). So

$$\partial_t u^+ + \mathrm{F}_\mu\partial_x u^+ + \frac{3\varepsilon}{2}u^+\partial_x u^+ = O(\mu\varepsilon). \tag{2.16}$$

The last equivalence coming from product estimates A.1 and $|(\mathrm{F}_\mu-1)[u]|_{H^\beta}\lesssim\mu|u|_{H^\beta}$ for any $\beta\geq 0$ and u in $H^{\beta+2}(\mathbb{R})$.

Supposing instead $u^+(0)=O(\mu)$, we have for all times $t\in[0,\frac{T}{\varepsilon}]$, $\varepsilon u^+(t)=O(\mu\varepsilon)$. Then, the second equation of (2.5) gives

$$\partial_t u^- - \mathrm{F}_\mu\partial_x u^- + \frac{3\varepsilon}{2}u^-\partial_x u^- = O(\mu\varepsilon).$$

□

It only remains to prove Proposition 1.12.

Proof. Let $U_e=\begin{pmatrix}u_e^+\\u_e^-\end{pmatrix}\in\mathcal{C}([0,\frac{T}{\varepsilon}],H^\alpha(\mathbb{R})^2)$ solve the system (2.7), with the initial data $u_{e,0}^-=O(\mu)$ and $u_{e,0}^+$, defined by (ζ_0,ψ_0) through the formulas (1.4). By the energy estimates of Lemma 2.11 applied on the second equation of system (2.7), we get for all times $t\in[0,\frac{T}{\varepsilon}]$,

$$|u_e^-|_{H^\alpha}\leq C|u_e^-(0)|_{H^\alpha},$$

where $C,T^{-1}=C(\frac{1}{h_{\min}},\mu_{\max},|\zeta_0|_{H^{\alpha+n}},|\psi_0|_{\dot{H}^{\alpha+n+1}})$. So that $u^+\in\mathcal{C}([0,\frac{T}{\varepsilon}],H^\alpha(\mathbb{R}))$, solution of the Whitham equation (which is well-posed in the Sobolev space $H^\alpha(\mathbb{R})$ [6])

$$\partial_t u^+ + \mathrm{F}_\mu\partial_x u^+ + \frac{3\varepsilon}{2}u^+\partial_x u^+ = 0,$$

satisfies (see (2.16))

$$\partial_t u^+ + (\varepsilon \frac{3u^+ + u_e^-}{2} + 1)\mathrm{F}_\mu \partial_x u^+ = \mu\varepsilon R,$$

where R is uniformly bounded in (μ,ε) in $H^\alpha(\mathbb{R})$. Using the stability estimates (2.15), with $U = \begin{pmatrix} u^+ \\ u_e^- \end{pmatrix}$ and $U_e(0) = U(0)$, we have for all times $t \in [0, \frac{T}{\varepsilon}]$,

$$|u_e^+ - u^+|_{H^\alpha} \leq C\mu\varepsilon t,$$

where $C, T^{-1} = C(\frac{1}{h_{\min}}, \mu_{\max}, |\zeta_0|_{H^{\alpha+n}}, |\psi_0|_{\dot{H}^{\alpha+n+1}})$.

Thus by Corollary 1.8 we have for all times $t \in [0, \frac{T}{\varepsilon}]$,

$$\begin{aligned} |(\zeta - \zeta_{\mathrm{Wh},+}, \psi - \psi_{\mathrm{Wh},+})|_{H^\alpha \times \dot{H}^{\alpha+1}} \leq & |(\zeta - \zeta_{\mathrm{c}}, \psi - \psi_{\mathrm{c}})|_{H^\alpha \times \dot{H}^{\alpha+1}} + |(\zeta_{\mathrm{c}} - \zeta_{\mathrm{Wh},+}, \psi_{\mathrm{c}} - \psi_{\mathrm{Wh},+})|_{H^\alpha \times \dot{H}^{\alpha+1}} \\ \leq & C(|u_e^-|_{H^{\alpha+1}} + \mu\varepsilon t) \\ \leq & C(|u_{e,0}^-|_{H^{\alpha+1}} + \mu\varepsilon t), \end{aligned}$$

where $C, T^{-1} = C(\frac{1}{h_{\min}}, \mu_{\max}, |\zeta_0|_{H^{\alpha+n}}, |\psi_0|_{\dot{H}^{\alpha+n+1}})$.

The reasoning is the same if instead $u_e^+(0) = O(\mu)$. □

3 Decoupling the water waves equations into two Whitham equations

The goal of this section is to prove Theorem 1.14. In the previous section, we proved the consistency (see Proposition 2.1 for the definition) of the water waves equations with the Whitham equations at a precision order $O(\mu\varepsilon)$ in the shallow water regime $\mathcal{A} := \{0 \leq \mu \leq \mu_{\max}, 0 \leq \varepsilon \leq 1\}$ (see Proposition 2.3). For that purpose we made a restrictive hypothesis on the initial conditions. We supposed either $u^-(0) = O(\mu)$ or $u^+(0) = O(\mu)$ (see Notation 1.10). In this section we use a generalisation of Birkhoff's normal form algorithm for almost smooth Hamiltonians, introduced by Bambusi in [1], on the water waves Hamiltonian. It will allow us to decouple the water waves equations into two Whitham equations satisfied by the two front waves, at a precision order $O(\mu\varepsilon + \varepsilon^2)$ in the shallow water regime, without any assumption of smallness on the initial conditions.

We will only recall the elements of Bambusi's theory which are useful to understand our reasoning. For more details see [1].

In this section we always suppose the parameters (μ,ε) in $\mathcal{A}$. Moreover every α will be a positive integer.

3.1 From the water waves equations to a Whitham-Boussinesq system

The starting point is the Hamiltonian of the water waves system

$$H_{\mathrm{WW}} = \frac{1}{2}\int_{\mathbb{R}} \zeta^2 \,\mathrm{dx} + \frac{1}{2}\int_{\mathbb{R}} \psi \frac{1}{\mu}\mathcal{G}^\mu[\varepsilon\zeta]\psi \,\mathrm{dx}\,. \tag{3.1}$$

We use the first shallow water estimates of Proposition 1.8 in [10], which we recall here for the sake of clarity.

Proposition 3.1. *Let $\alpha \geq 0$, and $\zeta \in H^{\alpha+4}(\mathbb{R})$ be such that (1.3) is satisfied. Let $\psi \in \dot{H}^{\alpha+3}(\mathbb{R})$, then*

$$|\frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi + \partial_x(h\mathrm{F}_{\mu}^2[\nabla\psi])|_{H^{\alpha}} \leq \mu\varepsilon M(\alpha+4)|\psi|_{\dot{H}^{\alpha+4}},$$

where $h = 1+\varepsilon\zeta$ and for any $\beta \geq 2$, $M(\beta) := C(\frac{1}{h_{\min}}, \mu_{\max}, |\zeta|_{H^{\beta}})$.

Proposition 3.2. *Let H_{WW} be the Hamiltonian of the water waves equations (3.1). Then for any $\alpha \geq 0$ and any $(\zeta,\psi) \in H^{\alpha+4}(\mathbb{R}) \times \dot{H}^{\alpha+4}(\mathbb{R})$ we have*

$$|J\nabla H_{\mathrm{WW}} - J\nabla(H_0+\varepsilon H_1)|_{H^{\alpha}\times\dot{H}^{\alpha+1}} < \mu\varepsilon M(\alpha+4)|\psi|_{\dot{H}^{\alpha+4}}, \tag{3.2}$$

where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $H_0 := \frac{1}{2}\int_{\mathbb{R}}\zeta^2\,\mathrm{dx} + \frac{1}{2}\int_{\mathbb{R}}(\mathrm{F}_{\mu}[\partial_x\psi])^2\,\mathrm{dx}$ and $H_1 := \frac{1}{2}\int_{\mathbb{R}}\zeta(\mathrm{F}_{\mu}[\partial_x\psi])^2\,\mathrm{dx}$.

Proof. We defined H_{WW} as the Hamiltonian of the water waves equations, so

$$J\nabla H_{\mathrm{WW}} = \begin{pmatrix} \frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi \\ -\zeta - \frac{\varepsilon}{2}(\partial_x\psi)^2 + \frac{\mu\varepsilon}{2}\frac{(\frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi + \varepsilon\partial_x\zeta\partial_x\psi)^2}{1+\varepsilon^2\mu(\partial_x\zeta)^2} \end{pmatrix}. \tag{3.3}$$

We can easily compute $J\nabla(H_0+\varepsilon H_1)$, we get

$$J\nabla(H_0+\varepsilon H_1) = \begin{pmatrix} -\mathrm{F}_{\mu}^2[\partial_x^2\psi] - \mathrm{F}_{\mu}[\partial_x(\varepsilon\zeta\mathrm{F}_{\mu}[\partial_x\psi])] \\ -\zeta - \frac{\varepsilon}{2}(\mathrm{F}_{\mu}[\partial_x\psi])^2 \end{pmatrix}. \tag{3.4}$$

We compare the first term of (3.3) and (3.4).

$$\begin{aligned} &|\frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi + \mathrm{F}_{\mu}^2[\partial_x^2\psi] + \mathrm{F}_{\mu}[\partial_x(\varepsilon\zeta\mathrm{F}_{\mu}[\partial_x\psi])]|_{H^{\alpha}} \\ \leq &|\frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi + \partial_x(h\mathrm{F}_{\mu}^2[\partial_x\psi])|_{H^{\alpha}} + |(\mathrm{F}_{\mu}-1)[\partial_x(\varepsilon\zeta\mathrm{F}_{\mu}[\partial_x\psi])]|_{H^{\alpha}} + |\partial_x(\varepsilon\zeta(\mathrm{F}_{\mu}-1)\mathrm{F}_{\mu}[\partial_x\psi])|_{H^{\alpha}}. \end{aligned}$$

Hence, using Proposition 3.1, the fact that for any $u \in H^{\alpha+2}(\mathbb{R})$, $|(\mathrm{F}_{\mu}-1)[u]|_{H^{\alpha}} \lesssim \mu|u|_{H^{\alpha+2}}$ and product estimates A.1, we get

$$|\frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi + \mathrm{F}_{\mu}^2[\partial_x^2\psi] + \mathrm{F}_{\mu}[\partial_x(\varepsilon\zeta\mathrm{F}_{\mu}[\partial_x\psi])]|_{H^{\alpha}} \lesssim \mu\varepsilon M(\alpha+4)|\psi|_{\dot{H}^{\alpha+4}}.$$

Then we compare the second term of (3.3) and (3.4). We first remark that there is a part of (3.3) which is clearly of order $O(\mu\varepsilon)$, we start by estimate it using quotient estimates A.3, product estimates A.1 and Proposition A.6

$$\begin{aligned} |\frac{(\frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi + \varepsilon\partial_x\zeta\partial_x\psi)^2}{1+\varepsilon^2\mu(\partial_x\zeta)^2}|_{\dot{H}^{\alpha+1}} \leq &|\frac{(\frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi + \varepsilon\partial_x\zeta\partial_x\psi)^2}{1+\varepsilon^2\mu(\partial_x\zeta)^2}|_{H^{\alpha+1}} \\ \leq &C(\mu_{\max}, |(\partial_x\zeta)^2|_{H^{\alpha+1}})|(\frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi + \varepsilon\partial_x\zeta\partial_x\psi)^2|_{H^{\alpha+1}} \\ \leq &C(\mu_{\max}, |\zeta|_{H^{\alpha+2}})|\frac{1}{\mu}\mathcal{G}^{\mu}[\varepsilon\zeta]\psi + \varepsilon\partial_x\zeta\partial_x\psi|_{H^{\alpha+1}}^2 \\ \leq &M(\alpha+3)|\psi|_{\dot{H}^{\alpha+3}}. \end{aligned}$$

Then, it only remains to see that using product estimates A.1 and $|(\mathrm{F}_\mu - 1)[u]|_{H^\alpha} \lesssim \mu |u|_{H^{\alpha+2}}$ for any $u \in H^{\alpha+2}(\mathbb{R})$, we have

$$\begin{aligned}|(\partial_x\psi)^2 - (\mathrm{F}_\mu[\partial_x\psi])^2|_{\dot{H}^{\alpha+1}} \leq & |(\partial_x\psi)^2 - (\mathrm{F}_\mu[\partial_x\psi])^2|_{H^{\alpha+1}} \\ \leq & |\partial_x\psi - \mathrm{F}_\mu[\partial_x\psi]|_{H^{\alpha+1}} |\partial_x\psi + \mathrm{F}_\mu[\partial_x\psi]|_{H^{\alpha+1}} \lesssim \mu |\psi|_{\dot{H}^{\alpha+4}} |\psi|_{\dot{H}^{\alpha+2}}.\end{aligned}$$

This ends the proof. □

Remark 3.3.

- *When deriving the Hamilton's equations associated with $H_0 + \varepsilon H_1$ we get the following Whitham-Boussinesq system*

$$\begin{cases} \partial_t\zeta + \mathrm{F}_\mu^2[\partial_x^2\psi] + \varepsilon \mathrm{F}_\mu \partial_x[\zeta \mathrm{F}_\mu[\partial_x\psi]] = 0, \\ \partial_t\psi + \zeta + \frac{\varepsilon}{2}(\mathrm{F}_\mu[\partial_x\psi])^2 = 0. \end{cases}$$

  It is equivalent to the system (2.1) *at a precision order $O(\mu\varepsilon)$.*

- *We could approximate the Hamiltonian H_{WW} with any other perturbation of $H_0 + \varepsilon H_1$ of order $O(\mu\varepsilon)$. It wouldn't change the reasoning. Our choice is made to simplify the computations.*

Notations 3.4. *We will denote by $\widetilde{H}_0 + \varepsilon \widetilde{H}_1$ the Hamiltonian $H_0 + \varepsilon H_1$ rewritten in the variables (ζ, v) where $v = \partial_x\psi$. It is associated with the Poisson tensor $\widetilde{J} = \begin{pmatrix} 0 & -\partial_x \\ -\partial_x & 0 \end{pmatrix}$. We recall that a change of unknowns $\mathcal{T}(\eta,\omega)$ turns the initial Poisson tensor, denoted J, into $\left(\frac{\partial F}{\partial(\eta,\omega)}\right) J \left(\frac{\partial F}{\partial(\eta,\omega)}\right)^*$, where $*$ is the adjoint in $L^2(\mathbb{R})^2$.*

Before applying Birkhoff's algorithm, we change the unknowns of the approximated Hamiltonian $H_0 + \varepsilon H_1$ with those usually used to diagonalize the linear part of the water waves equations.

Proposition 3.5. *Let r and s be defined as*

$$r := \frac{\zeta + \mathrm{F}_\mu[\partial_x\psi]}{2}, \quad s := \frac{\zeta - \mathrm{F}_\mu[\partial_x\psi]}{2}. \tag{3.5}$$

We have

$$(\widetilde{H}_0 + \varepsilon \widetilde{H}_1)(\zeta, v) = (H_0 + \varepsilon H_1)(\zeta, \psi) = \int_{\mathbb{R}} (r^2 + s^2)\,\mathrm{dx} + \frac{\varepsilon}{2} \int (r^3 + s^3)\,\mathrm{dx} - \frac{\varepsilon}{2} \int_{\mathbb{R}} (r^2 s + r s^2)\,\mathrm{dx}\,. \tag{3.6}$$

We will denote the later Hamiltonian $H_{\mathrm{BW}}(r,s)$. The Poisson tensor associated with is $J_\mu = \begin{pmatrix} -\frac{\mathrm{F}_\mu\partial_x}{2} & 0 \\ 0 & \frac{\mathrm{F}_\mu\partial_x}{2} \end{pmatrix}$.

Proof. Easy computations. □

For later purposes we also define the inverse of the transformation (3.5).

Property 3.6. *Let $\alpha \geq 0$ and $\mathcal{T}_{\mathrm{D}} : H^{\alpha+1}(\mathbb{R}) \times H^{\alpha+1}(\mathbb{R}) \to H^{\alpha}(\mathbb{R}) \times H^{\alpha}(\mathbb{R})$ be defined by*

$$\mathcal{T}_{\mathrm{D}}(r,s) := \begin{pmatrix} r+s \\ \mathrm{F}_{\mu}^{-1}[r-s] \end{pmatrix}. \tag{3.7}$$

Then $(\widetilde{H}_0 + \varepsilon \widetilde{H}_1)(\mathcal{T}_{\mathrm{D}}(r,s)) = H_{\mathrm{BW}}(r,s)$.

Proof. For any $\alpha \geq 0$ and any $u \in H^{\alpha+1}(\mathbb{R})$, we have $|\mathrm{F}_{\mu}^{-1}[u]|_{H^{\alpha}} \lesssim |u|_{H^{\alpha+1}}$. □

To separate each important part of the Hamiltonian H_{BW} we set some notations.

Notations 3.7. *Let*

$$\begin{cases} L := \int_{\mathbb{R}} (r^2 + s^2)\,\mathrm{dx}, \\ Z := \frac{1}{2} \int (r^3 + s^3)\,\mathrm{dx}, \\ W := -\frac{1}{2} \int_{\mathbb{R}} (r^2 s + r s^2)\,\mathrm{dx}. \end{cases}$$

We have

$$H_{\mathrm{BW}} = L + \varepsilon Z + \varepsilon W.$$

Remark 3.8. *Our goal is to decouple the equations associated with the Hamiltonian H_{BW}. We remark that only W gives coupled terms in the Hamilton's equations.*

3.2 Application of Birkhoff's algorithm

We suppose first that all our objects of study are smooth. If $G(r,s)$ is a smooth function then it's corresponding Hamilton's equations are

$$\partial_t \begin{pmatrix} r \\ s \end{pmatrix} = J_{\mu} \nabla G.$$

We will denote by Φ_G^t the corresponding flow.

If $F(r,s)$ is another smooth function, we denote the Lie derivative of F with respect to G by

$$\{F, G\}_{\mu} := (\nabla F, J_{\mu} \nabla G),$$

where $(\cdot,\cdot)$ denotes the scalar product in $L^2(\mathbb{R})^2$.

Property 3.9. *Let $G(r,s)$ be a smooth function and Φ_G^t its flow associated with the Poisson tensor J_{μ}. Let also $F(r,s)$ be another smooth function. Then*

$$F \circ \Phi_G^{\varepsilon} = F + \varepsilon \{F, G\}_{\mu} + O(\varepsilon^2).$$

Proof. See Subsection 4.1 in [1]. □

Suppose for now that in our case the Hamiltonian H_{BW} is smooth. Then for any smooth function $G(r,s)$ we get

$$H_{\mathrm{BW}} \circ \Phi_G^{\varepsilon} = L + \varepsilon \{L, G\}_{\mu} + \varepsilon Z + \varepsilon W + O(\varepsilon^2).$$

Remark 3.10. *Here it seems that to get a normal form at the order of precision $O(\varepsilon^2)$, we need G to solve the homological equation*

$$\{L, G\}_\mu + W = 0. \tag{3.8}$$

In [1], the author gave the solution of such equations. If

$$\lim_{\tau \to +\infty} W(\Phi_L^\tau) + W(\Phi_L^{-\tau}) = 0, \tag{3.9}$$

(in our case, we can prove this assumption using a Littlewood-Paley decomposition) then the solution of (3.8) *is*

$$G(r, s) = -\frac{1}{2} \int_{\mathbb{R}} \operatorname{sgn}(\tau) W(\Phi_L^\tau(r, s)) d\tau, \tag{3.10}$$

at the additional condition that the above function is well-defined.

However the latter condition is not easy to verify in general. We expose here a naive attempt to do so in our context.

The flow Φ_L^τ satisfy

$$\frac{d}{d\tau} \Phi_L^\tau = J_\mu \nabla L(\Phi_L^\tau).$$

So that

$$\Phi_L^\tau(r, s) = \begin{pmatrix} e^{-iD\mathrm{F}_\mu \tau} r \\ e^{iD\mathrm{F}_\mu \tau} s \end{pmatrix},$$

and

$$W(\Phi_L^\tau(r, s)) = -\frac{1}{2} \int_{\mathbb{R}} (e^{-iD\mathrm{F}_\mu \tau} r)^2 e^{iD\mathrm{F}_\mu \tau} s \,\mathrm{dx} - \frac{1}{2} \int_{\mathbb{R}} (e^{iD\mathrm{F}_\mu \tau} s)^2 e^{-iD\mathrm{F}_\mu \tau} r \,\mathrm{dx}\,.$$

So

$$\begin{aligned} |W(\Phi_L^\tau(r, s))| &\le \int_{\mathbb{R}} |e^{-iD\mathrm{F}_\mu \tau} r|^2 |e^{iD\mathrm{F}_\mu \tau} s| \,\mathrm{dx} + \int_{\mathbb{R}} |e^{iD\mathrm{F}_\mu \tau} s|^2 |e^{-iD\mathrm{F}_\mu \tau} r| \,\mathrm{dx} \\ &\le |e^{iD\mathrm{F}_\mu \tau} s|_\infty |e^{-iD\mathrm{F}_\mu \tau} r|_2^2 + |e^{-iD\mathrm{F}_\mu \tau} r|_\infty |e^{iD\mathrm{F}_\mu \tau} s|_2^2. \end{aligned}$$

But $|e^{-iD\mathrm{F}_\mu \tau} r|_2 = |r|_2$, and the dispersive estimates only give a decrease in time of $|e^{iD\mathrm{F}_\mu \tau} s|_\infty$ and $|e^{-iD\mathrm{F}_\mu \tau} r|_\infty$ of order $1/t^{-1/2}$ [4].

The idea to overcome this issue is to only solve an approximation of the homological equation at order $O(\mu\varepsilon)$.

We define another Lie derivative, associated with the Poisson tensor $J_{\mathrm{simp}} := \begin{pmatrix} -\partial_x/2 & 0 \\ 0 & \partial_x/2 \end{pmatrix}$:

$$\{L, G\}_{\mathrm{simp}} := (\nabla L, J_{\mathrm{simp}} \nabla G)$$

Property 3.11. *We have*

$$\{L, G\}_\mu = \{L, G\}_{\mathrm{simp}} + O(\mu)$$

Proof. By definition $\{L,G\}_\mu = (\nabla L, J_\mu \nabla G)$, with $J_\mu = \begin{pmatrix} -\frac{\mathrm{F}_\mu \partial_x}{2} & 0 \\ 0 & \frac{\mathrm{F}_\mu \partial_x}{2} \end{pmatrix}$. So

$$\{L,G\}_\mu - \{L,G\}_{\text{simp}} = ((\nabla L, (J_\mu - J_{\text{simp}})\nabla G),$$

with $J_\mu - J_{\text{simp}} = \begin{pmatrix} -\frac{\mathrm{F}_\mu - 1}{2} & 0 \\ 0 & \frac{\mathrm{F}_\mu - 1}{2} \end{pmatrix}$. But for any $\alpha \geq 0$ and any $u \in H^{\alpha+2}(\mathbb{R})$, we have $|(\mathrm{F}_\mu - 1)[u]|_{H^\alpha} \lesssim \mu |u|_{H^{\alpha+2}}$. □

The Lie derivative associated with the Poisson tensor J_{simp} is the same one as in Bambusi's article. To write an explicit expression of the solution of the simplified homological equation associated with

$$\{L,G\}_{\text{simp}} + W = 0, \tag{3.11}$$

we need the definition of a classical primitive operator and some properties on it.

Definition/Property 3.12. *Let $\alpha \geq 0$. We define the operator $\partial^{-1} : W^{\alpha,1}(\mathbb{R}) \to W^{\alpha,\infty}(\mathbb{R})$ by the formula*

$$\partial^{-1} u(y) = \frac{1}{2} \int_{\mathbb{R}} \operatorname{sgn}(y - y_1) u(y_1)\, \mathrm{dy}_1\,. \tag{3.12}$$

Here are some properties on it:

- *It is continuous.*
- *It is skew-adjoint for the scalar product in $L^2(\mathbb{R})$.*
- *If $\lim_{x\to+\infty} u(x) + u(-x) = 0$ (it is the case for any function $u \in W^{1,1}(\mathbb{R})$, see Corollary 8.9 in [3]), then $\partial^{-1}(\partial_x u) = u$.*
- *$\partial_x(\partial^{-1} u) = u$.*

We can now write explicitly the solution of the simplified homological equation.

Proposition 3.13. *Let $W(r,s) = -\frac{1}{2}\int_{\mathbb{R}}(r^2 s + r s^2)\,\mathrm{dx}$, and $L(r,s) = \int_{\mathbb{R}}(r^2+s^2)\,\mathrm{dx}$ for any $(r,s) \in W^{1,1}(\mathbb{R})$. Then the solution of the simplified homological equation* (3.11) *is*

$$G(r,s) = \frac{1}{4}\int_{\mathbb{R}} [\partial^{-1}(r^2)s + \partial^{-1}(r)s^2]\,\mathrm{dx}\,.$$

Proof. See Lemma 5.2 in [1]. □

Here G is clearly well-defined for r and s in $W^{\alpha,1}(\mathbb{R})$ with $\alpha \geq 1$ large enough. However there is another problem, G does not generate a flow. For that reason we use the generalization of Birkhoff's algorithm in the case of "almost smooth" functions depending on small parameters (see [1]). We recall here what is an almost smooth function.

Definition 3.14. *Let $\{\mathcal{B}^\alpha\}_{\alpha\geq 0}$ be a Banach scale. A map $F(r,s,\mu,\varepsilon)$ is said to be almost smooth if $\forall \beta,\gamma\geq 0$, there exists δ and an open neighbourhood of the origin $\mathcal{U}\subset\mathcal{B}^\delta\times\mathbb{R}\times\mathbb{R}$ (where $\mathbb{R}\times\mathbb{R}$ is the domain of (μ,ε)), such that*

$$F\in\mathcal{C}^\beta(\mathcal{U},\mathcal{B}^\gamma\times\mathbb{R}\times\mathbb{R}).$$

Due to the definition of ∂^{-1}, for the rest of the article, we will work with the Banach scale $\{W^{\alpha,1}(\mathbb{R})\times W^{\alpha,1}(\mathbb{R})\}_{\alpha\geq 0}$. If

- for any $\alpha\geq 0$, there exists $\alpha'\geq 0$ such that the Poisson tensor J_μ is bounded from $W^{\alpha',1}(\mathbb{R})\times W^{\alpha',1}(\mathbb{R})$ to $W^{\alpha,1}(\mathbb{R})\times W^{\alpha,1}(\mathbb{R})$,
- the function $G(r,s)$ defined in Proposition 3.10 has an almost smooth vector field, i.e. $J_\mu\nabla G$ is almost smooth, same for $\{L,G\}_{\text{simp}}$,

then the theory developed in [1] gives the existence of a transformation $\mathcal{T}_{\text{B}}$ such that

$$\begin{aligned}H_{\text{BW}}\circ\mathcal{T}_{\text{B}} &= L+\varepsilon\{L,G\}_\mu+\varepsilon Z+\varepsilon W+O(\varepsilon^2)\\ &= L+\varepsilon\{L,G\}_{\text{simp}}+\varepsilon Z+\varepsilon W+O(\mu\varepsilon+\varepsilon^2)\\ &= L+\varepsilon Z+O(\mu\varepsilon+\varepsilon^2),\end{aligned}$$

where $A=B+O(\mu^k\varepsilon^l)$ now means $\frac{A-B}{\mu^k\varepsilon^l}$ has an almost smooth vector field.

We verify the latter two conditions.

Property 3.15. *For any $\alpha\geq 0$ the Fourier multiplier F_μ is bounded from $W^{\alpha+1,1}(\mathbb{R})$ to $W^{\alpha,1}(\mathbb{R})$. In particular, the Poisson tensor J_μ is bounded from $W^{\alpha+2,1}(\mathbb{R})\times W^{\alpha+2,1}(\mathbb{R})$ to $W^{\alpha,1}(\mathbb{R})\times W^{\alpha,1}(\mathbb{R})$.*

The inverse Fourier multiplier of F_μ, denoted $\mathrm{F}_\mu^{-1}:=\sqrt{\frac{\sqrt{\mu}|D|}{\tanh(\sqrt{\mu}|D|)}}$, is bounded from $W^{\alpha+2,1}(\mathbb{R})$ to $W^{\alpha,1}(\mathbb{R})$.

Proof. First, we remark that we only need to prove the boundedness of F_μ in $W^{1,1}(\mathbb{R})\to L^1(\mathbb{R})$. To do so, we consider an even cutoff function $\chi\in\mathcal{C}^\infty((-2,2))$ such that $\chi(\xi)=1$ for $|\xi|\leq 1$. Let f be in $W^{1,1}(\mathbb{R})$, we can decompose $\mathrm{F}_\mu[f]$ into

$$\begin{aligned}\mathrm{F}_\mu[f](x)=&\int_{\mathbb{R}}\int_{\mathbb{R}}e^{-i(y-x)\xi}F_\mu(\xi)\chi(\xi)f(y)\,\mathrm{d}y\,d\xi+\int_{\mathbb{R}}\int_{\mathbb{R}}e^{-i(y-x)\xi}F_\mu(\xi)(1-\chi(\xi))f(y)\,\mathrm{d}y\,d\xi\\ :=&I_1+I_2.\end{aligned}$$

We deal with each term separately.

- We begin with I_1. We can write it using the inverse Fourier transform, denoted here $\mathcal{F}^{-1}$:

  $$I_1=\int_{\mathbb{R}}\int_{\mathbb{R}}e^{-i(y-x)\xi}F_\mu(\xi)\chi(\xi)f(y)\,\mathrm{d}y\,d\xi=\mathcal{F}^{-1}[F_\mu(\xi)\chi(\xi)]*f.$$

  Using Young's inequality, we get

  $$|\mathcal{F}^{-1}[F_\mu(\xi)\chi(\xi)]*f|_{L^1}\leq|\mathcal{F}^{-1}[F_\mu(\xi)\chi(\xi)]|_{L^1}|f|_{L^1}.$$

But

$$|\mathcal{F}^{-1}[F_\mu(\xi)\chi(\xi)]|_{L^1} \lesssim |F_\mu(\xi)\chi(\xi))|_{H^1}.$$

And $\forall \xi \in \mathbb{R}$ $|F_\mu(\xi)|, |F'_\mu(\xi)| \leq 1$, so $\mathcal{F}^{-1}[F_\mu(\xi)\chi(\xi)]$ is in $L^1(\mathbb{R})$. It finishes the case of I_1.

- Now we deal with I_2. Doing an integration by parts, we obtain

$$\begin{aligned}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-i(y-x)\xi} F_\mu(\xi)(1-\chi(\xi))f(y)\,\mathrm{dy}\,d\xi &= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-i(y-x)\xi}\frac{F_\mu(\xi)}{i\xi}(1-\chi(\xi))\partial_y f(y)\,\mathrm{dy}\,d\xi \\ &= \mathcal{F}^{-1}[\frac{F_\mu(\xi)}{i\xi}(1-\chi(\xi))] * \partial_y f\end{aligned}$$

Using again Young's inequality, we get

$$|\mathcal{F}^{-1}[\frac{F_\mu(\xi)}{i\xi}(1-\chi(\xi))] * \partial_y f|_{L^1} \leq |\mathcal{F}^{-1}[\frac{F_\mu(\xi)}{i\xi}(1-\chi(\xi))]|_{L^1}|\partial_y f|_{L^1}.$$

But

$$|\mathcal{F}^{-1}[\frac{F_\mu(\xi)}{i\xi}(1-\chi(\xi))]|_{L^1} \lesssim |\frac{F_\mu(\xi)}{i\xi}(1-\chi(\xi))|_{H^1}.$$

So $\mathcal{F}^{-1}[\frac{F_\mu(\xi)}{i\xi}(1-\chi(\xi))] \in L^1(\mathbb{R})$. It finishes the case of I_2.

At the end, we have $|\mathrm{F}_\mu[f]|_{L^1} \leq |I_1|_{L^1} + |I_2|_{L^1} \lesssim |f|_{W^{1,1}}$.

With the same process we obtain the result on F_μ^{-1} (we need a second integration by parts in the high frequencies to get $\mathcal{F}^{-1}[\frac{F_\mu^{-1}(\xi)}{(i\xi)^2}(1-\chi(\xi))]$ in L^1). □

Property 3.16. *The function $G(r,s)$ defined in Proposition 3.13 and $\{L,G\}_{\text{simp}}$ have an almost smooth vector field.*

Proof. First we have $2\partial_r G = -r\partial^{-1}(s) - \frac{1}{2}\partial^{-1}(s^2)$, where we used the skew-adjoint property of ∂^{-1}. So $-2\mathrm{F}_\mu\partial_x\partial_r G = \mathrm{F}_\mu[\partial_x(r)\partial^{-1}(s)] + \mathrm{F}_\mu[rs] + \frac{1}{2}\mathrm{F}_\mu[s^2]$. But for any $\alpha \geq 0$, F_μ is bounded from $W^{\alpha+1,1}(\mathbb{R})$ to $W^{\alpha,1}(\mathbb{R})$, and $W^{\alpha+1,1}(\mathbb{R})$ is an algebra. We also know that ∂^{-1} is a linear operator bounded from $W^{\alpha+1,1}(\mathbb{R})$ to $W^{\alpha+1,\infty}(\mathbb{R})$, and the product of a $W^{\alpha+1,\infty}(\mathbb{R})$ with a $W^{\alpha+1,1}(\mathbb{R})$ function is still in $W^{\alpha+1,1}(\mathbb{R})$. Combining all these elements, we get the result for G.

The same arguments give the result for $\{L,G\}_{\text{simp}}$. □

We have an explicit formulation of the transformation $\mathcal{T}_{\mathrm{B}}$ given by (4.25) in [1], taking $X = J_{\text{simp}}\nabla G$, on which we can do precise estimates.

Proposition 3.17. *Let $\alpha \geq 0$ and $\mathcal{T}_{\mathrm{B}} : W^{\alpha+1,1}(\mathbb{R}) \times W^{\alpha+1,1}(\mathbb{R}) \to W^{\alpha,1}(\mathbb{R}) \times W^{\alpha,1}(\mathbb{R})$ be defined by*

$$\mathcal{T}_{\mathrm{B}}(r,s) = \begin{pmatrix} r \\ s \end{pmatrix} + \varepsilon J_{\text{simp}}\nabla G(r,s) = \begin{pmatrix} r + \frac{\varepsilon}{4}\partial_x(r)\partial^{-1}(s) + \frac{\varepsilon}{4}rs + \frac{\varepsilon}{8}s^2 \\ s + \frac{\varepsilon}{4}\partial_x(s)\partial^{-1}(r) + \frac{\varepsilon}{4}rs + \frac{\varepsilon}{8}r^2 \end{pmatrix}. \tag{3.13}$$

Then

$$H_{\mathrm{BW}} \circ \mathcal{T}_{\mathrm{B}} = L + \varepsilon Z + O(\mu\varepsilon + \varepsilon^2)$$

Proof. Let $r, s \in W^{\alpha+1,1}(\mathbb{R})$. The latter Sobolev space in an algebra, so the only terms we need to look at are $\partial_x r \partial^{-1} s$ and $\partial_x s \partial^{-1} r$. We have

$$|\partial_x r \partial^{-1} s|_{W^{\alpha,1}} \leq |\partial^{-1} s|_{W^{\alpha,\infty}} |\partial_x r|_{W^{\alpha,1}}.$$

Here using the continuity of ∂^{-1} from $L^1(\mathbb{R})$ to $L^\infty(\mathbb{R})$ for any $\beta \geq 0$ and the third property of Definition/Property 3.12, we get

$$\begin{cases} |\partial^{-1} s|_{W^{\alpha,\infty}} \lesssim |s|_{L^1} \quad \text{if} \quad \alpha = 0, \\ |\partial^{-1} s|_{W^{\alpha,\infty}} \lesssim |s|_{L^1} + \sum\limits_{\beta \leq \alpha} |\partial_x^\beta \partial^{-1} s|_{L^\infty} = |s|_{L^1} + \sum\limits_{\beta \leq \alpha} |\partial^{-1} \partial_x^\beta s|_{L^\infty} \lesssim |s|_{W^{\alpha,1}} \quad \text{if} \quad \alpha \geq 1. \end{cases} \tag{3.14}$$

Idem for the other term. Thus $\mathcal{T}_\mathrm{B} : W^{\alpha+1,1}(\mathbb{R}) \times W^{\alpha+1,1}(\mathbb{R}) \to W^{\alpha,1}(\mathbb{R}) \times W^{\alpha,1}(\mathbb{R})$.

From the definition of H_BW (see (3.6)) and (3.13) we have

$$\begin{aligned} H_\mathrm{BW}(\mathcal{T}_\mathrm{B}(r,s)) = & \int_\mathbb{R} \Big(r^2 + \frac{\varepsilon}{2} r \partial_x r \partial^{-1} s + \frac{\varepsilon}{2} r^2 s + \frac{\varepsilon}{4} r s^2 \Big) \,\mathrm{dx} \\ & + \int_\mathbb{R} \Big(s^2 + \frac{\varepsilon}{2} s \partial_x s \partial^{-1} r + \frac{\varepsilon}{2} r s^2 + \frac{\varepsilon}{4} s r^2 \Big) \,\mathrm{dx} \\ & - \frac{\varepsilon}{2} \int_\mathbb{R} (r^2 s + r s^2) \,\mathrm{dx} + \frac{\varepsilon}{2} \int_\mathbb{R} (r^3 + s^3) \,\mathrm{dx} + O(\varepsilon^2). \end{aligned}$$

But using the skew-adjointness of ∂^{-1} we get

$$\begin{cases} \int_\mathbb{R} r \partial_x r \partial^{-1} s \,\mathrm{dx} = \frac{1}{2} \int_\mathbb{R} \partial_x (r^2) \partial^{-1} s \,\mathrm{dx} = -\frac{1}{2} \int_\mathbb{R} \partial^{-1} (\partial_x r^2) s \,\mathrm{dx} = -\frac{1}{2} \int_\mathbb{R} r^2 s \,\mathrm{dx}, \\ \int_\mathbb{R} s \partial_x s \partial^{-1} r \,\mathrm{dx} = \frac{1}{2} \int_\mathbb{R} \partial_x (s^2) \partial^{-1} r \,\mathrm{dx} = -\frac{1}{2} \int_\mathbb{R} \partial^{-1} (\partial_x s^2) r \,\mathrm{dx} = -\frac{1}{2} \int_\mathbb{R} s^2 r \,\mathrm{dx}. \end{cases}$$

Thus

$$\begin{aligned} H_\mathrm{BW}(\mathcal{T}_\mathrm{B}(r,s)) &= \int_\mathbb{R} (r^2 + s^2) \,\mathrm{dx} + \frac{\varepsilon}{2} \int_\mathbb{R} (r^3 + s^3) \,\mathrm{dx} + O(\varepsilon^2) \\ &= L(r,s) + \varepsilon Z(r,s) + O(\varepsilon^2). \end{aligned}$$

□

Remark 3.18. *The theory developped in [1] tells us that the transformation $\mathcal{T}_\mathrm{B}$ should be $\mathcal{T}_\mathrm{B}(r,s) = \begin{pmatrix} r \\ s \end{pmatrix} + \varepsilon J_\mu \nabla G(r,s)$. But the precision we want is $O(\mu\varepsilon + \varepsilon^2)$, and $\varepsilon J_\mu = \varepsilon J_\mathrm{simp} + O(\mu\varepsilon)$.*

The transformation (3.13) preserves the Hamiltonian structure associated with the Poisson tensor J_μ at the order of precision $O(\mu\varepsilon + \varepsilon^2)$, see the following proposition.

Proposition 3.19. *Let $\alpha \geq 0$. Let $r, s \in W^{\alpha+4,1}(\mathbb{R})$ Let $\frac{\partial \mathcal{T}_\mathrm{B}}{\partial(r,s)}$ be the Jacobian matrix of the transformation T. We have*

$$\Big(\frac{\partial \mathcal{T}_\mathrm{B}}{\partial(r,s)} \Big) J_\mu \Big(\frac{\partial \mathcal{T}_\mathrm{B}}{\partial(r,s)} \Big)^* = J_\mu + (\mu\varepsilon + \varepsilon^2) R,$$

where $*$ is the adjoint in $L^2(\mathbb{R})^2$ and R is linear operator such that for any $U \in H^{\alpha+4}(\mathbb{R}) \times H^{\alpha+4}(\mathbb{R})$, $|RU|_{H^\alpha \times H^\alpha} \leq C(\mu_{\max}, |r|_{W^{\alpha+4,1}}, |s|_{W^{\alpha+4,1}}) |U|_{H^{\alpha+4} \times H^{\alpha+4}}$.

Proof. We easily compute the Jacobian matrix of $\mathcal{T}_{\mathrm{B}}$:

$$\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)} = \begin{pmatrix} 1+\frac{\varepsilon}{4}s+\frac{\varepsilon}{4}\partial^{-1}(s)\partial_x(\circ) & \frac{\varepsilon}{4}\partial_x(r)\partial^{-1}(\circ)+\frac{\varepsilon}{4}(r+s) \\ \frac{\varepsilon}{4}\partial_x(s)\partial^{-1}(\circ)+\frac{\varepsilon}{4}(r+s) & 1+\frac{\varepsilon}{4}r+\frac{\varepsilon}{4}\partial^{-1}(r)\partial_x(\circ) \end{pmatrix}.$$

Its L^2 adjoint is

$$\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^* = \begin{pmatrix} 1+\frac{\varepsilon}{4}s-\frac{\varepsilon}{4}\partial_x(\partial^{-1}(s)\circ) & \frac{\varepsilon}{4}(r+s)-\frac{\varepsilon}{4}\partial^{-1}(\partial_x(s)\circ) \\ \frac{\varepsilon}{4}(r+s)-\frac{\varepsilon}{4}\partial^{-1}(\partial_x(r)\circ) & 1+\frac{\varepsilon}{4}r-\frac{\varepsilon}{4}\partial_x(\partial^{-1}(r)\circ). \end{pmatrix}$$

Using the fact that both $\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}$ and $\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^*$ can be written under the form $I_d+O(\varepsilon)$, we write

$$\begin{aligned} \Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big) J_\mu \Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^* &= \Big[\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}-I_d\Big)+I_d\Big] J_\mu \Big[\Big(\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^*-I_d\Big)+I_d\Big] \\ &= \Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}-I_d\Big) J_\mu \Big(\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^*-I_d\Big) \\ &+ \Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}-I_d\Big) J_\mu + J_\mu \Big(\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^*-I_d\Big) + J_\mu. \end{aligned}$$

The first term is of order $O(\varepsilon^2)$, we will estimate it later. The second and third terms are of order $O(\varepsilon)$. We prove now that their addition is of order $O(\mu\varepsilon)$. We easily compute both terms:

$$\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}-I_d\Big) J_\mu = \frac{\varepsilon}{8}\begin{pmatrix} -s\mathrm{F}_\mu\partial_x[\circ]-\partial^{-1}(s)\mathrm{F}_\mu\partial_x^2[\circ] & (r+s)\mathrm{F}_\mu\partial_x[\circ]+\partial_x(r)\mathrm{F}_\mu[\circ] \\ -(r+s)\mathrm{F}_\mu\partial_x[\circ]-\partial_x(s)\mathrm{F}_\mu[\circ] & r\mathrm{F}_\mu\partial_x[\circ]+\partial^{-1}(r)\mathrm{F}_\mu\partial_x^2[\circ] \end{pmatrix},$$

and

$$J_\mu\Big(\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^*-I_d\Big) = \frac{\varepsilon}{8}\begin{pmatrix} \mathrm{F}_\mu\partial_x[\partial^{-1}(s)\partial_x(\circ)] & -\mathrm{F}_\mu\partial_x[(r+s)\circ]+\mathrm{F}_\mu[\partial_x(s)\circ] \\ \mathrm{F}_\mu\partial_x[(r+s)\circ]-\mathrm{F}_\mu[\partial_x(r)\circ] & -\mathrm{F}_\mu\partial_x[\partial^{-1}(r)\partial_x(\circ)] \end{pmatrix}.$$

For now, we look at the first term of both matrices. For any $u\in H^{\alpha+4}(\mathbb{R})$

$$\begin{aligned} &|\mathrm{F}_\mu\partial_x[\partial^{-1}(s)\partial_x u]-s\mathrm{F}_\mu\partial_x[u]-\partial^{-1}(s)\mathrm{F}_\mu\partial_x^2[u]|_{H^\alpha} \\ \leq&|[\mathrm{F}_\mu,s]\partial_x u|_{H^\alpha}+|[\mathrm{F}_\mu,\partial^{-1}s]\partial_x^2 u|_{H^\alpha} \\ =&|[\mathrm{F}_\mu-1;s]\partial_x u|_{H^\alpha}+|[\mathrm{F}_\mu-1,\partial^{-1}s]\partial_x^2 u|_{H^\alpha}. \end{aligned}$$

But using Remark 2.7 and (3.14) we have

$$\begin{aligned} |[\mathrm{F}_\mu-1,\partial^{-1}s]\partial_x^2 u|_{H^\alpha} &\lesssim \mu|\partial^{-1}(s)\partial_x^2 u|_{H^{\alpha+2}}+|\partial^{-1}s|_{W^{\alpha+2,\infty}}|(\mathrm{F}_\mu-1)[\partial_x^2 u|_{H^\alpha} \\ &\lesssim \mu|\partial^{-1}s|_{W^{\alpha+2,\infty}}|\partial_x^2 u|_{H^{\alpha+2}} \\ &\lesssim \mu|s|_{W^{\alpha+2,1}}|u|_{H^{\alpha+4}}. \end{aligned}$$

Using algebra properties of $H^{\alpha+2}(\mathbb{R})$ and the Sobolev embedding $W^{\alpha+3,1}(\mathbb{R})\subset H^{\alpha+2}(\mathbb{R})$, we also have

$$|[\mathrm{F}_\mu-1,s]\partial_x u|_{H^\alpha}\lesssim \mu|s|_{H^{\alpha+2}}|u|_{H^{\alpha+3}}\lesssim \mu|s|_{W^{\alpha+3,1}}|u|_{H^{\alpha+3}}.$$

So that

$$|\mathrm{F}_\mu\partial_x[\partial^{-1}(s)\partial_x u] - s\mathrm{F}_\mu\partial_x[u] - \partial^{-1}(s)\mathrm{F}_\mu\partial_x^2[u]|_{H^\alpha} \lesssim \mu|s|_{W^{\alpha+3,1}}|u|_{H^{\alpha+4}}.$$

The same can be done for the addition of the fourth term of both matrices.

Now we look at the addition of the second term of both matrices. For any $u \in H^{\alpha+3}(\mathbb{R})$ we have

$$\begin{aligned}(r+s)\mathrm{F}_\mu[\partial_x u] + \partial_x(r)\mathrm{F}_\mu[u] &= \partial_x((r+s)\mathrm{F}_\mu[u]) - \partial_x(r+s)\mathrm{F}_\mu[u] \\ &= \partial_x((r+s)\mathrm{F}_\mu[u]) - \partial_x(s)\mathrm{F}_\mu[u].\end{aligned}$$

Using Lemma 2.6 and the Sobolev embedding $W^{\alpha+4,1}(\mathbb{R}) \subset H^{\alpha+3}(\mathbb{R})$, we get

$$\begin{aligned}&|\partial_x((r+s)\mathrm{F}_\mu[u]) - \partial_x(s)\mathrm{F}_\mu[u] - \mathrm{F}_\mu[\partial_x((r+s)u)] + \mathrm{F}_\mu[\partial_x(s)u]|_{H^\alpha} \\ \leq & |\partial_x([\mathrm{F}_\mu, r+s]u)|_{H^\alpha} + |[\mathrm{F}_\mu, \partial_x s]u|_{H^\alpha} \\ \leq & |[\mathrm{F}_\mu - 1, r+s]u|_{H^{\alpha+1}} + |[\mathrm{F}_\mu - 1, \partial_x s]u|_{H^\alpha} \\ \lesssim & \mu(|r|_{H^{\alpha+3}} + |s|_{H^{\alpha+3}})|u|_{H^{\alpha+3}} \\ \lesssim & \mu(|r|_{W^{\alpha+4,1}} + |s|_{W^{\alpha+4}})|u|_{H^{\alpha+3}}.\end{aligned}$$

The same can be done for the addition of the third term of both matrices. Thus we get for any $U \in H^{\alpha+4}(\mathbb{R}) \times H^{\alpha+4}(\mathbb{R})$

$$|\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)} - I_d\Big)J_\mu U + J_\mu\Big(\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^* - I_d\Big)U|_{H^\alpha} \lesssim \mu\varepsilon(|r|_{W^{\alpha+4,1}} + |s|_{W^{\alpha+4,1}})|U|_{H^{\alpha+4}}$$

It remains to estimate $\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)} - I_d\Big)J_\mu\Big(\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^* - I_d\Big)$. Using the boundedness of F_μ and the same tools as before, we get for any $U \in H^{\alpha+3}(\mathbb{R}) \times H^{\alpha+3}(\mathbb{R})$

$$|\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)} - I_d\Big)J_\mu\Big(\Big(\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}\Big)^* - I_d\Big)U|_{H^\alpha} \lesssim \varepsilon^2 C(|r|_{W^{\alpha+4,1}}, |r|_{W^{\alpha+4,1}})|U|_{H^{\alpha+3}}.$$

□

3.3 From the Hamiltonian of the Whitham-Boussinesq system under normal form to two decoupled Whitham equations

In the previous subsections, we proved the existence of a transformation $\mathcal{T}_{\mathrm{B}}$ such that

$$\begin{aligned}H_{\mathrm{WW}} \circ \mathcal{T}_{\mathrm{B}} &= L + \varepsilon Z + O(\mu\varepsilon + \varepsilon^2) \\ &= \int_{\mathbb{R}} (r^2 + s^2)\,\mathrm{dx} + \varepsilon\frac{1}{2}\int (r^3 + s^3)\,\mathrm{dx} + O(\mu\varepsilon + \varepsilon^2).\end{aligned}$$

The equations associated with the Hamiltonian $L + \varepsilon Z$ and the Poisson tensor J_μ are easily computed.

Property 3.20. *The Hamilton's equations associated with the normal form $L + \varepsilon Z$ are*

$$\begin{cases}\partial_t r + \mathrm{F}_\mu[\partial_x r] + \frac{3\varepsilon}{2}\mathrm{F}_\mu[r\partial_x r] = 0, \\ \partial_t s - \mathrm{F}_\mu[\partial_x s] - \frac{3\varepsilon}{2}\mathrm{F}_\mu[s\partial_x s] = 0.\end{cases} \tag{3.15}$$

Proof. The Hamilton's equations are

$$\partial_t \begin{pmatrix} r \\ s \end{pmatrix} = J_\mu \nabla (L + \varepsilon Z).$$

$\square$

Notation 3.21. *We write $H_{\mathrm{Wh}}(r,s) := L(r,s) + \varepsilon Z(r,s)$.*

Remark 3.22. *The two equations* (3.15) *are almost two Whitham equations. One just need to use that for any $\alpha \geq 0$ and any $u \in H^{\alpha+2}(\mathbb{R})$ one has $|(\mathrm{F}_\mu - 1)[u]|_{H^\alpha} \lesssim \mu |u|_{H^{\alpha+2}}$ to get the existence of a remainder (R_1, R_2) controllable in Sobolev spaces H^α such that*

$$\begin{cases} \partial_t r + \mathrm{F}_\mu[\partial_x r] + \frac{3\varepsilon}{2} r \partial_x r = \mu\varepsilon R_1, \\ \partial_t s - \mathrm{F}_\mu[\partial_x s] - \frac{3\varepsilon}{2} s \partial_x s = \mu\varepsilon R_2. \end{cases}$$

Lemma 3.23. *Let $\theta > 1/2$ and $\alpha \geq 0$. Then for any $r_0, s_0 \in W^{\alpha+5,1}(\mathbb{R})$ there exists a time $T > 0$ such that both equations* (3.15) *admit a unique solution $r, s \in \mathcal{C}([0, \frac{T}{\varepsilon}], H^{\alpha+4}(\mathbb{R}))$, with initial conditions r_0 and s_0, which satisfy for any times $t \in [0, \frac{T}{\epsilon}]$*

$$\begin{cases} |r(t)|_{W^{\alpha,1}} \lesssim (1+\mu t)^\theta C(\mu_{\max}, T, |r_0|_{W^{\alpha+5,1}}), \\ |s(t)|_{W^{\alpha,1}} \lesssim (1+\mu t)^\theta C(\mu_{\max}, T, |s_0|_{W^{\alpha+5,1}}). \end{cases}$$

In particular $r, s \in \mathcal{C}([0, \frac{T}{\varepsilon}], W^{\alpha,1}(\mathbb{R}))$ are uniformly bounded in (μ, ε) for times $t \in [0, \frac{T}{\max(\mu,\varepsilon)}]$ in $W^{\alpha,1}(\mathbb{R})$.

Moreover, for any $r_0, s_0 \in W^{\alpha+7,1}(\mathbb{R})$, both equations (3.15) *admit a unique solution $r, s \in \mathcal{C}^1([0, \frac{T}{\max(\mu,\varepsilon)}], W^{\alpha,1}(\mathbb{R})) \cap \mathcal{C}([0, \frac{T}{\max(\mu,\varepsilon)}], W^{\alpha+2,1}(\mathbb{R}))$ with initial conditions r_0 and s_0.*

Proof. We have the Sobolev embeddings $W^{\alpha+5,1}(\mathbb{R}) \subset H^{\alpha+4}(\mathbb{R})$ and the equations (3.15) are well-posed in these Sobolev spaces (see [6]). So there exists $T > 0$ such that the latter equations admit a unique solution in $\mathcal{C}([0, \frac{T}{\varepsilon}], H^{\alpha+2}(\mathbb{R}))$.

We do the rest of the reasoning for r. We use the Duhamel formula to say that

$$r(t) = e^{-itD\mathrm{F}_\mu} r_0 + \frac{3\varepsilon}{2} \int_0^t e^{-i(t-t')D\mathrm{F}_\mu} \mathrm{F}_\mu[r\partial_x r] \,\mathrm{d}t'\,.$$

We prove first that $|e^{-itD\mathrm{F}_\mu} r_0|_{W^{\alpha,1}} \lesssim (1+\mu t)^\theta |r_0|_{W^{\alpha+2,1}}$.

We begin by doing the same reasoning as in the proof of Property 3.15. We remark that we only need to prove the previous estimate in $L^1(\mathbb{R})$. Let $\chi \in \mathcal{C}^\infty((-2,2))$ be a cutoff function such that $\chi(\xi) = 1$ for $|\xi| \leq 1$. We decompose $e^{-itD\mathrm{F}_\mu} r_0 = e^{-itD(\mathrm{F}_\mu - 1)} e^{-itD} r_0$ into

$$\begin{aligned} e^{-itD(\mathrm{F}_\mu - 1)} e^{-itD} r_0(x) &= \mathcal{F}^{-1}[e^{-it\xi(F_\mu(\xi)-1)}\chi(\xi)] * (e^{-itD} r_0) \\ &\quad + \mathcal{F}^{-1}[e^{it\xi(F_\mu(\xi)-1)} \frac{1-\chi(\xi)}{(i\xi)^2}] * \partial_y^2 e^{-itD} r_0 := I_1 + I_2. \end{aligned}$$

- We prove here that $|I_1|_{L^1} \lesssim (1+\mu t)^{\frac{1}{2}+}|r_0|_{L^1}$. Using Young's inequality, we get

$$|I_1|_{L^1} \leq |\mathcal{F}^{-1}[e^{-it\xi(F_\mu(\xi)-1)}\chi(\xi)]|_{L^1}|e^{-itD}r_0|_{L^1} \leq |\mathcal{F}^{-1}[e^{-it\xi(F_\mu(\xi)-1)}\chi(\xi)]|_{L^1}|r_0|_{L^1}.$$

Now we pose $G(x) = \mathcal{F}^{-1}[e^{-it\xi(F_\mu(\xi)-1)}\chi(\xi)]$. Let $\beta \in \mathcal{C}^\infty((-2,2))$ be a cutoff function such that $\beta(x) = 1$ for $|x| \leq 1$. We have for any $\theta > 1/2$

$$\begin{aligned}|G|_{L^1} &\leq |\beta G|_{L^1} + |(1-\beta)G|_{L^1} \\ &\lesssim 1 + |\frac{1-\beta(x)}{x^\theta}|_{L^2}|x^\theta G|_{L^2}.\end{aligned}$$

We will prove the inequality $|xG|_{L^2} \lesssim 1+\mu t$ and interpolate it with the obvious one $|G|_{L^2} \lesssim 1$.

We remark that

$$|xG|_{L^2} = |\partial_\xi \widehat{G}|_{L^2} = |\partial_\xi(e^{-it\xi(F_\mu(\xi)-1)}\chi(\xi))|_{L^2}.$$

But we have the two following estimates on F_μ: for any $\xi \in \mathbb{R}$

$$|F_\mu(\xi) - 1| \lesssim \mu\xi^2, \quad |F'_\mu(\xi)| \lesssim \mu\xi^2. \tag{3.16}$$

So

$$|\partial_\xi(e^{-it\xi(F_\mu(\xi)-1)}\chi(\xi))|_{L^2} \leq t|(F_\mu(\xi)-1)\chi(\xi)|_{L^2} + t|\xi F'_\mu(\xi)\chi(\xi)|_{L^2} + |\chi'(\xi)|_{L^2} \lesssim 1+\mu t.$$

Then Hölder's inequality gives us

$$|x^\theta G|_{L^2} = |(xG)^\theta G^{1-\theta}|_{L^2} \leq |(xG)^\theta|_{L^p}|G^{1-\theta}|_{L^q},$$

with $\frac{1}{2} = \frac{1}{p} + \frac{1}{q}$. It remains to take $p = \frac{2}{\theta}$ and $q = \frac{2}{1-\theta}$ to get

$$|x^\theta G|_{L^2} \leq |xG|^\theta_{L^2}|G|^{1-\theta}_{L^2} \lesssim (1+\mu t)^\theta.$$

- We prove here that $|I_2|_{L^1} \lesssim (1+\mu t)^\theta |r_0|_{W^{2,1}}$. Again we use Young's inequality to get

$$|I_2|_{L^1} \leq |\mathcal{F}^{-1}[e^{it\xi(F_\mu(\xi)-1)}\frac{1-\chi(\xi)}{(i\xi)^2}]|_{L^1}|r_0|_{W^{2,1}}.$$

Then the rest of the proof is the same, with a little exception: here, using the estimates (3.16), we have

$$\begin{aligned}&|\partial_\xi(e^{-it\xi(F_\mu(\xi)-1)}\frac{1-\chi(\xi)}{\xi^2})|_{L^2} \\ \leq &|\frac{t(F_\mu(\xi)-1)(1-\chi(\xi))\xi^2 + t\xi F'_\mu(\xi)(1-\chi(\xi))\xi^2 - \chi'(\xi)\xi^2 - 2(1-\chi(\xi))\xi}{\xi^4}|_{L^2} \\ \lesssim &1+\mu t.\end{aligned}$$

Now we prove $|\int_0^t e^{-i(t-t')D\mathrm{F}_\mu}\mathrm{F}_\mu[r\partial_x r]\,\mathrm{d}t'\,|_{W^{\alpha,1}} \lesssim (1+\mu t)^\theta \int_0^t |r|_{H^{\alpha+4}}^2\,\mathrm{d}t'$, so that for all times $t \in [0, \frac{T}{\varepsilon}]$ one has

$$
\begin{aligned}
|\frac{3\varepsilon}{2}\int_0^t e^{-i(t-t')D\mathrm{F}_\mu}\mathrm{F}_\mu[r\partial_x r]\,\mathrm{d}t'\,|_{W^{\alpha,1}} &\leq (1+\mu t)^\theta C(\mu_{\max}, T, |r_0|_{H^{\alpha+4}}) \\
&\leq (1+\mu t)^\theta C(\mu_{\max}, T, |r_0|_{W^{\alpha+5,1}}).
\end{aligned}
$$

Using what we did before, we have

$$
\begin{aligned}
|\int_0^t e^{-i(t-t')D\mathrm{F}_\mu}\mathrm{F}_\mu[r\partial_x r]\,\mathrm{d}t'\,|_{W^{\alpha,1}} &\leq \int_0^t |e^{-i(t-t')D\mathrm{F}_\mu}\mathrm{F}_\mu[r\partial_x r]|_{W^{\alpha,1}}\,\mathrm{d}t' \\
&\lesssim \int_0^t (1+\mu(t-t'))^\theta |\mathrm{F}_\mu[r\partial_x r]|_{W^{\alpha+2,1}}\,\mathrm{d}t'\,.
\end{aligned}
$$

Using the continuity of F_μ from $W^{\alpha+3,1}(\mathbb{R})$ to $W^{\alpha+2,1}(\mathbb{R})$ (see Property 3.15), we get

$$
\begin{aligned}
|\int_0^t e^{-i(t-t')D\mathrm{F}_\mu}\mathrm{F}_\mu[r\partial_x r]\,\mathrm{d}t'\,|_{W^{\alpha,1}} &\lesssim (1+\mu t)^\theta \int_0^t |r\partial_x r|_{W^{\alpha+3,1}}\,\mathrm{d}t' \\
&\lesssim (1+\mu t)^\theta \int_0^t |r|_{H^{\alpha+4}}^2\,\mathrm{d}t'\,.
\end{aligned}
$$

It remains to prove that taking r_0 and s_0 in $W^{\alpha+7}(\mathbb{R})$, $\partial_t r$ and $\partial_t s$ are in $W^{\alpha,1}(\mathbb{R})$. We do the reasoning for r. From what we proved above, r is in $W^{\alpha+2,1}(\mathbb{R})$. It is also solution of the equation

$$
\partial_t r + \mathrm{F}_\mu[\partial_x r] + \frac{3\varepsilon}{2}\mathrm{F}_\mu[r\partial_x r] = 0.
$$

So that

$$
\begin{aligned}
|\partial_t r|_{W^{\alpha,1}} \lesssim& |\mathrm{F}_\mu[\partial_x r]|_{W^{\alpha,1}} + |\mathrm{F}_\mu[r\partial_x r]|_{W^{\alpha,1}} \\
\leq& |r|_{W^{\alpha+2,1}} + |r\partial_x r|_{W^{\alpha+1,1}} \lesssim |r|_{W^{\alpha+2,1}}^2.
\end{aligned}
$$

□

Theorem 3.24. *Let $\alpha \geq 0$. Let $r, s \in \mathcal{C}^1([0, \frac{T}{\varepsilon}], W^{\alpha+6,1}(\mathbb{R}))$ be solutions of the equations* (3.15)*. Let also $\mathcal{T}_\mathrm{B}$ be the transformation* (3.13)*, and define:*

$$
\begin{pmatrix} r_\mathrm{c} \\ s_\mathrm{c} \end{pmatrix} := \mathcal{T}_\mathrm{B}(r,s). \tag{3.17}
$$

Then, there exists $R \in \mathcal{C}^1([0, \frac{T}{\varepsilon}], H^\alpha(\mathbb{R}) \times H^\alpha)(\mathbb{R})$, such that one has:

$$
\partial_t \begin{pmatrix} r_\mathrm{c} \\ s_\mathrm{c} \end{pmatrix} = J_\mu \nabla(H_\mathrm{BW})(r_\mathrm{c}, s_\mathrm{c}) + (\mu\varepsilon + \varepsilon^2)R, \quad \forall t \in [0, \frac{T}{\varepsilon}],
$$

where H_BW is the Hamiltonian defined in (3.6) *and $|R|_{H^\alpha \times H^\alpha} \leq C(\mu_{\max}, |r|_{W^{\alpha+6,1}}, |s|_{W^{\alpha+6,1}})$. In particular, R is uniformly bounded in (μ, ε) for times $t \in [0, \frac{T}{\max(\mu,\varepsilon)}]$ in $H^\alpha(\mathbb{R}) \times H^\alpha(\mathbb{R})$.*

Proof. The existence of the solutions (r,s) is given by the previous lemma 3.23.

We write $\mathcal{T}_{\rm B}(r,s)=\begin{pmatrix}T_1(r,s)\\T_2(r,s)\end{pmatrix}=\begin{pmatrix}r+\frac{\varepsilon}{4}\partial_x r\partial^{-1}s+\frac{\varepsilon}{4}rs+\frac{\varepsilon}{8}s^2\\ s+\frac{\varepsilon}{4}\partial_x s\partial^{-1}r+\frac{\varepsilon}{4}rs+\frac{\varepsilon}{8}r^2\end{pmatrix}$. By definition of $H_{\rm BW}$ (see (3.6)) and $H_{\rm Wh}$ (see Notation 3.21), and the computations in the proof of Proposition 3.17 we have

$$\begin{aligned}
H_{\rm BW}(\mathcal{T}_{\rm B}(r,s)) &= \int_{\mathbb{R}}(T_1(r,s)^2+T_2(r,s)^2)\,\mathrm{dx}+\frac{\varepsilon}{2}\int(T_1(r,s)^3+T_2(r,s)^3)\,\mathrm{dx}\\
&\quad-\frac{\varepsilon}{2}\int_{\mathbb{R}}(T_1(r,s)^2T_2(r,s)+T_1(r,s)T_2(r,s)^2)\,\mathrm{dx}\\
&= H_{\rm Wh}(r,s)+\int_{\mathbb{R}}[(\frac{\varepsilon}{4}\partial_x(r)\partial^{-1}(s)+\frac{\varepsilon}{4}rs+\frac{\varepsilon}{8}s^2)^2+(\frac{\varepsilon}{4}\partial_x(s)\partial^{-1}(r)+\frac{\varepsilon}{4}rs+\frac{\varepsilon}{8}r^2)^2]\,\mathrm{dx}\\
&\quad+\frac{\varepsilon}{2}\int_{\mathbb{R}}(T_1(r,s)^3-r^3+T_2(r,s)^3-s^3)\,\mathrm{dx}\\
&\quad-\frac{\varepsilon}{2}\int_{\mathbb{R}}(T_1(r,s)^2T_2(r,s)-r^2s+T_1(r,s)T_2(r,s)^2-rs^2)\,\mathrm{dx}
\end{aligned}$$

So $H_{\rm BW}(\mathcal{T}_{\rm B}(r,s))=H_{\rm Wh}(r,s)+\varepsilon^2\int_{\mathbb{R}}p(r,s,\partial_x(r)\partial^{-1}(s),\partial_x(s)\partial^{-1}(r))\,\mathrm{dx}$, where $p:\mathbb{R}^4\to\mathbb{R}$ is polynomial.

But $(r_{\rm c},s_{\rm c})=\mathcal{T}_{\rm B}(r,s)$. So

$$\partial_t\begin{pmatrix}r_{\rm c}\\s_{\rm c}\end{pmatrix}=\frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}\partial_t\begin{pmatrix}r\\s\end{pmatrix},$$

where $\frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}$ is the Jacobian matrix of $\mathcal{T}_{\rm B}(r,s)$ computed in the proof of Proposition 3.19. Since r and s are solutions of the Hamilton's equations associated with $H_{\rm Wh}$, we have

$$\begin{aligned}
\partial_t\begin{pmatrix}r_{\rm c}\\s_{\rm c}\end{pmatrix} &= \frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}J_\mu\nabla(H_{\rm Wh})(r,s)\\
&= \frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}J_\mu\nabla(H_{\rm BW}(\mathcal{T}_{\rm B}(r,s))-\varepsilon^2\frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}J_\mu\nabla\int_{\mathbb{R}}p(r,s,\partial_x r\partial^{-1}s,\partial_x s\partial^{-1}r)\,\mathrm{dx}\\
&= \frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}J_\mu\Big(\frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}\Big)^*\nabla(H_{\rm BW})(r_{\rm c},s_{\rm c})-\varepsilon^2\frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}J_\mu\nabla\int_{\mathbb{R}}p(r,s,\partial_x r\partial^{-1}s,\partial_x s\partial^{-1}r)\,\mathrm{dx}\,.
\end{aligned}$$

Moreover by Proposition 3.19, we know that $\frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}J_\mu\Big(\frac{\partial\mathcal{T}_{\rm B}}{\partial(r,s)}\Big)^*=J_\mu+(\mu\varepsilon+\varepsilon^2)R$, with $|RU|_{H^\alpha\times H^\alpha}\leq C(\mu_{\max},|r|_{W^{\alpha+4,1}},|s|_{W^{\alpha+4,1}})|U|_{H^{\alpha+4}}$ for any $U\in H^{\alpha+4}(\mathbb{R})\times H^{\alpha+4}(\mathbb{R})$. So, here, we need to estimate $\nabla(H_{\rm BW})(r_{\rm c},s_{\rm c})$ in $H^{\alpha+4}(\mathbb{R})\times H^{\alpha+4}(\mathbb{R})$.

$$\nabla(H_{\rm BW})(\mathcal{T}_{\rm B}(r,s))=\begin{pmatrix}2T_1(r,s)+\frac{3\varepsilon}{2}T_1(r,s)^2-\varepsilon T_1(r,s)T_2(r,s)-\frac{\varepsilon}{2}T_2(r,s)^2\\ 2T_2(r,s)+\frac{3\varepsilon}{2}T_2(r,s)^2-\varepsilon T_1(r,s)T_2(r,s)-\frac{\varepsilon}{2}T_1(r,s)^2\end{pmatrix}.$$

Combining the algebra properties of $H^{\alpha+4}(\mathbb{R})$ with the estimates

$$\begin{cases}|\partial_x(r)\partial^{-1}(s)|_{H^{\alpha+4}}\lesssim|\partial^{-1}s|_{W^{\alpha+4,\infty}}|\partial_x r|_{H^{\alpha+4}}\lesssim|s|_{W^{\alpha+4,1}}|r|_{H^{\alpha+5}}\lesssim|s|_{W^{\alpha+4,1}}|r|_{W^{\alpha+6,1}},\\ |\partial_x(s)\partial^{-1}(r)|_{H^{\alpha+4}}\lesssim|r|_{W^{\alpha+4,1}}|s|_{W^{\alpha+6,1}}\end{cases}$$

(see (3.14)), we have

$$|\nabla(H_{\rm BW})(\mathcal{T}_{\rm B}(r,s))|_{H^{\alpha+4}}\leq C(||r|_{W^{\alpha+6,1}},|s|_{W^{\alpha+6,1}}).$$

It remains to prove that $\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)} J_\mu \nabla \int_{\mathbb{R}} p(r,s,\partial_x(r)\partial^{-1}(s),\partial_x(s)\partial^{-1}(r))\,\mathrm{dx}$ is in $\mathcal{C}^1([0,\frac{T}{\varepsilon}],H^\alpha(\mathbb{R})\times H^\alpha(\mathbb{R}))$.

The polynomial p is a linear combination of terms of the form $r^{n_1}s^{n_2}(\partial_x(r)\partial^{-1}(s))^{n_3}(\partial_x(s)\partial^{-1}(r))^{n_4}$ where n_1,n_2,n_3 and n_4 are non-negative integers such that $n_1+n_2+n_3+n_4\geq 2$. Moreover

$$J_\mu \nabla \int_{\mathbb{R}} r^{n_1}s^{n_2}(\partial_x(r)\partial^{-1}(s))^{n_3}(\partial_x(s)\partial^{-1}(r))^{n_4}\,\mathrm{dx} = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix},$$

where

$$\begin{cases} M_1 = & -\frac{1}{2}\mathrm{F}_\mu\Big[\partial_x\Big(n_1 r^{n_1-1}s^{n_2}(\partial_x(r)\partial^{-1}(s))^{n_3}(\partial_x(s)\partial^{-1}(r))^{n_4}\Big)\Big] \\ & +\frac{1}{2}\mathrm{F}_\mu\Big[\partial_x^2\Big(r^{n_1}s^{n_2}n_3(\partial_x r)^{n_3-1}(\partial^{-1}s)^{n_3}(\partial_x(s)\partial^{-1}(r))^{n_4}\Big)\Big] \\ & +\frac{1}{2}\mathrm{F}_\mu\Big[r^{n_1}s^{n_2}(\partial_x(r)\partial^{-1}(s))^{n_3}(\partial_x s)^{n_4}n_4(\partial^{-1}r)^{n_4-1}\Big], \\ M_2 = & \frac{1}{2}\mathrm{F}_\mu\Big[\partial_x\Big(r^{n_1}n_2 s^{n_2-1}(\partial_x(r)\partial^{-1}(s))^{n_3}(\partial_x(s)\partial^{-1}(r))^{n_4}\Big)\Big] \\ & -\frac{1}{2}\mathrm{F}_\mu\Big[r^{n_1}s^{n_2}(\partial_x r)^{n_3}n_3(\partial^{-1}s)^{n_3-1}(\partial_x(s)\partial^{-1}(r))^{n_4}\Big] \\ & -\frac{1}{2}\mathrm{F}_\mu\Big[\partial_x^2\Big(r^{n_1}s^{n_2}(\partial_x(r)\partial^{-1}(s))^{n_3}n_4(\partial_x s)^{n_4-1}(\partial^{-1}r)^{n_4}\Big)\Big], \end{cases}$$

where we used $\partial_x\partial^{-1}u = u$ when $u\in L^1(\mathbb{R})$ (see Definition/Property 3.12). The property $n_1+n_2+n_3+n_4\geq 2$ implies that the terms containing $\partial^{-1}s$ or $\partial^{-1}r$ are always multiplied by a function in a Sobolev space $W^{\beta,1}(\mathbb{R})$ with $\beta\geq 0$.

We recall now the expression of $\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)}$:

$$\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)} = \begin{pmatrix} 1+\frac{\varepsilon}{4}s+\frac{\varepsilon}{4}\partial^{-1}(s)\partial_x(\circ) & \frac{\varepsilon}{4}\partial_x(r)\partial^{-1}(\circ)+\frac{\varepsilon}{4}(r+s) \\ \frac{\varepsilon}{4}\partial_x(s)\partial^{-1}(\circ)+\frac{\varepsilon}{4}(r+s) & 1+\frac{\varepsilon}{4}r+\frac{\varepsilon}{4}\partial^{-1}(r)\partial_x(\circ) \end{pmatrix}.$$

From the above computations we get

$$\frac{\partial \mathcal{T}_{\mathrm{B}}}{\partial(r,s)} J_\mu \nabla \int_{\mathbb{R}} p(r,s,\partial_x(r)\partial^{-1}(s),\partial_x(s)\partial^{-1}(r))\,\mathrm{dx} \tag{3.18}$$

$$= \begin{pmatrix} M_1+\frac{\varepsilon}{4}sM_1+\frac{\varepsilon}{4}\partial^{-1}(s)\partial_x(M_1)+\frac{\varepsilon}{4}\partial_x(r)\partial^{-1}(M_2)+\frac{\varepsilon}{4}(r+s)M_2 \\ \frac{\varepsilon}{4}\partial_x(s)\partial^{-1}(M_1)+\frac{\varepsilon}{4}(r+s)M_1+M_2+\frac{\varepsilon}{4}rM_2+\frac{\varepsilon}{4}\partial^{-1}(r)\partial_x(M_2) \end{pmatrix}, \tag{3.19}$$

We estimate the terms thus obtained. Using product estimates A.1 and (3.14) we have

$$\begin{cases} |sM_1|_{H^\alpha} \lesssim |s|_{H^{\alpha+1}}|M_1|_{H^\alpha}, \\ |\partial^{-1}(s)\partial_x(M_1)|_{H^\alpha} \lesssim |\partial^{-1}s|_{W^{\alpha,\infty}}|M_1|_{H^{\alpha+1}} \lesssim |s|_{W^{\alpha,1}}|M_1|_{H^{\alpha+1}}, \\ |\partial_x(r)\partial^{-1}(M_1)|_{H^\alpha} \lesssim |r|_{H^{\alpha+1}}|\partial^{-1}M_1|_{W^{\alpha,\infty}} \lesssim |r|_{W^{\alpha+2,1}}|M_1|_{W^{\alpha,1}}. \end{cases}$$

Idem for the terms with M_2. So that it only remains to estimate M_1 and M_2. We do it for M_1. Using product estimates A.1 and $n_1+n_2+n_3+n_4\geq 2$, we get

$$|M_1|_{H^\alpha} \lesssim |\partial^{-1}r|^{N_1}_{W^{\alpha+2,\infty}}|\partial^{-1}s|^{N_2}_{W^{\alpha+2,\infty}}|r|^{N_3}_{H^{\alpha+3}}|s|^{N_4}_{H^{\alpha+3}},$$

where N_1,N_2,N_3 and N_4 are non negative integers such that $(N_3,N_4)\neq(0,0)$. Then, using (3.14) and Sobolev embeddings, we have

$$|M_1|_{H^\alpha} \lesssim |r|^{N_1+N_3}_{W^{\alpha+4,1}}|s|^{N_2+N_4}_{W^{\alpha+4,1}}.$$

Moreover, using the continuity of F_μ from $W^{\alpha+1,1}(\mathbb{R})$ to $W^{\alpha,1}(\mathbb{R})$ and the algebra properties of $W^{\beta,1}(\mathbb{R})$ for any $\beta \geq 1$, we get

$$|M_1|_{W^{\alpha,1}} \lesssim |\partial^{-1} r|^{N_1}_{W^{\alpha+3,\infty}} |\partial^{-1} s|^{N_2}_{W^{\alpha+3,\infty}} |r|^{N_3}_{W^{\alpha+4,1}} |s|^{N_4}_{W^{\alpha+4,1}} \lesssim |r|^{N_1+N_3}_{W^{\alpha+4,1}} |s|^{N_2+N_4}_{W^{\alpha+4,1}}.$$

$\square$

Combining Proposition 3.2 and Theorem 3.24 we get Theorem 1.14 which we recall here for the sake of clarity.

Theorem 3.25. *Let $\alpha \geq 0$. Let $r, s \in \mathcal{C}^1([0, \frac{T}{\varepsilon}], W^{\alpha+7,1}(\mathbb{R}))$ be solutions of the equations* (3.15)*. Let also $\mathcal{T}_\mathrm{I}, \mathcal{T}_\mathrm{D}$ and $\mathcal{T}_B$ be the transformations of Definition 1.13 and define:*

$$\begin{pmatrix} \zeta_{\mathrm{Wh}} \\ \psi_{\mathrm{Wh}} \end{pmatrix} := \mathcal{T}_\mathrm{I}(\mathcal{T}_\mathrm{D}(\mathcal{T}_\mathrm{B}(r, s))). \tag{3.20}$$

Then, there exists $R \in \mathcal{C}^1([0, \frac{T}{\varepsilon}], H^\alpha(\mathbb{R}) \times H^\alpha(\mathbb{R}))$, such that one has:

$$\partial_t \begin{pmatrix} \zeta_{\mathrm{Wh}} \\ \psi_{\mathrm{Wh}} \end{pmatrix} = J\nabla(H_{\mathrm{WW}})(\zeta_{\mathrm{Wh}}, \psi_{\mathrm{Wh}}) + (\mu\varepsilon + \varepsilon^2)R, \quad \forall t \in [0, \frac{T}{\varepsilon}],$$

where H_{WW} is the Hamiltonian defined in (1.2) *and $|R|_{H^\alpha \times \dot{H}^{\alpha+1}} \leq C(\mu_{\max}, |r|_{W^{\alpha+7,1}}, |s|_{W^{\alpha+7,1}})$. In particular, R is uniformly bounded in (μ, ε) for times $t \in [0, \frac{T}{\max(\mu,\varepsilon)}]$ in $H^\alpha(\mathbb{R}) \times \dot{H}^{\alpha+1}(\mathbb{R})$.*

Proof. We start by defining the quantities

$$\begin{pmatrix} \zeta_{\mathrm{Wh}} \\ v_{\mathrm{Wh}} \end{pmatrix} := \mathcal{T}_\mathrm{D}(\mathcal{T}_\mathrm{B}(r, s)). \tag{3.21}$$

Differentiating (3.21) in time and using Theorem 3.24 we get

$$\begin{aligned} \partial_t \begin{pmatrix} \zeta_{\mathrm{Wh}} \\ v_{\mathrm{Wh}} \end{pmatrix} &= \frac{\partial \mathcal{T}_\mathrm{D}}{\partial(r,s)}(\mathcal{T}_\mathrm{B}(r,s)) \frac{\partial \mathcal{T}_\mathrm{B}}{\partial(r,s)} \partial_t \begin{pmatrix} r \\ s \end{pmatrix} \\ &= \frac{\partial \mathcal{T}_\mathrm{D}}{\partial(r,s)}(\mathcal{T}_\mathrm{B}(r,s)) J_\mu \nabla(H_{\mathrm{BW}})(\mathcal{T}_\mathrm{B}(r,s)) + (\mu\varepsilon + \varepsilon^2) \frac{\partial \mathcal{T}_\mathrm{D}}{\partial(r,s)}(\mathcal{T}_\mathrm{B}(r,s)) R_1, \end{aligned}$$

where $\frac{\partial \mathcal{T}_\mathrm{D}}{\partial(r,s)}(\mathcal{T}_\mathrm{B}(r,s)) = \begin{pmatrix} 1 & 1 \\ \mathrm{F}_\mu^{-1}[\circ] & -\mathrm{F}_\mu^{-1}[\circ] \end{pmatrix}$ and R_1 is the rest in Theorem 3.24. Using now Proposition 3.6 we get

$$\begin{aligned} \partial_t \begin{pmatrix} \zeta_{\mathrm{Wh}} \\ v_{\mathrm{Wh}} \end{pmatrix} &= \widetilde{J}\nabla(\widetilde{H}_0 + \varepsilon \widetilde{H}_1)(\mathcal{T}_\mathrm{D}(\mathcal{T}_\mathrm{B}(r,s)) + (\mu\varepsilon + \varepsilon^2)R_2 \\ &= \widetilde{J}\nabla(\widetilde{H}_0 + \varepsilon \widetilde{H}_1)(\zeta_{\mathrm{Wh}}, v_{\mathrm{Wh}}) + (\mu\varepsilon + \varepsilon^2)R_2, \end{aligned}$$

where $\widetilde{J} = \begin{pmatrix} 0 & -\partial_x \\ -\partial_x & 0 \end{pmatrix}$ and $|R_2|_{H^\alpha \times H^\alpha} \leq C(|r|_{W^{\alpha+7,1}}, |s|_{W^{\alpha+7,1}})$ (because for any $\beta \geq 0$ and any $u \in W^{\beta+1,1}(\mathbb{R})$, $|\mathrm{F}_\mu^{-1}[u]|_{H^\beta} \lesssim |u|_{H^{\beta+1/2}} \lesssim |u|_{W^{\beta+1}}$). Changing the unknowns into $(\zeta_{\mathrm{Wh}}, \psi_{\mathrm{Wh}})$ we obtain

$$\partial_t \begin{pmatrix} \zeta_{\mathrm{Wh}} \\ \psi_{\mathrm{Wh}} \end{pmatrix} = J\nabla(H_0 + \varepsilon H_1)(\zeta_{\mathrm{Wh}}, \psi_{\mathrm{Wh}}) + (\mu\varepsilon + \varepsilon^2)R_3,$$

where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $|R_3|_{H^\alpha \times \dot{H}^{\alpha+1}} \leq C(|r|_{W^{\alpha+7,1}}, |s|_{W^{\alpha+7,1}})$ So that

$$\begin{aligned}\partial_t \begin{pmatrix} \zeta_{\mathrm{Wh}} \\ \psi_{\mathrm{Wh}} \end{pmatrix} &= J\nabla(H_{\mathrm{WW}})(\zeta_{\mathrm{Wh}}, \psi_{\mathrm{Wh}}) + (J\nabla(H_0 + \varepsilon H_1)(\zeta_{\mathrm{Wh}}, \psi_{\mathrm{Wh}}) - J\nabla(H_{\mathrm{WW}})(\zeta_{\mathrm{Wh}}, \psi_{\mathrm{Wh}})) \\ &\quad + (\mu\varepsilon + \varepsilon^2)R_2.\end{aligned}$$

Using Proposition 3.2 we get the result. The uniform boundedness in (μ, ε) for times $t \in [0, \frac{T}{\max(\mu,\varepsilon)}]$ of the remainder comes from the lemma 3.23. □

It only remains to prove Corollary 1.16.

Proof. We construct (r_0, s_0) such that

$$|\mathcal{T}_{\mathrm{I}}(\mathcal{T}_{\mathrm{D}}(\mathcal{T}_{\mathrm{B}}(r_0, s_0))) - (\zeta_0, \psi_0)|_{H^\alpha \times \dot{H}^{\alpha+1}} \lesssim \varepsilon^2. \tag{3.22}$$

We remark first that the transformations $\mathcal{T}_{\mathrm{I}}$ and $\mathcal{T}_{\mathrm{D}}$ are invertible with, for any $\beta \geq 0$, $\mathcal{T}_{\mathrm{I}}^{-1} : W^{\beta,1}(\mathbb{R}) \times \dot{W}^{\beta+1,1}(\mathbb{R}) \to W^{\beta,1}(\mathbb{R}) \times W^{\beta,1}(\mathbb{R})$ and $\mathcal{T}_{\mathrm{D}}^{-1} : W^{\beta+1,1}(\mathbb{R}) \times W^{\beta+1,1}(\mathbb{R}) \to W^{\beta,1}(\mathbb{R}) \times W^{\beta,1}(\mathbb{R})$ due to the continuity of F_μ from $W^{\beta+1,1}(\mathbb{R})$ to $W^{\beta,1}(\mathbb{R})$. Moreover, by definition (see (3.13)), we can write $\mathcal{T}_{\mathrm{B}}$ under the form

$$\mathcal{T}_{\mathrm{B}} = \mathrm{Id} + \varepsilon\widetilde{\mathcal{T}_{\mathrm{B}}},$$

where Id is the identity and for any $\beta \geq 0$, $\widetilde{\mathcal{T}_{\mathrm{B}}} : W^{\beta+1,1}(\mathbb{R}) \times W^{\beta+1,1}(\mathbb{R}) \to W^{\beta,1}(\mathbb{R}) \times W^{\beta,1}(\mathbb{R})$. We define the transformation $\mathcal{T}_{\mathrm{B}}^{\mathrm{inv}} : W^{\beta+1,1}(\mathbb{R}) \times W^{\beta+1,1}(\mathbb{R}) \to W^{\beta,1}(\mathbb{R}) \times W^{\beta,1}(\mathbb{R})$ by

$$\mathcal{T}_{\mathrm{B}}^{\mathrm{inv}} = \mathrm{Id} - \varepsilon\widetilde{\mathcal{T}_{\mathrm{B}}}.$$

One has for any $(\eta, w) \in W^{\beta+2,1}(\mathbb{R}) \times W^{\beta+2,1}(\mathbb{R})$, $\mathcal{T}_{\mathrm{B}}(\mathcal{T}_{\mathrm{B}}^{\mathrm{inv}}(\eta, w)) = \begin{pmatrix} \eta \\ w \end{pmatrix} + \varepsilon^2 R$ where $|R|_{W^{\beta,1} \times W^{\beta,1}} \leq C(|\eta|_{W^{\beta+2,1}}, |w|_{W^{\beta+2,1}})$. So that

$$(r_0, s_0) = \mathcal{T}_{\mathrm{B}}^{\mathrm{inv}}(\mathcal{T}_{\mathrm{D}}^{-1}(\mathcal{T}_{\mathrm{I}}^{-1}(\zeta_0, \psi_0))) \tag{3.23}$$

satisfy (3.22). □

A Technical tools

Proposition A.1. *(Product estimates)*

1. *Let $t_0 > 1/2$, $s \geq -t_0$ and $f \in H^\alpha \cap H^{t_0}(\mathbb{R}), g \in H^\alpha(\mathbb{R})$. Then $fg \in H^\alpha(\mathbb{R})$ and*

$$|fg|_{H^\alpha} \lesssim |f|_{H^{\max(t_0,\alpha)}} |g|_{H^\alpha}$$

2. *Let $\alpha_1, \alpha_2 \in \mathbb{R}$ be such that $\alpha_1 + \alpha_2 \geq 0$. Then for all $\alpha \leq \alpha_j$ ($j = 1, 2$) and $\alpha < \alpha_1 + \alpha_2 - 1/2$, and all $f \in H^{\alpha_1}(\mathbb{R}), g \in H^{\alpha_2}(\mathbb{R})$, one has $fg \in H^\alpha(\mathbb{R})$ and*

$$|fg|_{H^\alpha} \lesssim |f|_{H^{\alpha_1}} |g|_{H^{\alpha_2}}$$

Proof. See Appendix B.1 in [13]. □

Proposition A.2. *(Composition estimates) Let $G : \mathbb{R} \to \mathbb{R}$ be a smooth function such that $G(0) = 0$. Also, let $t_0 > 1/2$, $\alpha \geq 0$ and $f \in H^{\max(t_0,\alpha)}(\mathbb{R})$. Then $G(f) \in H^\alpha(\mathbb{R})$ and*

$$|G(f)|_{H^\alpha} \leq C(|f|_{H^{\max(t_0,\alpha)}})|f|_{H^\alpha}.$$

Proof. See Appendix B.1 in [13]. □

Proposition A.3. *(Quotient estimates) Let $t_0 > 1/2, \alpha \geq -t_0$ and $c_0 > 0$. Also let $f \in H^\alpha(\mathbb{R})$ and $g \in H^\alpha \cap H^{t_0}(\mathbb{R})$ be such that for all $x \in \mathbb{R}$, one has $1 + g(X) \geq c_0$. Then $\frac{f}{1+g}$ belongs to $H^\alpha(\mathbb{R})$ and*

$$|\frac{f}{1+g}|_{H^\alpha} \leq C(\frac{1}{c_0}, |g|_{H^{\max(t_0,\alpha)}})|f|_{H^\alpha}$$

Proof. See Appendix B.1 in [13]. □

Definition A.4. *We say that a Fourier multiplier* $\mathrm{F}(D)$ *is of order* α $(\alpha \in \mathbb{R})$ *and write* $\mathrm{F} \in \mathcal{S}^\alpha$ *if $\xi \in \mathbb{R} \mapsto F(\xi) \in \mathbb{C}$ is smooth and satisfies*

$$\forall \xi \in \mathbb{R}, \forall \beta \in \mathbb{N}, \quad \sup_{\xi \in \mathbb{R}} \langle \xi \rangle^{\beta - \alpha} |\partial^\beta F(\xi)| < \infty.$$

We also introduce the seminorm

$$\mathcal{N}^\alpha(F) = \sup_{\beta \in \mathbb{N}, \beta \leq 4} \sup_{\xi \in \mathbb{R}} \langle \xi \rangle^{\beta - \alpha} |\partial^\beta F(\xi)|.$$

Proposition A.5. *Let $t_0 > 1/2$, $\alpha \geq 0$ and $F \in \mathcal{S}^\alpha$. If $f \in H^\alpha \cap H^{t_0+1}$ then, for all $g \in H^{\alpha-1}$,*

$$|[F(D), f]g|_2 \leq \mathcal{N}^\alpha(F)|f|_{H^{\max(t_0+1,\alpha)}}|g|_{H^{\alpha-1}}.$$

Proof. See Appendix B.2 in [13] for a proof of this proposition. □

Proposition A.6. *Let $\alpha \geq 2$. Let $\zeta \in H^{\alpha+2}(\mathbb{R})$ be such that* (1.3) *is satisfied and $\psi \in \dot{H}^{\alpha+2}(\mathbb{R})$. Then one has*

$$|\frac{1}{\mu}\mathcal{G}^\mu[\varepsilon\zeta]\psi|_{H^\alpha} \leq M(s+2)|\psi|_{\dot{H}^{\alpha+2}}.$$

Proof. This is a direct consequence of Theorem 3.15 in [13]. □

References

[1] Dario Bambusi. Hamiltonian studies on counter-propagating water waves. *Water Waves*, 2020.

[2] Jerry L. Bona, Thierry Colin, and David Lannes. Long wave approximations for water waves. *Arch. Ration. Mech. Anal.*, 178(3):373–410, 2005.

[3] Haïm Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer-Verlag New York, 2011.

[4] Aynur Bulut. An optimal decay estimate for the linearized water wave equation in 2d. *Proceedings of the American Mathematical Society*, 144:4733–4742, 2016.

[5] Evgueni Dinvay, Sigmund Selberg, and Achenef Tesfahun. Well-posedness for a dispersive system of the Whitham-Boussinesq type. arXiv:1902.09438v3, 2019.

[6] Mats Ehrnström, Joachim Escher, and Long Pei. A note on the local well-posedness for the Whitham equation. In *Elliptic and parabolic equations*, volume 119 of *Springer Proc. Math. Stat.*, pages 63–75. Springer, Cham, 2015.

[7] Mats Ehrnström and Henrik Kalisch. Traveling waves for the Whitham equation. *Differential Integral Equations*, 22(11-12):1193–1210, 2009.

[8] Mats Ehrnström and Erik Wahlén. On Whitham's conjecture of a highest cusped wave for a nonlocal dispersive equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 36(6):1603–1637, 2019.

[9] Mats Ehrnström and Yuexun Wang. Enhanced existence time of solutions to evolution equations of Whitham type. arXiv:2008.12722, 2020.

[10] Louis Emerald. Rigorous derivation from the water waves equations of some full dispersion shallow water models. arXiv:2004.09240, 2020.

[11] Vera Mikyoung Hur. Wave breaking in the Whitham equation. *Adv. Math.*, 317:410–437, 2017.

[12] Christian Klein, Felipe Linares, Didier Pilod, and Jean-Claude Saut. On Whitham and related equations. *Stud. Appl. Math.*, 140(2):133–177, 2018.

[13] Lannes. *The water waves problem: mathematical analysis and asymptotics*. Mathematical surveys and monographs; volume 188. American Mathematical Society, Rhode Island, United-States, 2013.

[14] Guy Métivier. *Para-differential calculus and applications to the Cauchy problem for nonlinear systems*, volume 5 of *Centro di Ricerca Matematica Ennio De Giorgi (CRM) Series*. Edizioni della Normale, Pisa, 2008.

[15] Daulet Moldabayev, Henrik Kalisch, and Denys Dutykh. The Whitham equation as a model for surface waves. *Physica D: Nonlinear Phenomena*, 309:99–107, 2015.

[16] Nathan Sanford, Keri Kodama, John D. Carter, and Henrik Kalisch. Stability of traveling wave solutions to the Whitham equation. *Phys. Lett. A*, 378(30-31):2100–2107, 2014.

[17] Jean-Claude Saut and Yuexun Wang. The wave breaking for Whitham-type equations revisited. arXiv:2006.03803, 2020.

[18] Guido Schneider and C. Eugene Wayne. Corrigendum: The long-wave limit for the water wave problem I. The case of zero surface tension [mr1780702]. *Comm. Pure Appl. Math.*, 65(5):587–591, 2000.

[19] Tien Truong, Erik Wahlén, and H. Miles Wheeler. Global bifurcation of solitary waves for the Whitham equation. arXiv:2009.04713, 2020.

[20] G. B. Whitham. Variational methods and applications to water waves. *Royal Society, London*, 299:6–25, 1967.

[21] G. B. Whitham. *Linear and Nonlinear Waves.* Wiley, New York, 1974.

[22] V. E. Zakharov. Stability of periodic waves of finite amplitude on the surface of a deep fluid. *J. Appl. Mech. Tech. Phys.*, 9:190–194, 1968.