Hamiltonian studies on counter-propagating water waves

Dario Bambusi*

Abstract

We use a Hamiltonian normal form approach to study the dynamics of the water wave problem in the small amplitude long wave regime (KdV regime). If μ is the small parameter corresponding to the inverse of the wave length, we show that the normal form at order μ^5 consists of two decoupled equation, one describing right going waves and the other describing left going waves. Each of these equations is integrable: it is a linear combination of the first three equations in the KdV hierarchy. At order μ^7 we find nontrivial terms coupling the two counter-propagating waves.

Keywords: Gravity waves, KdV, Hamiltonian partial differential equations, normal form

1 Introduction

In this paper we study the dynamics of the free surface of a fluid which evolves under the influence of gravitation. The aim is to find the effective equation governing the dynamics in the regime of small amplitude and long wave. It is well known that, at the first nontrivial order, the effective equation is the Kortweg de Vries equation; more precisely, the dynamics is described by two KdV equations [SW00], one describing right going waves and the other describing left going waves, moreover the two counter-propagating waves do not interact, at least at the order of approximation controlled by KdV.

Here starting from the so called Zakharov-Craig-Sulem Hamiltonian approach to the water wave dynamics [Zak68, CG94, CS93] we use Birkhoff Normal form theory in order to attack the problem. As a first result we get

*Dipartimento di Matematica, Università degli Studi di Milano, Via Saldini 50, I-20133 Milano. *Email:* `dario.bambusi@unimi.it`

that the two decoupled KdV mentioned above are just the Hamilton equations of the first order Birkhoff Normal Form of the system. More generally, it turns out that *at any order*, the normal form of the system consists just of two decoupled equations, one describing right going waves and the other describing left going waves. The problem is that, in order to put the system in normal form, one has to construct a canonical transformation conjugating the original Hamiltonian to its normal form, and the existence of such a transformation is not ensured by any known general argument. So, we investigate the existence of the normalizing transformation; we prove that the transformation putting the system in second order normal form exists, while we find an obstruction to the existence of the transformation putting the system in third order normal form. To be slightly more precise, let μ be a small parameter, and consider an initial datum of size of order μ^2 and wave length of order μ^{-1}, then KdV is the normal form at order μ^3; we show that the system can be put in normal form at order μ^5 and that there is an obstruction to put the system in normal form at order μ^7.

So we stop our Hamiltonian construction at order 5 and analyze the equation that we get. It turns out that this equation falls in a class analyzed by Kodama (see [Kod85, Kod87b, Kod87a, HK09]), who showed that there always exists a non-Hamiltonian transformation conjugating, at order μ^5, such an equation to a linear combination of the first three equations of the KdV hierarchy. Remarkably enough, this is not true at order μ^7. Thus we apply Kodama's result getting that, up to order μ^7, counterpropagating waves are described by *two decoupled non interacting equations, each of which is an integrable equation* which is a linear combination of the first three equations in the KdV hierarchy.

We emphasize that the idea of using the Hamiltonian approach to show the appearance of KdV in water wave theory appeared in [CG94], where Craig and Groves made an expansion of the Hamiltonian in powers of the parameter μ (the one we just introduced) and then studied the first terms of the so obtained Hamiltonian in order to find the effective equations. A fundamental step in their procedure (a step which plays a crucial role also in the present paper) consists in parametrizing the surface of the fluid using suitable functions $r(y,t)$, $s(y,t)$, where t is a rescaled time variable and y is a rescaled space variable. Then the equations of motion of the unperturbed system turn out to be given simply by

$$\frac{\partial r}{\partial t} = -\frac{\partial r}{\partial y} , \quad \frac{\partial s}{\partial t} = \frac{\partial s}{\partial y} , \tag{1.1}$$

whose solution is of course a right going wave non interacting with a left going wave. For this reason we will call such functions characteristic variables.

Then the main remark of [CG94] (concerning KdV) is that, if one restricts the Hamiltonian to the submanifold $s = 0$, then the Hamiltonian turns out to coincide with the Hamiltonian of the KdV equation. The same is true for the Poisson tensor so that, in this submanifold, the equation of motion coincide with the KdV equation. However, with this procedure one does not see the appearance of the second KdV equation, and furthermore one has the problem that the manifold $s = 0$ is not invariant under the dynamics. Here normal form theory comes into play: indeed, using the characteristic variables, it is very easy to compute the first order normal form of the system and to get that it consists just of a couple of decoupled KdV equations. This method was already used in the context of the FPU problem in [BP06] and a similar point of view was also used in [BCP02] in order to deduce the NLS equation as a normal form for the Klein Gordon equation. Now, once one has computed the first term of the normal form, it is very natural to try to iterate the procedure. In this way we get our Hamiltonian result, and then, as anticipated above we perform Kodama's transformation in order to reduce our equations to a couple of decoupled integrable equations.

We now recall a few results on the deduction of modulation equations for the water wave problem. First, it is by now quite standard to obtain KdV as an equation describing unidirectional waves; KdV_5 has also been deduced as an higher order approximation for such unidirectional waves (see e.g. [DGH03]). For the case of more general initial data, giving rise to counterpropagating waves, we recall that the corrections of order μ^5 to the modulation equation were studied in [Wri05], where the author obtained that the first correction to the KdV equation contains terms which fulfill a linear time dependent equation plus terms in which an interaction of the counter-propagating waves is actually present. We emphasize that this description is compatible with the description that we get here. In particular the interaction between the counter-propagating waves is a product of the coordinate transformation that we use to put the system in normal form. A remarkable fact that our description yields concerning the interaction of counter-propagating waves is that the effects of interaction between the two waves disappear after the interaction, so that, if two spatially localized waves cross, then after the interaction they should return to the original shape, at least at the considered order of approximation.

We also recall that the modulation equation that describes the solutions can depend on the kind of initial data that one considers, in particular on the decay at infinity of the data. An interesting discussion of this phenomenon can be found in [BCL05] (see also [Lan19]).

A final consideration pertains the dynamics of the water waves in the complete model: after proving that a solution of the normal form equation fulfills the equations of the water wave problem up to an error of order μ^7, we apply Theorem 4.18 of [Lan13] to prove that the solution of the water wave problem remains $O(\mu^4)$ close to the solution of the normal form equation for a time of order μ^{-3}. We emphasize that the result applies to *all* initial data for the water wave problem which are of class $W^{\mathtt{s},1}$ with $\mathtt{s}$ large enough[1]. However such a time scale is not very satisfactory, since the time over which the dynamics of the fifth order normal form becomes visible is μ^{-5}. It would be interesting to try to apply the technique recently introduced in [BD18] (see also [BFP18]) in order to reach such a longer time scale.

From a technical point of view, the proof of our result requires some nontrivial steps. First one has to develop a normal form technique in the case where the unperturbed system is essentially a transport equation on $\mathbb{R}$. Actually some averaging techniques adapted to this situation were already developed in [BCP02]. Here, due to the particular structure of the water wave problem, we find that such techniques are particularly effective, and in particular we find a general algorithm to solve the so called homological equation.

The main difficulty is related to the fact that, in Hamiltonian perturbation theory, the transformation conjugating the system to its normal form is typically generated as the flow of some auxiliary Hamiltonian system. However, it turns out that the auxiliary Hamiltonian system one finds does not generate a flow (it is very similar an inverse heat equation). In Sect.4 we develop a technique allowing to put the system in normal form in the case of vector fields not generating a flow. The idea is to approximate the flow through its truncated expansion in the small parameter involved in the construction. The nontrivial point is that the so obtained transformation is not canonical, but only approximately canonical, thus one has to show that it can actually be used to normalize the system at the wanted order of approximation. We mention that an alternative technique that one could try to use in order to normalize the systems is that introduced in [Bam05] (also used in [BP06]), which is based on the use of Galerkin truncations. Maybe this would work also here, but this is not clear due to the difficulty of defining in a suitable way the operator ∂^{-1} in Fourier transform.

The paper is organized as follows In Sect. 2 we give our main result; in Sect. 3 we prepare the Hamiltonian of the water wave problem for the application of the normal form procedure. In particular this section reproduces the procedure by Craig and Groves in order to deduce KdV. In Sect. 4 we

[1]These are the L^1 based Sobolev spaces. See below for a precise definition.

develop an abstract framework for Hamiltonian normal form in the case of vector fields that do not generate a flow. In Sect. 5 we develop the tools needed to solve the so called homological equation in the case of the water wave problem and we prove our result on Hamiltonian normal form. We also show the obstruction that one finds when trying to put the system in normal form at order μ^7. Finally, in Sect. 6, we recall Kodama's transformation and conclude the proof of our main theorem.

This paper is dedicated to the memory of Walter Craig, he was a good friend and from a scientific point of view he had a great influence on my work. It was always a great pleasure to meet Walter and to spend time with him discussing about science or doing sport and tourism. I miss his great humanity and his enthusiasm.

Acknowledgments. Part of the material presented in this paper is the content of some lectures that I gave more than 10 years ago in order to prepare a visit by Walter Craig. I thank all the people who attended such lectures and contributed with their comments to improve the material, in particular Antonio Ponno with whom I had a lot of discussions on the subject. The discussions with Antonio Ponno were also the key to the understanding of the relevance of Kodama's work in the present context. I also would like to thank Doug Wright and David Lannes who gave me some relevant feedback on higher order corrections to KdV and on some technical issues. Finally I warmly thank the two referees of the paper whose comments allowed to greatly improve the paper.

2 Main result

Consider an ideal fluid occupying, at rest, the domain

$$\Omega_0 := \left\{(x, z) \in \Re^2 \ : \ -h < z < 0\right\} \ ,$$

we study the evolution of the free surface under the action of gravity, in the irrotational regime. Thus, given a function $\eta(x)$, we define the domain

$$\Omega_\eta := \left\{(x, z) \in \Re^2 \ : \ -h < z < \eta(x)\right\} \ . \tag{2.1}$$

and introduce the velocity potential ϕ, which is related to the velocity of the fluid by $u = \nabla\phi$. It is well known that the problem admits a Hamiltonian formulation [Zak68, CG94, CS93], the conjugated canonical variables being the wave profile η and the trace of the velocity potential at the free surface, namely

$$\psi(x) := \phi(x, \eta(x)) \ . \tag{2.2}$$

For the moment, just to fix ideas we work in the phase space of functions $z \equiv (\eta, \psi)$ of Schwartz class, later we will work in a more general setting. We endow the phase space by the L^2 scalar product, namely

$$\langle z; z' \rangle = \langle (\eta, \psi); (\eta', \psi') \rangle := \langle \eta; \eta' \rangle_{L^2} + \langle \psi; \psi' \rangle_{L^2} \ .$$

and by the Poisson tensor

$$J(\eta, \psi) := (-\psi, \eta) \ , \tag{2.3}$$

so that, given a Hamiltonian function $H = H(z)$, and defining its L^2 gradient (which is defined by $dH(z)h = \langle \nabla H(z); h \rangle$) the Hamilton equations are given by

$$\dot{z} = J \nabla H(z) \iff \begin{cases} \dot{\eta} = \nabla_\psi H(\eta, \psi) \\ \dot{\psi} = -\nabla_\eta H(\eta, \psi) \end{cases} . \tag{2.4}$$

The Hamiltonian of the water wave problem is given by

$$H(\eta, \psi) = \int \left(\frac{1}{2} g \eta^2 + \frac{1}{2} \psi G(\eta) \psi \right) dx \tag{2.5}$$

and G is the Dirichlet Neumann operator (see Definition 3.5).

We will look for solutions of the form

$$\eta(x) = \mu^2 h^3 \sqrt{2}\ \tilde{\eta}(\mu x) \ , \qquad \psi(x) = \mu \sqrt{2gh} h^2\ \tilde{\psi}(\mu x) \ , \quad \mu \ll 1 \ , \tag{2.6}$$

where the factors depending on g and h have been inserted for future convenience. In terms of the variables $\tilde{\eta}$ and $\tilde{\psi}$ the system is still Hamiltonian with a scaled Hamiltonian (see Subsect. 3.2), which takes the form

$$\mu \sqrt{gh} H_{WW}(\tilde{\eta}, \tilde{\psi}) \ , \tag{2.7}$$

with a suitable smooth H_{WW}. So, dividing by $\mu\sqrt{gh}$, which is equivalent to pass to the scaled time

$$\tilde{t} := \frac{t}{\mu \sqrt{gh}} \ , \tag{2.8}$$

one is reduced to the Hamiltonian system with Hamiltonian H_{WW}. Expanding in μ it takes the form (see Subsect. 3.3)

$$H_{WW} = H_0 + \epsilon H_1 + \epsilon^2 H_2 + ... \tag{2.9}$$

where $\epsilon := (h\mu)^2$

$$H_0 = \int \frac{\eta^2 + \psi_y^2}{2} dy \tag{2.10}$$

and the expressions of the higher order terms are not relevant for the moment. Here we also omitted the tildes.

Then, following [CG94], it is convenient to introduce the characteristic variables

$$r = \frac{\eta + \psi_y}{\sqrt{2}} \ , \quad s = \frac{\eta - \psi_y}{\sqrt{2}} \ , \tag{2.11}$$

which transform the Poisson tensor essentially in the Poisson tensor of the KdV equation (see Remark 3.3 and Subsect. 3.4). Precisely, the Hamilton equations of a Hamiltonian $H(r, s)$ turn out to be given by

$$\dot{r} = -\partial_y \nabla_r H \ , \quad \dot{s} = \partial_y \nabla_s H \ . \tag{2.12}$$

In particular one has that H_0 takes the form

$$H_0 = \int \frac{r^2 + s^2}{2} dy \ , \tag{2.13}$$

whose equations of motion are given by (1.1).

We remark that, (2.11) is just a change of variables, so that, if a solution is written in terms of the variables $r = r(y, t)$ and $s = s(y, t)$, then one can go back to the rescaled physical variables

$$\eta(y, t) := \frac{1}{\sqrt{2}} \left[r\Big(y, t\Big) + s\Big(y, t\Big) \right] \ , \tag{2.14}$$

$$\psi_y(y, t) := \frac{1}{\sqrt{2}} \left[r\Big(y, t\Big) - s\Big(y, t\Big) \right] \tag{2.15}$$

(originally denoted by $\tilde{\eta}$, $\tilde{\psi}$) in order to get the wave profile and the trace of the velocity potential. We remark that the integration constant allowing to pass from ψ_y to ψ is invariant with respect to the dynamics, so it is irrelevant in the following.

In particular it is possible to rewrite the Hamiltonian H_{WW} (c.f. (2.9)) in terms of the variables (r, s). *We will still say that this is the Hamiltonian of the water wave problem (in the variables* (r, s)*).*

Definition 2.1. *In the following, given a couple of function* $r(y, t)$, $s(y, t)$, *we say that*

$$z^p(y, t) := (\eta(y, t), \psi(y, t)) \ , \tag{2.16}$$

with η, ψ *given by* (2.14) *and* (2.15) *is called* the corresponding function in scaled physical variables.

As anticipated in the introduction, our goal is to put the system in normal form at second order.

Before stating the main result we still need a few preliminaries.

First we recall that the KdV hierarchy consists of a sequence of Hamiltonian systems with Hamiltonians $K_0, K_1, K_2...$, each of which is integrable and which pairwise commute, so that, in some sense, they form a complete set of integral of motion. Given an abstract function $u = u(y)$, the first three Hamiltonians are given explicitly (with a suitable choice of a normalization parameter) by

$$K_0(u) = \frac{1}{2}\int_{\mathbb{R}} u^2 dy \ , \tag{2.17}$$

$$K_1(u) = \int_{\mathbb{R}} \left(-\frac{1}{12}u_y^2 + \frac{1}{3}u^3 \right) dy \ , \tag{2.18}$$

$$K_2(u) = \int_{\mathbb{R}} \left(\frac{1}{2}u_{yy}^2 - \frac{5}{2}u_x^2 u + \frac{5}{8}u^4 \right) dy \ , \tag{2.19}$$

and the corresponding Hamilton vector fields are

$$\frac{du}{dt} = -\partial_y \nabla K_j(u) \ , \quad j = 0, 1, 2 \ .$$

Consider also the following Hamiltonian

$$H_{NF}(r, s) := K_0(r) + \epsilon K_1(r) + \epsilon^2 c_2 K_2(r) \tag{2.20}$$

$$+ K_0(s) + \epsilon K_1(s) + \epsilon^2 c_2 K_2(s) \ , \tag{2.21}$$

(with an arbitrary $c_2 \in \mathbb{R}$) and remark that its Hamilton equations are two decoupled equations, one for r and one for s. Each of these equations is integrable and one passes from one to the other just by inverting the space, namely by the transformation $y \to -y$.

Then, in order to precisely specify the properties of the transformation T_ϵ used to conjugate to the final normal form, we need to define the operator ∂^{-1} by

$$(\partial^{-1}u)(y) := \frac{1}{2}\left[\int_{-\infty}^{y} u(y_1)dy_1 - \int_{y}^{+\infty} u(y_1)dy_1 \right] \ . \tag{2.22}$$

Finally, we define the precise phase space we are going to use: we will work in the scale of Banach spaces $\mathcal{B}^{\mathtt{s}} := W^{\mathtt{s},1} \times W^{\mathtt{s},1} \ni (r, s) \equiv z$ where $W^{\mathtt{s},1}$ is the Sobolev space of the L^1 functions which have weak derivatives of order $\mathtt{s}$ of class L^1. We consider the case $\mathtt{s} \gg 1$. We will denote by $B_1^{\mathtt{s}} \subset \mathcal{B}^{\mathtt{s}}$ the ball of radius 1 centered at the origin. We will also consider the Sobolev spaces $W^{\mathtt{s},2}$ based on L^2.

Theorem 2.2. *There exists $c_2^* \in \mathbb{R}$ s.t. the following holds true. For any $\mathtt{s}'$ there exists $\epsilon_* > 0$ and $\mathtt{s}, \mathtt{s}''$, s.t., if $0 < \epsilon < \epsilon_*$, then there exists a map $T_\epsilon : B_1^{\mathtt{s}} \to W^{\mathtt{s}'',1} \times W^{\mathtt{s}'',1}$, with the following properties*

(i) $T_\epsilon(r,s) - (r,s)$ *is a polynomial in* $\partial^k r, \partial^k s$, $k = -1, ..., 5$,

(ii) $\sup_{(r,s)\in B_1^{\mathtt{s}}} \|T_\epsilon(r,s) - (r,s)\|_{W^{\mathtt{s}'',1}\times W^{\mathtt{s}'',1}} \leq C\epsilon$,

(iii) *Let I_ϵ be an interval containing the origin and $z(.) = (r(.), s(.)) \in C^1(I_\epsilon; B_1^{\mathtt{s}})$ be a solution of the Hamiltonian system* (2.20), (2.21), *with $c_2 = c_2^*$, define*
$$z_a \equiv (r_a, s_a) := T_\epsilon(r,s) \ . \tag{2.23}$$
Then there exists $R \in C^1(I_\epsilon, W^{\mathtt{s}',2} \times W^{\mathtt{s}',2})$ s.t. one has
$$\dot{z}_a(t) = J\nabla H_{WW}(z_a(t)) + \epsilon^3 R(t) \ , \quad \forall t \in I_\epsilon \ , \tag{2.24}$$
where H_{WW} is the Hamiltonian (2.9) *of the water wave problem rewritten in the variables (r,s).*

Remark 2.3. *The form of the transformation T_ϵ is not computed explicitly in the paper, but it can be extracted by slghtly developing the computations of Sect. 5. It is given by $T_1 \circ T_2 \circ T_K$, with T_1, T_2 and T_K given by* (5.38), (5.39) *and* (6.3) *respectively.*

Remark 2.4. *The value of c_2^* can be computed in a straightforward way, however, its computation is quite long and I think that if one wants to use the theory of this paper to actually compute in a quantitative way the wave profile it is better to check all the computation, for example through a symbolic manipulator.*

By applying Theorem 4.18 of [Lan13] one immediately gets the following result (here the smoothness indexes have a value different from those of Theorem 2.2) which gives some dynamical information on all solutions with smooth enough integrable initial data[2].

Corollary 2.5. *For any $\mathtt{s}'$ there exist $\mathtt{s}$, $\epsilon_* > 0$ and $\mathtt{T} > 0$ s.t., if $\epsilon < \epsilon_*$ then the following holds true. Consider the Cauchy problem for the water wave problem H_{WW} with initial datum (η_0, ψ_0) fulfilling*
$$\|(\eta_0, \psi_0)\|_{W^{\mathtt{s},1}\times W^{\mathtt{s}+1,1}} \leq 1 \tag{2.25}$$

[2] Actually one could get a slightly more precise statement using the Beppo Levi spaces, here, for the sake of simplicity, I decided not to use them.

and denote by $z^p(t)$ the corresponding solution (of course still in the scaled physical variables (η, ψ)). Then there exists a solution $(r(t), s(t))$ of the Hamiltonian system (2.20), (2.21) *with the following property: denote by z_a the function defined by* (2.23)*, and by z_a^p the corresponding solution in scaled physical variables, then, for all times such that*

$$|t| \leq \frac{\mathtt{T}}{\epsilon} \ ,$$

one has

$$\|z_a^p(t) - z^p(t)\|_{W^{\mathtt{s}',2}\times W^{\mathtt{s}',2}} \leq C\epsilon^3 t \ . \tag{2.26}$$

From this Corollary, it is possible to go back to the *original non scaled* variables. For example, exploiting the embedding of $W^{\mathtt{s}',2} \subset L^\infty$ (provided $\mathtt{s}' > 1/2$), one can get for

$$|t| \leq \mathtt{T}/\mu^3\sqrt{gh} \ ,$$

the following estimate on the profile of the wave:

$$\|\eta(t) - \eta_a(t)\|_{L^\infty} \leq C\mu^9 t \ . \tag{2.27}$$

We remark that in these results we never tried to get optimality concerning the regularity loss related to our procedure.

A more serious problem with our deduction is that we deal with solutions of the normal form equations which belong to the Sobolev spaces based on L^1. It is not clear if this can be avoided. This plays a role in the definition of ∂^{-1}. One could try to work in spaces of square integrable functions in which there is a better theory of existence and uniqueness for the equations of the KdV hierarchy, but this requires some nontrivial work. Probably one could work with some regularized version of the operator ∂^{-1} defined in (2.22).

Finally we remark that, as anticipated in the introduction, one would like to get results of correspondence between the approximate solution and the true solutions over longer time scales, but this is beyond the scope of this paper and is left for future work.

3 Preliminaries: scaling, expansions and characteristic variables

3.1 Canonical transformations

In this subsection we recall a few basic facts of Hamiltonian mechanics

Consider a change of variables $z := T(\zeta)$. Exploiting the formula

$$\nabla(H \circ T)(\zeta) = [dT(\zeta)]^*(\nabla H)(T(\zeta)) \ , \tag{3.1}$$

which is true for any smooth function H, one immediately sees that if $\zeta(t)$ fulfills the Hamiltonian equations

$$\dot{\zeta} = J\nabla(H \circ T)(\zeta) \ , \tag{3.2}$$

then, $z(t) = T(\zeta(t))$ fulfills

$$\dot{z} = dT(\zeta)\dot{\zeta} = dT(\zeta)J[dT(\zeta)]^*(\nabla H)(T(\zeta)) \ . \tag{3.3}$$

Vice-versa, if z fulfills

$$\dot{z} = J\nabla H(z) \ ,$$

then one has

$$\dot{\zeta} = [dT^{-1}(T(\zeta))]J[dT^{-1}(T(\zeta))]^*\nabla(H \circ T)(\zeta) \ ,$$

where the star denotes the adjoint with respect to the L^2 scalar product.

With this remark one can characterize the coordinate transformations leaving invariant the Hamiltonian formalism.

Definition 3.1. *A coordinate transformation $z = T(\zeta)$ is said to be canonical if it transforms the Hamilton equation of any Hamiltonian H into the Hamilton equations of $H \circ T$.*

As a consequence of eq. (3.3) we have the following proposition.

Proposition 3.2. *A coordinate transformation is canonical if and only if it fulfills*

$$dT(\zeta)J[dT(\zeta)]^* = J \ . \tag{3.4}$$

Then it is immediate to see that (3.4) is equivalent to

$$[dT^{-1}(z)]J[dT^{-1}(z)]^* = J \ .$$

3.2 Hamiltonian scalings

The main tool in order to perform the scaling at a Hamiltonian level is the following remark which is also needed in order to compute how the Poisson tensor changes when introducing the characteristic variables.

Remark 3.3. *A linear change of variables $z = B\zeta$, transforms the Hamilton equations of H into the equations $\dot{\zeta} = \tilde{J}\nabla\widehat{H}(\zeta)$, where $\widehat{H}(\zeta) := H(B\zeta)$ and $\tilde{J} := B^{-1}JB^{-*}$, and B^{-*} is the adjoint (with respect to the L^2 metric) of the inverse of B.*

In the particular case of the linear change of coordinates given by

$$z = B\zeta \qquad \Longleftrightarrow \qquad \begin{cases} \eta(x) = \epsilon_1 \tilde{\eta}(\mu x) \\ \psi(x) = \epsilon_2 \tilde{\psi}(\mu x) \end{cases} \tag{3.5}$$

One has the following Lemma

Lemma 3.4. *The transformation* (3.5) *transforms the Hamilton equations of H into the Hamilton equations*

$$\tilde{H}(\zeta) := \frac{\mu}{\epsilon_1\epsilon_2} H(B\zeta) \ . \tag{3.6}$$

Proof. We just compute B^{-1}, B^{-*} and $B^{-1}JB^{-*}$. First, one has that B^{-1} is given by

$$[B^{-1}(\eta,\psi)](y) = \left(\frac{1}{\epsilon_1}\eta\left(\frac{y}{\mu}\right), \frac{1}{\epsilon_2}\psi\left(\frac{y}{\mu}\right) \right) \ ,$$

from which one can compute its adjoint. Of course it enough to consider one of the components of the vector z. We have

$$\begin{aligned} \langle \eta'; B^{-1}\eta \rangle &= \int \eta'(y)\frac{1}{\epsilon_1}\eta\left(\frac{y}{\mu}\right) dy = \int \mu\eta'(y)\frac{1}{\epsilon_1}\eta\left(\frac{y}{\mu}\right) d\frac{y}{\mu} \\ &= \mu \int \eta(\mu x)\frac{1}{\epsilon_1}\eta\left(x\right) dx = \langle B^{-*}\eta'; \eta \rangle \end{aligned}$$

so that we have

$$[B^{-*}(\eta,\psi)](x) = \left(\frac{\mu}{\epsilon_1}\eta\left(\mu x\right), \frac{\mu}{\epsilon_2}\psi\left(\mu x\right) \right) \ .$$

It follows that

$$[B^{-1}JB^{-*}(\eta,\psi)](y) = B^{-1}J\left(\frac{\mu}{\epsilon_1}\eta\left(\mu x\right), \frac{\mu}{\epsilon_2}\psi\left(\mu x\right) \right) \tag{3.7}$$

$$= B\left(\frac{\mu}{\epsilon_2}\psi\left(\mu x\right), \frac{\mu}{\epsilon_1}\eta\left(\mu x\right) \right) = \left(\frac{\mu}{\epsilon_1\epsilon_2}\psi(y), -\frac{\mu}{\epsilon_1\epsilon_2}\eta(y) \right) = \frac{\mu}{\epsilon_1\epsilon_2}J(\eta,\psi). \tag{3.8}$$

Thus the Hamilton equations of H are transformed into $\dot{\zeta} = \frac{\mu}{\epsilon_1\epsilon_2}J\nabla\widehat{H}(\zeta) = J\nabla\frac{\mu}{\epsilon_1\epsilon_2}\widehat{H} = J\nabla\tilde{H}$ □

3.3 Expansion of the Hamiltonian

In this subsection and in Subsect. 3.4, essentially we repeat with minor changes the procedure developed in [CG94] in order to show the appearance of KdV in the water wave problem.

For the sake of completeness, we start by recalling the definition of the Dirichlet-Neumann operator, then we will recall its expansion, which was computed in [CS93].

Definition 3.5. *Given a function $\psi(x)$, consider the boundary value problem*

$$\Delta\phi = 0 \qquad (x,z) \in \Omega_\eta \tag{3.9}$$

$$\phi_z\Big|_{z=-h} = 0 \tag{3.10}$$

$$\lim_{x\to\infty} \phi = 0 \tag{3.11}$$

$$\phi\Big|_{z=\eta(x)} = \psi \qquad , \tag{3.12}$$

and let ϕ be its solution. Then the linear operator $G(\eta)$ defined by

$$G(\eta)\psi = \sqrt{1+\eta_x^2}\partial_n\phi\big|_{z=\eta(x)} \equiv (\phi_z - \eta_x\phi_x)\big|_{z=\eta(x,y)} \tag{3.13}$$

is called the Dirichlet Neumann operator*, where ∂_n is the derivative in the direction normal to $z = \eta(x)$.*

Formally, it is well known [CS93] that the Dirichlet Neumann operator has a Taylor expansion of the form $G(\eta) \simeq \sum_{j\geq 0} G^{(j)}(\eta)$ with $G^{(j)}(\eta)$ homogeneous of degree j in η. One has

$$G^{(0)} = D\tanh(hD) \ , \tag{3.14}$$

$$G^{(1)} = D\eta D - G^{(0)}\eta G^{(0)} \tag{3.15}$$

$$G^{(2)} = -\frac{1}{2}\left(D^2\eta^2G^{(0)} + G^{(0)}\eta^2D^2 - 2G^{(0)}\eta G^{(0)}\eta G^{(0)}\right) \tag{3.16}$$

where we used the standard notation $D := -i\partial_x$.

Substituting (3.5) in (3.14), denoting as above, by $y := \mu x$, and ∂_y the corresponding partial derivative and $D_y := -i\partial_y$, one gets

$$\begin{aligned} G^{(0)} &= \mu^2hD_y^2 - \frac{1}{3}\mu^4h^3D_y^4 + \frac{2}{15}\mu^6h^5D^6 + O(\mu^8) \\ &= -\mu^2h\partial_y^2 - \frac{1}{3}\mu^4h^3\partial_y^4 - \frac{2}{15}\mu^6h^5\partial_y^6 + O(\mu^8) \end{aligned} \tag{3.17}$$

$$\begin{aligned} G^{(1)} &= \mu^2\epsilon_1D_y\eta D_y - \mu^4\epsilon_1h^2D_y^2\eta D_y^2 + O(\epsilon_1\mu^6) \\ &= -\mu^2\epsilon_1\partial_y\eta\partial_y - \mu^4\epsilon_1h^2\partial_y^2\eta\partial_y^2 + O(\epsilon_1\mu^6) \end{aligned} \tag{3.18}$$

$$G^{(2)} = O(\epsilon_1^2\mu^4) \ .$$

Inserting in the Hamiltonian the scaling (3.5), and the expansions (3.17) and (3.18) and taking advantage of Lemma 3.4, one gets that the Hamiltonian for the scaled variables becomes $H = H_0 + H_1 + H_2 + h.o.t.$ with

$$H_0 := \frac{1}{2}\int \left(\frac{\epsilon_1}{\epsilon_2} g\eta^2 + \frac{\epsilon_2}{\epsilon_1}\mu^2 h\psi_y^2\right) dy \ , \tag{3.19}$$

$$H_1 := \frac{1}{2}\int \left(-\frac{\epsilon_2}{\epsilon_1}\frac{1}{3}\mu^4 h^3\psi_{yy}^2 + \epsilon_2\mu^2\eta\psi_y^2\right) dy \ , \tag{3.20}$$

$$H_2 := \frac{1}{2}\frac{\epsilon_2}{\epsilon_1}\int \left(\frac{2}{15}\mu^6 h^5\psi_{yyy}^2 - \mu^4\epsilon_1 h^2\eta\psi_{yy}^2\right) dy \ , \tag{3.21}$$

where we omitted the tildes (remark that, as a difference with the notation of Sect. 2, the small parameters are here included in H_j. We will come back to the original notation at the end of this subsection). The choice

$$\frac{\epsilon_1}{\epsilon_2} g = \frac{\epsilon_2}{\epsilon_1}\mu^2 h$$

makes the two terms of H_0 of equal order of magnitude, and gives it the form

$$H_0 := \mu\sqrt{gh}\frac{1}{2}\int \left(\eta^2 + \psi_y^2\right) dy \ ; \tag{3.22}$$

The choice $\epsilon_2 = \mu h^2\sqrt{2gh}$, which implies $\epsilon_1 = \sqrt{2}\mu^2 h^3$, also implies that the two terms of H_1 have the same order of magnitude ($\sqrt{2}$ has been inserted for future convenience). Remark in particular that the relationship (3.5) turns out to take the form (2.6).

Inserting in the Hamiltonian one gets

$$H_1 := \mu^3\sqrt{gh}h^2\frac{1}{2}\int \left(-\frac{1}{3}\psi_{yy}^2 + \sqrt{2}\eta\psi_y^2\right) dy \tag{3.23}$$

$$H_2 := \mu^5\sqrt{gh}h^4\frac{1}{2}\int \left(\frac{2}{15}\psi_{yyy}^2 - \sqrt{2}\eta\psi_{yy}^2\right) dy \ . \tag{3.24}$$

Finally passing to the scaled time $\tilde{t}$ (cf (2.8)) and separating the small parameter from H_j, the Hamiltonian takes the form

$$H_{WW} = H_0 + \epsilon H_1 + \epsilon^2 H_2 + O(\epsilon^3) \ , \tag{3.25}$$

with $\epsilon := (h\mu)^2$ and

$$H_0 = \int \frac{\eta^2 + \psi_y^2}{2} dy \tag{3.26}$$

$$H_1 = \frac{1}{2}\int \left(-\frac{1}{3}\psi_{yy}^2 + \sqrt{2}\eta\psi_y^2\right) dy \tag{3.27}$$

$$H_2 = \frac{1}{2}\int \left(\frac{2}{15}\psi_{yyy}^2 - \sqrt{2}\eta\psi_{yy}^2\right) dy \tag{3.28}$$

More precisely, we have the following result

Proposition 3.6. *Consider the Hamiltonian* (2.5) *and introduce the scaled variables* (3.5)*. Let H_{WW} be the scaled Hamiltonian of the water wave problem in scaled variables, then, for any* $\mathtt{s}'$ *there exists* $\mathtt{s}$*, s.t., for any ball* $\mathcal{U} \subset W^{\mathtt{s},2} \times W^{\mathtt{s},2}$ *centered at the origin, there exists* $\epsilon_* > 0$ *s.t.*

$$\sup_{0<\epsilon<\epsilon_*} \frac{\sup_{(\eta,\psi)\in\mathcal{U}} \|J\nabla H_{WW} - J\nabla(H_0 + \epsilon H_1 + \epsilon^2 H_2)\|_{W^{\mathtt{s}',2}\times W^{\mathtt{s}',2}}}{\epsilon^3} < \infty \ . \tag{3.29}$$

The proof is postponed to Sect. (4.2).

3.4 Characteristic variables

We introduce the characteristic variables (2.11). Applying Remark 3.3 it is easy to see that the Hamilton equations take the form (2.12). Inserting in the various part of the Hamiltonian, one gets

$$H_0 = \int \frac{r^2 + s^2}{2} dy \ , \tag{3.30}$$

$$H_1 = \int_{\mathbb{R}} \left(-\frac{1}{12}(r_y^2 + s_y^2) + \frac{r^3 + s^3}{4} \right. \tag{3.31}$$

$$\left. + \frac{r_y s_y}{6} - \frac{r^2 s + r s^2}{4} \right) dy \tag{3.32}$$

$$H_2 = \int \left(\frac{1}{2}\frac{r_{yy}^2 + s_{yy}^2}{15} - \frac{1}{4}(r r_y^2 + s s_y^2) \right. \tag{3.33}$$

$$\left. - \frac{1}{15} r_{yy} s_{yy} - \frac{1}{4}(r s_y^2 - 2 r r_y s_y + s r_y^2 - 2 s r_y s_y) \right) dy \tag{3.34}$$

Remark 3.7. *The Hamiltonian is the sum of terms, each of which is the integral over $\mathbb{R}$ of a polynomial in r, s and their derivatives. If a term is a function of r (and its derivatives) only, then it is invariant under the flow of H_0 and thus it Poisson commutes with it, which means that it is in normal form. The same is true if a term depends on s and its derivatives only.*

Remark 3.8. *If one restricts $H_0 + \epsilon H_1$ to the manifold $s = 0$ then one gets*

$$H_{res} = \int \left(\frac{r^2}{2} - \frac{1}{12} r_y^2 + \frac{r^3}{4} \right) dy \ , \tag{3.35}$$

namely the Hamiltonian of a KdV equation in a reference frame translating with velocity 1.

This is the procedure used by Craig and Groves in [CG94] in order to deduce KdV as an equation describing the dynamics of water waves in this approximation.

4 Abstract Birkhoff normal form with no flow

4.1 Birkhoff normal form in the finite dimensional case

In this subsection we recall the algorithm of Birkhoff normal form in the finite dimensional case. We will also present some explicit formulae that will play a role in the water wave problem.

We first introduce some notations. Let $\mathcal{P}$ be a $2n$-dimensional linear phase space endowed by a scalar product $\langle .; . \rangle$; we denote by J the Poisson tensor (namely a skew-symmetric invertible linear operator) and define the Hamiltonian vector field of a Hamiltonian G by $J\nabla G$. Furthermore, given a function F, we denote by

$$\mathcal{L}_G F := dF J\nabla G \equiv \langle \nabla F; J\nabla G \rangle$$

its Lie derivative with respect to the vector field of G. In a Hamiltonian framework this quantity is also called Poisson Brackets of F and G, and denoted by

$$\{F; G\} := \mathcal{L}_G H \ . \tag{4.1}$$

Consider a family of Hamiltonian systems

$$H(z, \epsilon) = \sum_{k \geq 0} \epsilon^k H_k(z) \ , \tag{4.2}$$

smooth in a neighborhood of the origin. In the following we will not be interested in the size of the neighborhood, so we will not specify the domain of functions, giving for understood that they are smooth in a suitable neighborhood of the origin.

We are interested in the situation in which H_0 is a quadratic form in z, whose Hamiltonian vector field generates a periodic flow. Then it is well known that one can put the system in normal form at any order. In particular the following version of Birkhoff normal form theorem holds.

Theorem 4.1. *Fix an arbitrary positive integer $r \geq 1$, then there exists a canonical transformation T (defined in a neighborhood of the origin) which puts the system (4.2) in Normal Form at order r, namely such that*

$$H \circ T = H_0 + \sum_{k=1}^{r} \epsilon^k Z_k + O(\epsilon^{r+1}) \tag{4.3}$$

where Z_k Poisson commutes with H_0, namely $\{H_0; Z_k\} \equiv 0$.

The idea of the proof is to construct iteratively a canonical transformation putting the system in normal form. This means to first construct a canonical transformation pushing the non normalized part of the Hamiltonian to order ϵ^2, then a transformation pushing it to order ϵ^3 and so on. Each of the transformations is constructed as the flow of a suitable auxiliary Hamiltonian system (Lie transform method).

We now perform explicitly the construction at order three which is the one relevant for the water wave problem.

Let G be a smooth function, and consider the corresponding Hamilton equations, namely

$$\dot{z} = J\nabla G(z) \ ,$$

denote by Φ_G^t the corresponding flow.

Definition 4.2. *The map Φ_G^ϵ will be called Lie transform generated by G.*

It is well known that Φ_G^ϵ is a canonical transformation.

We are now going to study the way a Hamiltonian changes when the coordinate are subjected to a Lie transformation. Thus, let F be a smooth function and let Φ_G^ϵ be the Lie transform generated by a function G. To compute the expansion of $F \circ \Phi_G^\epsilon$, first remark that

$$\frac{d}{dt} F \circ \Phi_G^t = \{F, G\} \circ \Phi_G^t \tag{4.4}$$

so that, defining the sequence

$$F^{(0)} := F \ , \quad F^{(l)} = \left\{F^{(l-1)}; G\right\} \ , \quad l \geq 1 \ , \tag{4.5}$$

one has $\forall r \geq 0$

$$F \circ \Phi_G^\epsilon = \sum_{l=0}^{r} \frac{\epsilon^l}{l!} F^{(l)} + O(\epsilon^{r+1}) \ . \tag{4.6}$$

We come to the normalization procedure. We look for an auxiliary Hamiltonian G_1 whose flow normalizes the Hamiltonian (4.2) at first order. For a generic G_1, one has

$$H \circ \Phi_{G_1}^\epsilon = (H_0 + \epsilon H_1 + \epsilon^2 H_2 + \epsilon^3 H_3) \circ \Phi_{G_1}^\epsilon + O(\epsilon^4)$$

$$= H_0 + \epsilon \{H_0; G_1\} + \frac{\epsilon^2}{2} \{\{H_0; G_1\}; G_1\} + \frac{\epsilon^3}{6} \{\{\{H_0; G_1\}; G_1\}; G_1\} \tag{4.7}$$

$$+\epsilon H_1 + \epsilon^2 \{H_1; G_1\} + \frac{\epsilon^3}{2} \{\{H_1; G_1\}; G_1\} \tag{4.8}$$

$$+\epsilon^2 H_2 + \epsilon^3 \{H_2; G_1\} + \epsilon^3 H_3 + O(\epsilon^4) \tag{4.9}$$

In order to determine G_1 in such a way that the terms of order ϵ are in normal form, we recall the following well known Lemma [BG93].

Lemma 4.3. *Assume that the flow $\Phi^t_{H_0}$ is periodic of period $\mathtt{T}$. Define*

$$Z_1(z) := \frac{1}{\mathtt{T}} \int_0^{\mathtt{T}} H_1(\Phi^\tau(z)) d\tau \ , \tag{4.10}$$

and $W_1 := H_1 - Z_1$, then Z_1 is in normal form and

$$G_1(z) := \frac{1}{\mathtt{T}} \int_0^{\mathtt{T}} \tau W_1(\Phi^\tau(z)) d\tau \tag{4.11}$$

solves the homological equation

$$\{H_0; G_1\} + W_1 = 0 \ . \tag{4.12}$$

Proof. Just compute

$$\begin{aligned} \{H_0; G_1\}(z) &= -\frac{d}{dt}\Big|_{t=0} G_1(\Phi^t_{H_0}(z)) = -\frac{1}{\mathtt{T}} \int_0^{\mathtt{T}} \tau \frac{d}{dt}\Big|_{t=0} W_1(\Phi^{t+\tau}_{H_0}(z)) d\tau \\ &= -\frac{1}{\mathtt{T}} \int_0^{\mathtt{T}} \tau \frac{d}{d\tau} W_1(\Phi^\tau_{H_0}(z)) d\tau = -\frac{\tau W_1(\Phi^\tau_{H_0}(z))}{\mathtt{T}}\Big|_0^{\mathtt{T}} \\ &\quad + \frac{1}{\mathtt{T}} \int_0^{\mathtt{T}} W_1(\Phi^\tau_{H_0}(z)) d\tau = -W_1(z) \ . \end{aligned}$$

□

Using such a G_1, exploiting also (4.12) in order to compute $\{H_0; G_1\}$, one gets

$$\begin{aligned} H \circ \Phi^\epsilon_{G_1} &= H_0 + \epsilon Z_1 \\ &+ \epsilon^2 \left(\{Z_1; G_1\} + H_2 + \frac{1}{2} \{W_1; G_1\} \right) && (4.13) \\ &+ \epsilon^3 \left(H_3 + \{H_2; G_1\} + \frac{1}{2} \{\{Z_1; G_1\}; G_1\} + \frac{1}{3} \{\{W_1; G_1\}; G_1\} \right) && (4.14) \\ &+ O(\epsilon^4) \\ &= H_0 + \epsilon Z_1 + \epsilon^2 H_{2,1} + \epsilon^3 H_{3,1} + O(\epsilon^4) \ , \end{aligned}$$

where we denoted by $H_{2,1}$, resp. $H_{3,1}$ the brackets in (4.13) resp. (4.14).

Let G_2 be a further auxiliary Hamiltonian. One has

$$\begin{aligned} H \circ \Phi^\epsilon_{G_1} \circ \Phi^{\epsilon^2}_{G_2} = H_0 + \epsilon^2 \{H_0; G_2\} + \epsilon Z_1 + \epsilon^3 \{Z_1; G_2\} \\ + \epsilon^2 H_{2,1} + \epsilon^3 H_{3,1} + O(\epsilon^4) \ . \end{aligned}$$

Decomposing $H_{2,1}$ as in Lemma 4.3, namely

$$H_{2,1} = Z_2 + W_2 \tag{4.15}$$

and determining G_2 as the solution of

$$\{H_0; G_2\} + W_2 = 0 \ , \tag{4.16}$$

one gets

$$H \circ \Phi^{\epsilon}_{G_1} \circ \Phi^{\epsilon^2}_{G_2} = H_0 + \epsilon Z_1 + \epsilon^2 Z_2 + \epsilon^3 H_{3,2} + O(\epsilon^4) \ ,$$

where, explicitly

$$H_{3,2} = H_3 + \{H_2; G_1\} + \frac{1}{2}\{\{Z_1; G_1\}; G_1\} + \frac{1}{3}\{\{W_1; G_1\}; G_1\} + \{Z_1; G_2\} \ . \tag{4.17}$$

To iterate a third time one has to decompose $H_{3,2} = Z_3 + W_3$, to solve the homological equation

$$\{H_0; G_3\} + W_3 = 0 \ , \tag{4.18}$$

and to transform using $\Phi^{\epsilon^3}_{G_3}$.

Of course one can iterate as many times as one wants. Here we described the procedure at order 3, since in the case of the water wave problem we do not have an abstract argument ensuring that G_l belongs to a good class of objects and we need to compute it explicitly. In particular, as we anticipated, at order 3 we find the first obstruction (see sect. 5).

4.2 Almost smooth maps

We are now going to generalize the above construction to the case where the vector field of the function G to be used to put the system in normal form does not generate a flow. The idea is to approximate all the objects we meet by their truncated expansion in ϵ.

We will work in a scale of Banach spaces $\mathcal{B} \equiv \{\mathcal{B}^{\mathtt{s}}\}$. In the case of the water wave problem we will use the space $\mathcal{B}^{\mathtt{s}} := W^{\mathtt{s},1} \times W^{\mathtt{s},1}$ (since we will work with the variables (r, s)). However it will be clear that everything works in an abstract context. We will assume that for $\mathtt{s}$ large enough the space $\mathcal{B}^{\mathtt{s}}$ is embedded in a Hilbert space, whose scalar product $\langle .; . \rangle$ will be used to define the gradient of functions.

In the case of the water wave problem the Hilbert space is $L^2 \times L^2$, so that the gradient will be with respect to the standard $L^2 \times L^2$ metric.

Furthermore we denote by J a skew-symmetric operator that we will use as the Poisson tensor. We assume that $\forall\mathtt{s}$ there exists $\mathtt{s}'$ such that $J:\mathtt{s}'\to\mathtt{s}$ is bounded.

In order to perform the proofs we will approximate the vector fields by smooth objects. To this end we assume that there exists a sequence of linear truncation operators $\{\Pi_N\}_{N\geq 0}$ which, for any $\mathtt{s},\mathtt{s}'$ are bounded as operators from $\mathcal{B}^{\mathtt{s}}$ to $\mathcal{B}^{\mathtt{s}'}$ and which converge to the identity as $N\to\infty$. Furthermore we assume that Π_N is self adjoint and commutes with J.

In the case of the water wave problem they are the standard truncations in Fourier space.

Following [Bam13], we will consider functions which have a weak smoothness property.

Let $\mathcal{B}\equiv\{\mathcal{B}^{\mathtt{s}}\}$ and $\tilde{\mathcal{B}}\equiv\left\{\tilde{\mathcal{B}}^{\mathtt{s}'}\right\}$ be two scales of Banach spaces, then we give the following definition.

Definition 4.4. *A map F will be said to be almost smooth if, $\forall\mathtt{r},\mathtt{s}'\geq 0$ there exist $\mathtt{s}$ and an open neighborhood of the origin $\mathcal{U}_{\mathtt{rss}'}\subset\mathcal{B}^{\mathtt{s}}$ such that*

$$F\in C^{\mathtt{r}}(\mathcal{U}_{\mathtt{rss}'};\tilde{\mathcal{B}}^{\mathtt{s}'})\ . \tag{4.19}$$

We will use the same notation also when one of the two scales, or both, is composed by a single space.

Furthermore, we will also deal with maps which depend on a small parameter ϵ. We will say that they are almost smooth if they fulfill the above definition with the scale $\mathcal{B}$ replaced by the scale $\{\mathcal{B}^{\mathtt{s}}\times\mathbb{R}\}$, where $\mathbb{R}$ has been added as the domain of ϵ. *In this case we will assume that the domain $\mathcal{U}_{\mathtt{rss}'}$ of* (4.19) *has the form $\mathcal{U}_{\mathtt{rss}'}=\mathcal{V}_{\mathtt{rss}'}\times I_{\mathtt{rss}'}$ with $\mathcal{V}_{\mathtt{rss}'}\subset\mathcal{B}^{\mathtt{s}}$ and $I_{\mathtt{rss}'}$ an interval.* The important point is that the size of the open set $\mathcal{V}_{\mathtt{rss}'}$ does not depend on ϵ.

In the following the width of open sets does not play any role so we will avoid to specify it. In particular we will often consider maps from a Banach space to some other space, *by this we* **always** *mean a map defined in an open neighborhood of the origin.*

We remark that, according to the above definition, if F is an almost smooth map, then its differential has the property that

$$\forall l,r,\ \exists k_1,k_2\ ,\quad s.t.\quad dF(.)\in C^{\mathtt{r}}(\mathcal{B}^{k_1};B(\mathcal{B}^{k_2},\mathcal{B}^{l}))\ . \tag{4.20}$$

In the following we will have to consider also the adjoint $dF(z)^*$ of $dF(z)$ with respect to the scalar product of the Hilbert space we use for the computation of gradients. With a small abuse of notation we will say that dF^* is almost smooth if it has the property (4.20).

Definition 4.5. *In the rest of the paper we will write*

$$A = B + O(\epsilon^{r+1})$$

if

$$\frac{A-B}{\epsilon^{r+1}}$$

is an almost smooth map.

As a first application of this notion we give the proof of Proposition 3.6. *Proof of Proposition 3.6.* First we recall that, in the original non scaled physical variables, the Hamilton equations of (2.5) are given by

$$\partial_t \eta = G(\eta)\psi \ , \tag{4.21}$$

$$\partial_t \psi = -g\eta - \frac{1}{2}\psi_x^2 + \frac{1}{2}\frac{\left(G(\eta)\psi + \eta_x\psi_x\right)^2}{1+\eta_x^2} \ . \tag{4.22}$$

After the scaling (3.5), in particular the operator G is substituted by the scaled Dirichlet Neumann operator, that we now denote by G_{scal}, studied in [Lan13], whose properties are summarized in Proposition 3.44 of that book. In particular, by such a proposition (and by Theorem 3.21 of [Lan13]) one has that the map $(\psi, \eta, \mu^2) \mapsto G_{scal}(\eta)\psi$ is almost smooth in the scale $W^{\mathtt{s},2} \times W^{\mathtt{s},2}$. It follows that the vector field $J\nabla H_{WW}$ is almost smooth. Thus, from the formal computation of Sect. 3.2, its truncated Taylor expansion in ϵ has the structure

$$J\nabla H_{WW} = X_0 + \epsilon X_1 + \epsilon^2 X_2 + O(\epsilon^3) \ ,$$

with $X_j = J\nabla H_j$, $j = 0, 1, 2$, and this is the thesis. □

4.3 Lie transform with no flow

Consider now an almost smooth vector field X and define the sequence of almost smooth vector fields

$$X^{(0)} := X \ , \quad X^{(k)} := dX^{(k-1)}X \ , \quad k \geq 1, \tag{4.23}$$

Remark 4.6. *If X is smooth as a map from $\mathcal{B}^{\mathtt{s}}$ to itself, for some $\mathtt{s}$, then denoting by Φ^ϵ the flow it generates, for any r one has*

$$\Phi^\epsilon(z) = z + \sum_{k\geq 0}^{r} \frac{\epsilon^{k+1}}{(k+1)!} X^{(k)}(z) + O(\epsilon^{r+1}) \ , \tag{4.24}$$

This follows from the formula

$$\frac{d^k}{dt^k}\left(X \circ \Phi^t\right) = dX^{(k-1)} \circ \Phi^t \ ,$$

which is easily proven by induction.

Having fixed X and $r \geq 1$, we define

$$T_{X,r,\epsilon}(z) := z + \sum_{k=0}^{r-1} \frac{\epsilon^{k+1}}{(k+1)!} X^{(k)}(z) \ , \tag{4.25}$$

$$\mathcal{T}_{X,r,\epsilon}(z) := z + \sum_{k=0}^{r-1} \frac{(-\epsilon)^{k+1}}{(k+1)!} X^{(k)}(z) \ . \tag{4.26}$$

In the following we will systematically omit the indexes X, r, ϵ from T and $\mathcal{T}$.

We remark that both T and $\mathcal{T}$ are almost smooth maps Therefore also $T \circ \mathcal{T}$ and $\mathcal{T} \circ T$ are almost smooth.

Lemma 4.7. *One has*

$$T \circ \mathcal{T} = 1 + O(\epsilon^{r+1}) \ , \quad \mathcal{T} \circ T = 1 + O(\epsilon^{r+1}) \ . \tag{4.27}$$

Proof. The proof is based on a regularization procedure. Using the truncation operator Π_N we define the truncated vector field by

$$X_N(z) := \Pi_N X(\Pi_N z) \ . \tag{4.28}$$

The flow it generates will be denoted by Φ_N^t.

We consider $T \circ \mathcal{T}$, the other case being equal. Remark first that such a quantity is smooth in ϵ so that it can be expanded in Taylor series at any order. Thus the statement is equivalent to the fact that the coefficient of order zero in the expansion of $T \circ \mathcal{T}$ is the identity, while the coefficients of order from 1 to r vanish. To prove this consider the sequence $X_N^{(k)}$ generated by the vector field X_N according to (4.23). Since X_N is smooth (in the standard sense), (4.24) holds for it. Define the maps T_N and $\mathcal{T}_N$ according to (4.25) and (4.26) with X_N in place of X, then one has

$$\Phi_N^{\epsilon} = T_N + O(\epsilon^{r+1}) \ , \quad \Phi_N^{-\epsilon} = \mathcal{T}_N + O(\epsilon^{r+1}) \ ,$$

and

$$1 = \Phi_N^{\epsilon} \circ \Phi_N^{-\epsilon} = T_N \circ \mathcal{T}_N + \big(T_N \circ (\mathcal{T}_N + O(\epsilon^{r+1})) - T_N \circ \mathcal{T}_N\big) + O(\epsilon^{r+1}) \ ,$$

from which

$$1 = T_N \circ \mathcal{T}_N + O(\epsilon^{r+1}) \ .$$

It follows that

$$\frac{d^k}{d\epsilon^k}\Big|_{\epsilon=0} T_N \circ \mathcal{T}_N \equiv 0 \ , \quad \forall 1 \leq k \leq r \ , \quad \forall N \ . \tag{4.29}$$

However, by construction $T_N \to T$ and $\mathcal{T}_N \to \mathcal{T}$ in $C^{\mathtt{r}}(\mathcal{B}^{\mathtt{s}'}, \mathcal{B}^{\mathtt{s}})$ for all $\mathtt{r}$ as $N \to \infty$, thus one gets

$$\frac{d^k}{d\epsilon^k}\Big|_{\epsilon=0} T \circ \mathcal{T} \equiv 0 \ , \quad \forall 1 \le k \le r \ . \tag{4.30}$$

which is the thesis. □

An immediate corollary of the above result is the following one.

Corollary 4.8. *Let Y and X be almost smooth vector fields; fix $\mathtt{s}'$, then there exist $\mathtt{s}$, $\mathtt{s}''$ and $\mathcal{U}_{\mathtt{s}\mathtt{s}'} \subset \mathcal{B}^{\mathtt{s}}$ with the following property: let $\zeta \in C^1([-\mathtt{T}_0, \mathtt{T}_0]; \mathcal{U}_{\mathtt{s}\mathtt{s}'})$ be a solution of*

$$\dot{\zeta} = d\mathcal{T}(\zeta) Y(T(\zeta)) \ , \tag{4.31}$$

then there exists $R \in C^1([-\mathtt{T}_0, \mathtt{T}_0]; \mathcal{B}^{\mathtt{s}''})$ s.t. $z(.) := T(\zeta(.)) \in C^1([-\mathtt{T}_0, \mathtt{T}_0]; \mathcal{B}^{\mathtt{s}'})$ fulfills the equation

$$\dot{z} = Y(z) + \epsilon^{r+1} R(t) \ . \tag{4.32}$$

This is immediately seen by remarking that

$$\dot{z} = \frac{d}{dt} T(\zeta(t)) = dT(\zeta(t))\dot{\zeta} = dT(\zeta) d\mathcal{T}(\zeta) Y(T(\zeta)) = (1 + O(\epsilon^{r+1})) Y(T(\zeta)) \ .$$

We come to the Hamiltonian case. Let $G \in C^1(\mathcal{B}^{\mathtt{s}}, \mathbb{R})$ and $H \in C^1(\mathcal{B}^{\mathtt{s}}, \mathbb{R})$ be two Hamiltonian functions such that the corresponding Hamiltonian vector fields $X := J\nabla G$ and $J\nabla H$ are almost smooth. Define the transformation T according to (4.25) and define the sequence $H^{(l)}$ according to the recursive definition (4.5) and define

$$\widetilde{H} := \sum_{l=0}^{r} \frac{\epsilon^l}{l!} H^{(l)} \ , \tag{4.33}$$

then the main result of this section is the following Theorem

Theorem 4.9. *Fix $\mathtt{s}'$, then there exists $\mathtt{s}$, $\mathtt{s}''$ and $\mathcal{U}_{\mathtt{s}\mathtt{s}'} \subset \mathcal{B}^{\mathtt{s}}$ with the following property: let $\zeta \in C^1([-\mathtt{T}_0, \mathtt{T}_0]; \mathcal{U}_{\mathtt{s}\mathtt{s}'})$, $0 < \mathtt{T}_0 \le \infty$, be a solution of*

$$\dot{\zeta} = J\nabla \widetilde{H}(\zeta) \ , \tag{4.34}$$

then there exists $R \in C^1([-\mathtt{T}_0, \mathtt{T}_0]; \mathcal{B}^{\mathtt{s}'})$ s.t. $z(.) := T(\zeta(.)) \in C^1([-\mathtt{T}_0, \mathtt{T}_0]; \mathcal{B}^{\mathtt{s}''})$ fulfills the equation

$$\dot{z} = J\nabla H(z) + \epsilon^{r+1} R(t) \ . \tag{4.35}$$

The rest of the section is devoted to the proof of this theorem. We will proceed step by step.

Due to (3.1) and (3.3), we have to study $[dT(\zeta)]^*$, in particular in the case where $X = J\nabla G$. First we remark that, in this case, for any $z \in \mathcal{B}^{\mathtt{s}}$, $\mathtt{s}$ sufficiently large, we have

$$(d(\nabla G(z))^* = d\nabla G(z) \ , \tag{4.36}$$

where $(d\nabla G(z))^*$ is the adjoint with respect to the L^2 scalar product. Indeed, for $k \in \mathcal{B}^{\mathtt{s}}$, consider the differential of the map $z \mapsto \langle k; \nabla G(z)\rangle$ applied to a vector h. We have

$$d(\langle k; \nabla G\rangle)h = \langle k; d(\nabla G)h\rangle = d(dGh)k = d^2G(h,k) = d^2G(k,h) = \langle h; d(\nabla G)k\rangle \ ,$$

which is the thesis.

Furthermore, since $J^* = -J$ is bounded, it follows that if $X = J\nabla G$ is almost smooth, then also $(dX(z))^* = -d(\nabla G)J$ is almost smooth.

Lemma 4.10. *Let $X := J\nabla G$ be an almost smooth vector field, then $(dT(z))^*$ is also almost smooth.*

Proof. We prove the result by induction on the vector fields $X^{(k)}$. By the above remark the result is true for $X^{(0)}$. By (4.23) one has

$$dX^{(k)}h = d^2X^{(k-1)}(X,h) + dX^{(k-1)}dXh \ .$$

the adjoint of the second addendum is $dX^*[dX^{(k-1)}]^*$, so, by the induction assumption it is almost smooth. Consider now the first addendum. The adjoint $L(z)$ of the linear operator $d^2X^{(k-1)}(X,.)$ is defined by

$$\langle L(z)h_1; h_2\rangle = \langle d^2X^{(k-1)}(z)(X(z),h_2); h_1\rangle \ .$$

Therefore one has to show that $\forall l\ \exists k_1k_2$ s.t., if $z \in \mathcal{B}^{k_1}$, then $L(z) \in B(\mathcal{B}^{k_1}, \mathcal{B}^l)$ and furthermore the dependence on z is smooth. We start by fixing z, so that the statement is equivalent to the existence of a constant C s.t.

$$|\langle L(z)h_1; h_2\rangle| \leq C\|h_1\|_{\mathcal{B}^{k_2}}\|h_2\|_{\mathcal{B}^{-l}} \ . \tag{4.37}$$

Actually, it is convenient to fix the argument of X and to define the operator $L_1(z)$ by

$$\langle L_1(z)h_1; h_2\rangle = \langle d^2X^{(k-1)}(z)(X(z_1),h_2); h_1\rangle \tag{4.38}$$

with fixed z_1 in a sufficiently smooth space. It is clear that, due to the smooth dependence on z_1 it is sufficient to study the operator L_1. Furthermore we denote simply $X(z_1) = X$ Now (4.38) is equal to

$$\langle d\left(dX^{(k-1)}(z)h_2\right)X; h_2\rangle = d\left(\langle dX^{(k-1)}(z)h_2; h_2\rangle\right)X$$
$$= d\left(\langle h_2; [dX^{(k-1)}(z)]^* h_2\rangle\right)X \ ,$$

but, by the inductive assumption one has $[dX^{(k-1)}(.)]^* \in C^{\mathbf{r}}(\mathcal{B}^{k_1}, B(\mathcal{B}^{k_2}, \mathcal{B}^l))$, therefore, if $X \in \mathcal{B}^{k_1}$, which can be ensured by taking z_1 smooth enough, the above quantity is estimated by

$$C\|X\|_{\mathcal{B}^{k_1}}\|h_1\|_{\mathcal{B}^{k_2}}\|h_2\|_{\mathcal{B}^{-l}} \ ,$$

which is the estimate that we had to prove. Smooth dependence on z follows from the smooth dependence of $[dX^{(k-1)}(z)]^*$ on z. □

Lemma 4.11. *Assume that $X = J\nabla G$ is an almost smooth vector field, then one has*

$$dT(\zeta)J[dT(\zeta)]^* = J + O(\epsilon^{r+1}) \ . \tag{4.39}$$

Proof. Let $G_N(\zeta) := G(\Pi_N\zeta)$ and denote $X^N := J\nabla G_N(\zeta) = \Pi_N J(\nabla G)(\Pi_N\zeta)$. As before, consider the corresponding flow $\Phi^\epsilon_N = T_N + O(\epsilon^{r+1})$, which is a canonical transformation. Thus one has

$$\Pi_N J\Pi_N = d\Phi^\epsilon_N(\zeta)\Pi_N J\Pi_N[d\Phi^\epsilon_N(\zeta)]^* = dT_N(\zeta)\Pi_N J\Pi_N[dT_N(\zeta)]^* + O(\epsilon^{r+1}) \ . \tag{4.40}$$

It follows that, for all N and for $1 \le l \le r$, one has

$$dT_N(\zeta)\Pi_N J\Pi_N[dT_N(\zeta)]^*\big|_{\epsilon=0} = \Pi_N J\Pi_N \ , \tag{4.41}$$
$$\frac{d^l}{d\epsilon^l}\Big|_{\epsilon=0} dT_N(\zeta)\Pi_N J\Pi_N[dT_N(\zeta)]^* = 0 \ , \tag{4.42}$$

but all these objects converge as almost smooth operators when $N \to \infty$, and thus the thesis follows. □

Corollary 4.12. *Let $\zeta(t)$ be a sufficiently smooth solution of*

$$\dot{\zeta} = J\nabla(H \circ T)(\zeta) \ , \tag{4.43}$$

then $z(t) := T(\zeta(t))$ fulfills

$$\dot{z} = J\nabla H(z) + \epsilon^{r+1}R \ . \tag{4.44}$$

We are now ready for the proof of Theorem 4.9. The main point is that, from Corollary 4.12, in terms of the variables ζ, the system is Hamiltonian (up to a remainder of order ϵ^{r+1}) with Hamiltonian $H \circ T$. We now have have the following Lemma

Lemma 4.13. *One has*

$$H \circ T = \sum_{l=0}^{r} \frac{\epsilon^l}{l!} H^{(l)} + \epsilon^{r+1} R \ . \tag{4.45}$$

with $H^{(l)}$ defined by (4.5)*, and R having an almost smooth vector field.*

Proof. We start by showing that

$$H^{(l)} = \frac{d^l}{d\epsilon^l}\Big|_{\epsilon=0} H \circ T \ .$$

which would show that $\epsilon^{r+1}R$ is the remainder of the Taylor series of a smooth function (of ϵ) and therefore R is bounded uniformly with respect to ϵ. Consider the flow Φ_N^ϵ of the truncated vector field X_N, then one has

$$\begin{aligned} H \circ \Phi_N^\epsilon = H \circ (T_N + \epsilon^{r+1} R) = H \circ T_N + \big(H \circ (T_N + \epsilon^{r+1} R) - H \circ T_N\big) \\ = H \circ T_N + \epsilon^{r+1} R \ , \end{aligned}$$

so that

$$\frac{d^l}{d\epsilon^l}\Big|_{\epsilon=0} H \circ T_N = \frac{d^l}{d\epsilon^l}\Big|_{\epsilon=0} H \circ \Phi_N = H^{(l)}(\Pi_N .) \ , \quad \forall l \le r \ .$$

Since this quantity converges to $H^{(l)}$ as N tends to infinity, one has the thesis. Reasoning in the same way on $\nabla(H \circ T)$, we get

$$\frac{d^l}{d\epsilon^l}\Big|_{\epsilon=0} \left(\Pi_N \nabla H(\Pi_N \Phi_N^\epsilon)\right) = \Pi_N \nabla H^{(l)}(\Pi_N .) \ ,$$

which, passing to the limit $N \to \infty$ shows that

$$\frac{d^l}{d\epsilon^l}\Big|_{\epsilon=0} \nabla(H \circ T) = \nabla H^{(l)} \ .$$

Finally, by the almost smoothness of $\nabla(H \circ T)$, which follows from eq. (3.1) and Lemma 4.10, one has that $\epsilon^{r+1}\nabla R$ is the remainder of a Taylor series of a smooth function and thus the thesis follows. □

Proof of Theorem 4.9. By Lemma 4.13 one has $\widetilde{H} = H \circ T - \epsilon^{r+1} R_1$ with R_1 having an almost smooth vector field, thus $\zeta(t)$ fulfills

$$\dot{\zeta}(t) = J\nabla(H \circ T) - \epsilon^{r+1} J \nabla R_1(\zeta(t)) \ ,$$

therefore, using (3.3), we have

$$\begin{aligned} \dot{z}(t) &= dT(\zeta(t)) J \left[dT(\zeta(t))\right]^* \nabla H(T(\zeta(t))) - \epsilon^{r+1} dT(\zeta(t)) J \nabla R_1(\zeta(t)) \\ &= J\nabla H(z(t)) + \epsilon^{r+1} R_2(\zeta(t)) \nabla H(z(t)) - \epsilon^{r+1} dT(\zeta(t)) J \nabla R_1(\zeta(t)) \ , \end{aligned}$$

where we used Lemma 4.11. But such an equation is the thesis. □

5 Hamiltonian Normal form for the water wave problem.

In order to be able to apply the normal form procedure to the water wave problem, we must be able to solve the Homological equation. This is done with the help of a few lemmas. The first one is an abstract lemma, the other two are really adapted to water wave problem.

Consider the the homological equation

$$\{H_0; G\} + W = 0 \ . \tag{5.1}$$

Lemma 5.1. *Assume that, for* $\mathtt{s}$ *large enough, one has*

$$\lim_{\tau\to+\infty}(W(\Phi_{H_0}^{-\tau}(z)) + W(\Phi_{H_0}^{\tau}(z))) = 0 \ , \quad \forall z \in \mathcal{B}^{\mathtt{s}} \ ; \tag{5.2}$$

if the following function G *is well defined, then it solves the homological equation* (5.1)

$$G(z) := -\frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(\tau) W(\Phi_{H_0}^{\tau}(z)) d\tau \ . \tag{5.3}$$

Proof. Just compute

$$\{H_0; G\}(z) = -\frac{d}{dt}\Big|_{t=0} G(\Phi_{H_0}^{t}(z)) \tag{5.4}$$

$$= \frac{d}{dt}\Big|_{t=0}\frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(\tau) W(\Phi_{H_0}^{\tau+t}(z)) d\tau \tag{5.5}$$

$$= \frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(\tau)\frac{d}{d\tau} W(\Phi_{H_0}^{\tau}(z)) d\tau \tag{5.6}$$

$$= -\frac{1}{2}\int_{-\infty}^{0}\frac{d}{d\tau} W(\Phi_{H_0}^{\tau}(z)) d\tau + \frac{1}{2}\int_{0}^{+\infty}\frac{d}{d\tau} W(\Phi_{H_0}^{\tau}(z)) d\tau \tag{5.7}$$

$$= -W(\Phi_{H_0}^{0}(z)) + \frac{W(\Phi_{H_0}^{-\infty}(z)) + W(\Phi_{H_0}^{+\infty}(z))}{2} = -W(z) \tag{5.8}$$

□

Actually one can get an explicit formula for the solution of the Homological equation. Before giving the result, we study a few properties of the operator ∂^{-1} defined in (2.22). First we remark that one also has

$$(\partial^{-1}u)(y) = \frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(y - y_1) u(y_1) dy_1 \ , \tag{5.9}$$

and that $\partial^{-1} : L^1 \to L^\infty$ continuously. Then one has $\partial(\partial^{-1}u) = u$. Furthermore, if u is such that $\lim_{\tau\to+\infty}(u(\tau) + u(-\tau)) = 0$ then one also has

$\partial^{-1}u_y = u$. We also remark that the property is automatic for the functions of class $W^{2,1}$.

By the very definition of ∂^{-1}, its adjoint is $-\partial^{-1}$.

Finally we introduce a notation which is very useful in order to shorten the computations:

In the following we denote

$$r_k := \partial^k r \ , \quad s_k := \partial^k s \ , \quad k \geq -1 \ . \tag{5.10}$$

We will consider functionals W of the form

$$W(r,s) = \int_{\mathbb{R}} P_1(r_{-1}(y), r(y), r_1(y), ..., r_{n_1}(y)) P_2(s_{-1}(y), s(y), s_1(y), ..., s_{n_2}(y)) dy \ , \tag{5.11}$$

where $P_1 : \mathbb{R}^{n_1+2} \to \mathbb{R}$ and $P_2 : \mathbb{R}^{n_2+2} \to \mathbb{R}$ are polynomials. For brevity we will simply denote

$$P_1(r) := P_1(r_{-1}(y), r(y), r_1(y), ..., r_{n_1}(y)) \ .$$

Sometimes we will denote

$$P_1(r(y)) := P_1(r_{-1}(y), r(y), r_1(y), ..., r_{n_1}(y)) \ .$$

We have the following Lemma

Lemma 5.2. *Assume that,* $P_1(r) \in L^2(\mathbb{R})$ *whenever* $r \in W^{\mathtt{s},1}$*, for* $\mathtt{s} \gg 1$*, and similarly for* $P_2(s)$*. Then the solution* (5.3) *of the homological equation* (5.1) *with* W *given by* (5.11) *is given by*

$$G(r,s) := -\frac{1}{2} \int_{\mathbb{R}} \left[\partial^{-1} P_1(r)\right] P_2(s) dy \tag{5.12}$$

Proof. We start by verifying that W fulfills the assumption (5.2). Fix some K, one has

$$\begin{aligned} W(\Phi^t_{H_0}(r,s)) &= \int_{\mathbb{R}} P_1(r(y-t)) P_2(s(y+t)) dy = \int_{\mathbb{R}} P_1(r(y-2t)) P_2(s(y)) dy \\ &= \int_{-\infty}^{K} P_1(r(y-2t)) P_2(s(y)) dy + \int_{K}^{+\infty} P_1(r(y-2t)) P_2(s(y)) dy \ . \end{aligned} \tag{5.13}$$

Consider first the first integral. It is estimated by

$$\begin{aligned} &\left[\int_{-\infty}^{K} |P_1(r(y-2t))|^2 dy\right]^{1/2} \left[\int_{-\infty}^{K} |P_2(s(y))|^2 dy\right]^{1/2} \\ &\qquad \leq \|P_2(s)\|_{L^2} \left[\int_{-\infty}^{K-2t} |P_1(r(y))|^2 dy\right]^{1/2} \ , \end{aligned}$$

but the last factor tends to zero when $t \to +\infty$, due to the fact that $P_1(r)$ is square integrable. Treating the second integral in (5.13) in a similar way we get that $\lim_{t\to+\infty} W(\Phi^t_{H_0}(r,s)) = 0$. In a similar way one gets $\lim_{t\to-\infty} W(\Phi^t_{H_0}(r,s)) = 0$.

We now use the formula (5.3) to compute G. Making the change of variables

$$y_1 = y - \tau \ , \quad y_2 = y + \tau \ ,$$

one has

$$G = -\frac{1}{2}\int_{\mathbb{R}} d\tau \mathrm{sgn}(\tau) \int_{\mathbb{R}} dy P_1(r(y-\tau))P_2(s(y+\tau)) \tag{5.14}$$

$$= \frac{1}{2}(-)\frac{1}{2}\int_{\mathbb{R}^2} \mathrm{sgn}(y_2 - y_1)P_1(r(y_1))P_2(s(y_2))dy_1 dy_2 \tag{5.15}$$

$$= -\frac{1}{2}\int_{\mathbb{R}} dy_2[(\partial^{-1}P_1(r))(y_2)]P_2(s(y_2)) \ . \tag{5.16}$$

□

Actually we do not have an abstract theorem ensuring that the Hamiltonian vector field of G is an almost smooth map. We now compute explicitly the second order normal form and compute the structure of the first two generating functions in order to show that their vector field is almost smooth. Furthermore we compute some terms of G_3 in order to show that the corresponding vector field is not well defined, so that we cannot perform (at least with this algorithm) a third step completely eliminating the interaction between right going waves and left going waves.

In order to simplify the notation and the computation, given a functional which is of the form

$$W(r,s) = \int_{\mathbb{R}} w(r(y), s(y))dy \ , \tag{5.17}$$

with $w(r,s) = P_1(r)P_2(s)$, we will always denote by lower case letter the density which is integrated to get the functional denoted with the corresponding capital letter.

Remark 5.3. *Given a functional W as in* (5.17)*, the corresponding gradient is given by*

$$\nabla_r W(r,s) = \sum_{k\geq -1}^{n_1} (-\partial^k)\frac{\partial w}{\partial r_k} \tag{5.18}$$

and similarly for the gradient with respect to the s variable.

Lemma 5.4. *Assume that W is of the form* (5.11) *with P_1 and P_2 fulfilling the assumptions of Lemma 5.2. Assume also that P_1 and P_2 are monomyals that do not depend on r_{-1} and s_{-1} respectively. Then the solution G of the homological equation* (5.1) *defined by* (5.12) *has an almost smooth vector field.*

Proof. Up to the factor 1/2 and exploiting the skew symmetry of ∂^{-1}, one has $g = P_1(r)\partial^{-1}P_2(s)$, from which

$$-\partial\nabla_r G = -\partial \sum_{k\geq 0}^{n_1} (-\partial)^k \left(\frac{\partial P_1}{\partial r_k} \partial^{-1} P_2 \right)$$

$$= (\partial^{-1}P_2(s)) \sum_{k\geq 0}^{n_1} \left((-\partial)^{k+1} \frac{\partial P_1}{\partial r_k} \right) + \text{local terms}$$

where, by local terms, we mean terms not involving ∂^{-1}.

We prove now that $\forall k \geq 0$, $(-\partial)^{k+1}\frac{\partial P_1}{\partial r_k} \in W^{\mathtt{s},1}$. Indeed, if $\frac{\partial P_1}{\partial r_k}$ is not a constant, then the result follows from the algebra property of $W^{\mathtt{s},1}$, while, if it is a constant, then ∂^{k+1} annihilates it. Thus, since the product of a function of class L^1 and a function of class L^∞ is still of class L^1 the result follows for the r component. Similarly one gets the result for the s component. $\square$

We now proceed in the explicit computation of z_i, w_i and g_i.

Consider H_1 as given (3.31), (3.32), in which $Z_1 =$ (3.31) and $W_1 =$ (3.32), so that one has

$$z_1 = -\frac{1}{12}(r_1^2 + s_1^2) + \frac{r^3 + s^3}{4} \,, \quad w_1 = \frac{r_1 s_1}{6} - \frac{r^2 s + r s^2}{4} \,. \tag{5.19}$$

From this, by Lemma 5.2 and the skew-symmetry of ∂^{-1},

$$g_1 = \frac{r_1 s}{12} - \frac{r^2 s_{-1} - r_{-1} s^2}{8} \,, \tag{5.20}$$

In particular, by Lemma 5.4 we know that its vector field is almost smooth.

Furthermore, one has

$$\nabla_r W_1 = -\frac{1}{6}s_2 - \frac{rs}{2} - \frac{s^2}{4} \tag{5.21}$$

$$\nabla_s W_1 = -\frac{1}{6}r_2 - \frac{rs}{2} - \frac{r^2}{4} \tag{5.22}$$

$$\nabla_r G_1 = -\frac{r s_{-1}}{4} - \partial^{-1}\frac{s^2}{8} - \frac{s_1}{12} \,, \tag{5.23}$$

$$\nabla_s G_1 = \frac{s r_{-1}}{4} + \partial^{-1}\frac{r^2}{8} + \frac{r_1}{12} \,. \tag{5.24}$$

So, in particular the vector field of G_1 is

$$(r-\text{component}) = \frac{r_1 s_{-1} + rs}{4} + \frac{s^2}{8} + \frac{s_2}{12} \tag{5.25}$$

$$(s-\text{component}) = \frac{s_1 r_{-1} + rs}{4} + \frac{r^2}{8} + \frac{r_2}{12} \; . \tag{5.26}$$

Remark 5.5. *If in the expressions of the vector field of G_1 we neglect the nonlinear terms, the corresponding equations of motion turn out to be*

$$\begin{cases} \dot r = \frac{s_2}{12} \\ \dot s = \frac{r_2}{12} \end{cases} \Longrightarrow \begin{cases} \ddot r = \frac{r_4}{144} \\ \ddot s = \frac{s_4}{144} \end{cases} \tag{5.27}$$

which is clearly ill posed. It follows in particular that the problem of existence and uniqueness for the Hamilton equations of G_1 is a nontrivial one. With our approach we do not need to study it.

We now compute $H_{2,1} = W_2 + Z_2$ (cf. (4.13)), in particular we will get an explicit expression for the terms contributing to Z_2. For the terms contributing to W_2, we will neglect the precise value of the coefficients of the various terms, that will be conventionally put equal to 1.

First remark that $\{Z_1; G_1\}$ does not contribute to Z_2, so we will only compute its general structure.

To start with compute (with a small abuse of notation)

$$\begin{aligned}
\{W_1; G_1\} &= \langle \nabla_s W_1; \partial \nabla_s G_1 \rangle - \langle \nabla_r W_1; \partial \nabla_r G_1 \rangle \\
&= \left(-\frac{1}{6} r_2 - \frac{rs}{2} - \frac{r^2}{4}\right)\left(\frac{s_1 r_{-1} + rs}{4} + \frac{r^2}{8} + \frac{r_2}{12}\right) \\
&\quad + \left(-\frac{1}{6} s_2 - \frac{rs}{2} - \frac{s^2}{4}\right)\left(\frac{r_1 s_{-1} + rs}{4} + \frac{s^2}{8} + \frac{s_2}{12}\right) \\
&= -\frac{1}{24} r^2 r_2 - \frac{1}{72} r_2^2 - \frac{1}{32} r^4 - \frac{1}{24} s^2 s_2 - \frac{1}{72} s_2^2 - \frac{1}{32} s^4 \\
&\quad + r_2 s_1 r_{-1} + r_2 rs + rss_1 r_{-1} + r^2 s^2 + sr^3 + r^2 s_1 r_{-1} + s_2 r_1 s_{-1} \\
&\quad + s_2 rs + rsr_1 s_{-1} + rs^3 + rss_2 + s^2 r_1 s_{-1} + rs^3
\end{aligned}$$

while we have

$$\begin{aligned}
\{Z_1; G_1\} &= (s_2 + s^2)(s_1 r_{-1} + rs + r^2 + r_2) + (r_2 + r^2)(r_1 s_{-1} + rs + s^2 + s_2) \\
&= s_2 s_1 r_{-1} + s_2 r^2 + s_2 r_2 + s^2 r^2 + s^2 r_2 + r_2 r_1 s_{-1} + r_2 s^2 + r_2 r_1 s_{-1} \\
&\quad + \text{terms already contained in } \{W_1; G_1\}
\end{aligned}$$

Integrating by parts the first term of $\frac{1}{2}\{W_1; G_1\}$ and adding the terms coming from H_2, we have

$$z_2 = -\frac{5}{24} r r_1^2 + \frac{19}{720} r_2^2 - \frac{1}{64} r^4 \tag{5.28}$$

$$-\frac{5}{24} s s_1^2 + \frac{19}{720} s_2^2 - \frac{1}{64} s^4 \ , \tag{5.29}$$

so that, its gradient and the corresponding vector field are given by

$$\nabla_r Z_2 = \frac{5}{24} r_1^2 + \frac{5}{12} r r_2 + \frac{19}{360} r_4 - \frac{1}{16} r^3 \tag{5.30}$$

$$-\partial \nabla_r Z_2 = -\frac{5}{6} r_1 r_2 - \frac{5}{12} r r_3 - \frac{19}{360} r_5 + \frac{3}{16} r^2 r_1 \ . \tag{5.31}$$

Concerning W_2, one has

$$w_2 = r_2 s_1 r_{-1} + r_2 r s + r s s_1 r_{-1} + r^2 s^2 + s r^3 + r^2 s_1 r_{-1} + s_2 r_1 s_{-1} \tag{5.32}$$

$$+ s_2 r s + r s r_1 s_{-1} + r s^3 + r s s_2 + s^2 r_1 s_{-1} + r s^3 \tag{5.33}$$

$$+ s_2 s_1 r_{-1} + s_2 r^2 + s_2 r_2 + s^2 r^2 + s^2 r_2 + r_2 r_1 s_{-1} + r_2 s^2 \ . \tag{5.34}$$

Lemma 5.6. *Let G_2 be given by* (5.12) *with P_1 and P_2 given by the different terms of* (5.32)-(5.34)*. Then the vector field of G_2 is almost smooth.*

Proof. According to Lemma 5.4, we only have to check the terms coming from non local terms in w_2, namely

$$\begin{aligned} w_2^{nl} &= r_2 s_1 r_{-1} + r s s_1 r_{-1} + r^2 s_1 r_{-1} + s_2 r_1 s_{-1} + r s r_1 s_{-1} + s^2 r_1 s_{-1} \\ &\qquad + s_2 s_1 r_{-1} + r_2 r_1 s_{-1} \\ &= r_2 r_{-1} s_1 + r r_{-1} \partial(s^2) + r^2 r_{-1} s_1 + r_1 s_2 s_{-1} + \partial(r^2) s s_{-1} + r_1 s_{-1} s^2 \\ &\qquad + r_{-1} \partial(s_1^2) + \partial(r_1^2) s_{-1} \ , \end{aligned}$$

from which

$$g_2^{nl} = r_2 r_{-1} s + r r_{-1} s^2 + r^2 r_{-1} s + r s_2 s_{-1} + r^2 s s_{-1} + r s_{-1} s^2 + r_{-1} s_1^2 + r_1^2 s_{-1} \ .$$

By the same argument as in the proof of Lemma 5.4, the vector field corresponding to each term of the above equation has an almost smooth vector field. □

As a consequence one can use G_2 to put the system in normal form at order ϵ^2. To give a precise statement consider the Hamiltonian

$$H_Z(r, s) := H_0(r, s) + \epsilon Z_1(r, s) + \epsilon^2 Z_2(r, s) \ , \tag{5.35}$$

with Z_1 given by (5.19) and Z_2 by (5.28).

Theorem 5.7. *For any* $\mathtt{s}'$ *there exists* $\epsilon_* > 0$ *and* $\mathtt{s}$, $\mathtt{s}''$, *s.t., if* $0 < \epsilon < \epsilon_*$, *then there exists a map* $T_H : B_1^{\mathtt{s}} \to \mathcal{B}^{\mathtt{s}''}$, *with the following properties*

(i) $T_H(r,s) - (r,s)$ *is a polynomial in* r_k, s_k, $k = -1, ..., 5$,

(ii) $\sup_{(r,s)\in B_1^{\mathtt{s}}} \|T_H(r,s) - (r,s)\|_{\mathcal{B}^{\mathtt{s}''}} \le \epsilon$,

(iii) Let I_ϵ *be an interval containing 0 and let* $z(.) = (r(.), s(.)) \in C^1(I_\epsilon; B_1^{\mathtt{s}})$ *be a solution of the Hamiltonian system* (5.35) *define*

$$z_h \equiv (r_h, s_h) := T_H(r,s) \ . \tag{5.36}$$

Then there exists $R \in C^1(I_\epsilon, W^{\mathtt{s}',2} \times W^{\mathtt{s}',2})$ *s.t. one has*

$$\dot{z}_h(t) = J\nabla H_{WW}(z_h(t)) + \epsilon^3 R(t) \ , \quad \forall t \in I_\epsilon \ , \tag{5.37}$$

where H_{WW} *is the Hamiltonian* (2.9) *of the water wave problem rewritten in the variables* (r,s).

Proof. Define $X_1 := J\nabla G_1$ with G_1 given by (5.20) and

$$T_1(z) := z + \epsilon X_1(z) + \epsilon^2 dX_1(z) X_1(z) \ , \tag{5.38}$$

define also

$$T_2(z) = z + J\nabla G_2(z) \ , \tag{5.39}$$

with G_2 as described in the statement of Lemma 5.6. Define $T_H := T_1 \circ T_2$, then Theorem 4.9 shows that there exists $R_h \in C^1(I_\epsilon, W^{\mathtt{s}'+1,1} \times W^{\mathtt{s}'+1,1})$ s.t.

$$\dot{z}_h = J\nabla(H_0 + \epsilon H_1 + \epsilon H_2) + \epsilon^3 R_h \ .$$

Adding the remainder coming from the truncation of the Hamiltonian (c.f. Proposition 3.6) and exploiting the embedding $W^{\mathtt{s}'+1,1} \times W^{\mathtt{s}'+1,1} \subset W^{\mathtt{s}',2} \times W^{\mathtt{s}',2}$ one gets the result. □

Then one would like to make at least a third step. As we anticipated there are obstructions that we now describe. Using the formula (4.17), one sees that W_3 contains in particular the term $\{Z_2, G_1\}$. Thus in particular it contains a monomyal coming from the terms r_2^2 in Z_2 and the term $r^2 s_{-1}$ in g_1. This gives rise to a non local term in W_3 which is

$$w_3^{bad} := r_4 r_1 s_{-1} \ ,$$

which in turn give rise to

$$g_3^{bad} = \partial^{-1}(r_4 r_1) s_{-1} \ ,$$

whose integral over $\mathbb{R}$ is, in general infinite. Even working formally, one can compute the corresponding term in the Hamiltonian vector field. It is given by

$$\partial_4(r_1 s_{-2}) + \partial(r_4 s_{-2}) = r_5 s_{-2} + \text{local terms} \ ,$$

which is not well defined, since the operator ∂^{-2} is in general not defined on $W^{\mathtt{s},1}$.

Actually this argument is not conclusive, since there could be terms compensating w_3^{bad} or additive terms which transform such a term in something of the form $\partial^3(r_1)s_{-1}$, which would give rise to well behaved terms. However the verification of this requires much longer computations that we leave for future work.

6 Kodama's theory

By using Kodama's theory, one immediately deduces Theorem 2.2 from Theorem 5.7. For the sake of completeness we are now going to summarize Kodama's theory.

Given a Hamiltonian system of the form

$$K_0(s) + \epsilon K_1(s) + \epsilon^2 Z_2(s) \ , \tag{6.1}$$

with K_0 and K_1 given by (2.17) and (2.18) and

$$Z_2(s) = \int_{\mathbb{R}} \left(b_1 s s_1^2 + b_2 s_2^2 + b_3 s^4 \right) dy \ , \quad b_j \in \mathbb{R} \tag{6.2}$$

Kodama [Kod85, Kod87b, Kod87a] (but we make here reference to the review paper [HK09]) has shown that there exists a coordinate transformation of the form

$$s = T_K(u) := u + \epsilon \mathcal{X}(u) + \epsilon^2 d\mathcal{X}(u)\mathcal{X}(u) \ , \tag{6.3}$$

$$\mathcal{X}(u) := a_1 u^2 + a_2 u_2 + a_3 u_1 u_{-1} \ , \quad a_j \in \mathbb{R} \tag{6.4}$$

which conjugates the Hamilton equations of (6.1) to the Hamilton equations of

$$K_0(u) + \epsilon K_1(u) + \epsilon^2 c_2 K_2(u) \ , \tag{6.5}$$

with K_2 given by (2.19) and a suitable c_2 to be determined.

To see this, first consider the Hamilton equations of (6.1), which have the form

$$\dot{s} = \mathcal{Y}_0(s) + \epsilon \mathcal{Y}_1(s) + \epsilon^2 \mathcal{Y}_2(s) \ , \tag{6.6}$$

with $\mathcal{Y}_0(s) = \partial_y \nabla K_0(s)$ and so on. Then remark that (6.3) is just the second order expansion of the time ϵ flow of the auxiliary equation $\dot{u} = \mathcal{X}(u)$. Thus, by repeating at a non Hamiltonian level the computations of sect. 4.3, one gets that the equations fulfilled by u (as defined by (6.3)) are

$$\dot{u} = \widetilde{\mathcal{Y}}(u) \ , \tag{6.7}$$

with

$$\begin{aligned}\widetilde{\mathcal{Y}} &= \mathcal{Y}_0 + \epsilon[\mathcal{Y}_0; \mathcal{X}] + \frac{\epsilon^2}{2}\left[[\mathcal{Y}_0; \mathcal{X}] \, ; \mathcal{X}\right] + \epsilon \mathcal{Y}_1 + \epsilon^2 \left[\mathcal{Y}_1; \mathcal{X}\right] + \epsilon^2 \mathcal{Y}_2 + O(\epsilon^3) \\ &= \mathcal{Y}_0 + \epsilon \mathcal{Y}_1 + \epsilon^2 \left(\mathcal{Y}_2 + [\mathcal{Y}_1; \mathcal{X}]\right) + O(\epsilon^3) \ ,\end{aligned}$$

where we denoted

$$\left[\mathcal{Y}; \mathcal{X}\right](u) := d\mathcal{Y}(u)\mathcal{X}(u) - d\mathcal{X}(u)\mathcal{Y}(u) \ ,$$

and used the fact that $[\mathcal{Y}_0; \mathcal{X}] = 0$, since $\mathcal{Y}_0$ is the generator of the translations (and also follows by direct computation).

Thus, one has to look for the values of the constants a_j in (6.4) such that

$$\mathcal{Y}_2 + [\mathcal{Y}_1; \mathcal{X}] = c_2 \partial \nabla K_2 \equiv c_2(u_5 + 5u_3 u + 10 u_2 u_1 + \frac{35}{8} u_1 u^2) \ . \tag{6.8}$$

Warning: *I am not sure that the coefficient present in the formulae below are correct. Nevertheless Kodama's theory ensures that there exist correct coefficients and that the linear system that one gets is solvable.*

Anyway, I think that the following presentation can be useful in order to allow to repeat, maybe with a symbolic manipulator, the computation and to get a value, in particular of the constant c_2^*, which is surely correct.

Now, a long, but straightforward computation shows that

$$\begin{aligned}\mathcal{Y}_2 + [\mathcal{Y}_1; \mathcal{X}] &= \left(\frac{1}{3}a_2 + 2b_2\right) u_5 + \left(\frac{1}{3}a_1 + \frac{1}{2}a_3 - 2b_1\right) u_3 u \\ &+ \left(a_1 - 2a_2 + \frac{5}{6}a_3 - 4b_1\right) u_2 u_1 + \left(\frac{7}{2}a_1 + \frac{3}{4}a_3 + 7b_3\right) u_1 u^2 \ .\end{aligned}$$

this leads to impose the system for the unknowns (a_1, a_2, a_3, c_2)

$$\begin{aligned}c_2 &= \frac{1}{3}a_2 + 2b_2 \\ 5c_2 &= \frac{1}{3}a_1 + \frac{1}{2}a_3 - 2b_1 \\ 10c_2 &= a_1 - 2a_2 + \frac{5}{6}a_3 - 4b_1 \\ \frac{35}{8}c_2 &= \frac{7}{2}a_1 + \frac{3}{4}a_3 + 7b_3\end{aligned}$$

which can be solved explicitly, giving in particular

$$c_2 = (7b_3 + 3b_1 + 81b_2)\,\frac{8}{389} = \frac{299}{389} =: c_2^* \ , \tag{6.9}$$

where the last equality is obtained by inserting the values of b_j coming from (5.28).

To conclude the proof one has just to define the transformation

$$T_\epsilon := T_H \circ T_K \tag{6.10}$$

and remark that it still has the property (i) of Theorem 2.2.

References

[Bam05] Dario Bambusi. Galerkin averaging method and Poincaré normal form for some quasilinear PDEs. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 4(4):669–702, 2005.

[Bam13] Dario Bambusi. Asymptotic stability of ground states in some Hamiltonian PDEs with symmetry. *Comm. Math. Phys.*, 320(2):499–542, 2013.

[BCL05] Jerry L. Bona, Thierry Colin, and David Lannes. Long wave approximations for water waves. *Arch. Ration. Mech. Anal.*, 178(3):373–410, 2005.

[BCP02] D. Bambusi, A. Carati, and A. Ponno. The nonlinear Schrödinger equation as a resonant normal form. *Discrete Contin. Dyn. Syst. Ser. B*, 2(1):109–128, 2002.

[BD18] Massimiliano Berti and Jean-Marc Delort. *Almost global solutions of capillary-gravity water waves equations on the circle*, volume 24 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Cham; Unione Matematica Italiana, [Bologna], 2018.

[BFP18] Massimiliano Berti, Roberto Feola, and Fabio Pusateri. Birkhoff normal form and long time existence for periodic gravity water waves, 2018.

[BG93] D. Bambusi and A. Giorgilli. Exponential stability of states close to resonance in infinite-dimensional Hamiltonian systems. *J. Statist. Phys.*, 71(3-4):569–606, 1993.

[BP06] Dario Bambusi and Antonio Ponno. On metastability in FPU. *Comm. Math. Phys.*, 264(2):539–561, 2006.

[CG94] Walter Craig and Mark D. Groves. Hamiltonian long-wave approximations to the water-wave problem. *Wave Motion*, 19(4):367–389, 1994.

[CS93] W. Craig and C. Sulem. Numerical simulation of gravity waves. *J. Comput. Phys.*, 108(1):73–83, 1993.

[DGH03] Holger R. Dullin, Georg A. Gottwald, and Darryl D. Holm. Camassa-Holm, Korteweg-de Vries-5 and other asymptotically equivalent equations for shallow water waves. volume 33, pages 73–95. 2003. In memoriam Prof. Philip Gerald Drazin 1934–2002.

[HK09] Y. Hiraoka and Y. Kodama. Normal form and solitons. In *Integrability*, volume 767 of *Lecture Notes in Phys.*, pages 175–214. Springer, Berlin, 2009.

[Kod85] Yuji Kodama. Normal forms for weakly dispersive wave equations. *Phys. Lett. A*, 112(5):193–196, 1985.

[Kod87a] Yuji Kodama. Normal form and solitons. In *Topics in soliton theory and exactly solvable nonlinear equations (Oberwolfach, 1986)*, pages 319–340. World Sci. Publishing, Singapore, 1987.

[Kod87b] Yuji Kodama. On solitary-wave interaction. *Phys. Lett. A*, 123(6):276–282, 1987.

[Lan13] David Lannes. *The water waves problem*, volume 188 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2013. Mathematical analysis and asymptotics.

[Lan19] David Lannes. Modeling shallow water waves. *Preprint*, 2019.

[SW00] Guido Schneider and C. Eugene Wayne. The long-wave limit for the water wave problem. I. The case of zero surface tension. *Comm. Pure Appl. Math.*, 53(12):1475–1535, 2000.

[Wri05] J. Douglas Wright. Corrections to the KdV approximation for water waves. *SIAM J. Math. Anal.*, 37(4):1161–1206, 2005.

[Zak68] 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.