*Proceedings of the Royal Society of Edinburgh*, **149**, 191–217, *2019*
DOI:10.1017/S0308210518000227

# The Korteweg–de Vries, Burgers and Whitham limits for a spatially periodic Boussinesq model

**Roman Bauer**, **Wolf-Patrick Düll** and **Guido Schneider**
Institut für Analysis, Dynamik und Modellierung,
Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany
(roman.bauer@mathematik.uni-stuttgart.de;
duell@mathematik.uni-stuttgart.de;
guido.schneider@mathematik.uni-stuttgart.de)

(MS received 19 August 2016; accepted 30 December 2016)

We are interested in the Korteweg–de Vries (KdV), Burgers and Whitham limits for a spatially periodic Boussinesq model with non-small contrast. We prove estimates of the relations between the KdV, Burgers and Whitham approximations and the true solutions of the original system that guarantee these amplitude equations make correct predictions about the dynamics of the spatially periodic Boussinesq model over their natural timescales. The proof is based on Bloch wave analysis and energy estimates and is the first justification result of the KdV, Burgers and Whitham approximations for a dispersive partial differential equation posed in a spatially periodic medium of non-small contrast.

*Keywords:* spatially periodic Boussinesq model; Korteweg–de Vries (KdV) equation; Burgers equation; Whitham system; approximation

2010 *Mathematics subject classification:* Primary 35Q53

## 1. Introduction

In the long wave limit there exist a zoo of amplitude equations that can be derived via multiple scaling analysis for various dispersive wave systems with conserved quantities. Among these there are only three generic, nonlinear amplitude equations that are independent of the small perturbation parameter: the Korteweg–de Vries (KdV) equation, the inviscid Burgers equation and the Whitham system. In this paper we shall discuss the validity of these approximations for a spatially periodic Boussinesq model with non-small contrast.

### 1.1. The formal approximations in the spatially homogeneous situation

The KdV equation occurs as an amplitude equation in the description of small spatial and temporal modulations of long waves in various dispersive wave systems. Examples are the water wave problem or equations from plasma physics (see [3]). For the Boussinesq equation

$$\partial_t^2 u(x,t) = \partial_x^2 u(x,t) - \partial_x^4 u(x,t) + \partial_x^2 (u(x,t)^2), \tag{1.1}$$

with $x \in \mathbb{R}$, $t \in \mathbb{R}$ and $u(x,t) \in \mathbb{R}$, by the ansatz

$$u(x,t) = \varepsilon^2 A(X,T), \tag{1.2}$$

where $X = \varepsilon(x-t)$, $T = \varepsilon^3 t$, $A(X,T) \in \mathbb{R}$ and $0 < \varepsilon \ll 1$ is a small perturbation parameter, the KdV equation

$$\partial_T A = \tfrac{1}{2}\partial_X^3 A - \tfrac{1}{2}\partial_X(A^2) \tag{1.3}$$

can be derived by inserting (1.2) into (1.1) and by equating the coefficients in front of $\varepsilon^6$ to zero. This ansatz can be generalized to

$$u(x,t) = \varepsilon^\alpha A(X,T), \tag{1.4}$$

where $X = \varepsilon(x-t)$, $T = \varepsilon^{1+\alpha} t$ and $A(X,T) \in \mathbb{R}$ with $\alpha > 0$. For $\alpha > 2$ the Airy equation $\partial_T A = \frac{1}{2}\partial_X^3 A$ occurs. The KdV equation is recovered for $\alpha = 2$, and for $\alpha \in (0,2)$ the inviscid Burgers equation,

$$\partial_T A = -\tfrac{1}{2}\partial_X(A^2), \tag{1.5}$$

is obtained. There is another long wave limit that leads to an $\varepsilon$-independent non-trivial amplitude equation. With the ansatz

$$u(x,t) = U(X,T), \tag{1.6}$$

where $X = \varepsilon x$, $T = \varepsilon t$ and $U(X,T) \in \mathbb{R}$, we obtain

$$\partial_T^2 U = \partial_X^2 U + \partial_X^2(U^2), \tag{1.7}$$

which can be written as a system of conservation laws:

$$\partial_T U = \partial_X V, \qquad \partial_T V = \partial_X U + \partial_X(U^2). \tag{1.8}$$

In the following, (1.7) and (1.8) are called the Whitham system (see [25]).

### 1.2. Justification by error estimates

Estimating that the formal KdV approximation and true solutions of the original system stay close together over the natural KdV timescale is a non-trivial task, since solutions of order $\mathcal{O}(\varepsilon^2)$ have to be shown to exist on an $\mathcal{O}(1/\varepsilon^3)$ timescale. For (1.1), an approximation result is formulated as follows.

Theorem 1.1. *Let $s \geqslant 0$ and let $A \in C([0,T_0], H^{5+s})$ be a solution of the KdV equation (1.3). Then there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have solutions $u$ of (1.1) with*

$$\sup_{t\in[0,T_0/\varepsilon^3]} \|u(\cdot,t) - \varepsilon^2 A(\varepsilon(\cdot - t), \varepsilon^3 t)\|_{H^{1+s}} \leqslant C\varepsilon^{7/2}.$$

There are two fundamentally different approaches to proving such an approximation result. For analytic initial conditions of the KdV equation, a Cauchy–Kowalevskaya-based approach can be chosen (cf. [19] with the comments given in [21] for the water wave problem). Working in analytic function spaces gives some artificial smoothing that allows us to gain the missing order with respect to (w.r.t.)

$\varepsilon$ between the inverse of the amplitude of $\mathcal{O}(\varepsilon^2)$ and the timescale of $\mathcal{O}(1/\varepsilon^3)$ via the derivative in front of the nonlinear terms in the KdV equation. This approach is very robust and works without a detailed analysis of the underlying problem (see [7] for another example), but gives non-optimal results.

For initial conditions in Sobolev spaces, the underlying idea to gain such estimates is conceptually rather simple, i.e. the construction of a suitable chosen energy that includes $\mathcal{O}(\varepsilon^2)$ terms in the equation for the error, such that, for the energy, $\mathcal{O}(\varepsilon^3 t)$ growth rates finally occur. However, the method is less than robust, since a different energy occurs for every original system and the construction of this energy is a major difficulty. Estimates using this approach showing that the formal KdV approximation and true solutions of the different formulations of the water wave problem stay close together over the natural KdV timescale appear in, for example, [1, 10, 13, 23, 24]. Another example is the justification of the KdV approximation for modulations of periodic waves in the nonlinear Schrödinger (NLS) equation (see [8]). For (1.1) the energy approach is rather short and very instructive for the subsequent analysis. Therefore, we recall it in § 2.

Interestingly, it turns out that the proofs given for the KdV approximations transfer more or less line for line into proofs for the justification of the inviscid Burgers equation and of the Whitham system. Since only the scaling has to be adapted, whenever a KdV approximation result holds, inviscid Burgers and Whitham approximation results can also be established. This will be explained in detail in § 2.

As above, obtaining such approximation results is a non-trivial task since solutions of order $\mathcal{O}(\varepsilon^\alpha)$ have to be shown to exist on an $\mathcal{O}(\varepsilon^{1+\alpha})$ timescale. For the inviscid Burgers equation the approximation result is formulated along the lines of theorem 1.1. However, due to the notational complexity in achieving general estimates for the residual (the terms that do not cancel after inserting the approximation into (1.9)), in remark 2.3 we restrict ourselves to the $\alpha = 1$ case.

THEOREM 1.2. *Let $s \geqslant 0$, $\alpha = 1$ and let $A \in C([0, T_0], H^{3+s})$ be a solution of the inviscid Burgers equation (1.5). Then there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have solutions $u$ of (1.1) with*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u(\cdot,t) - \varepsilon A(\varepsilon(\cdot - t), \varepsilon^2 t)\|_{H^{1+s}} \leqslant C\varepsilon^{(3+2\alpha)/2}.$$

Since solutions of order $\mathcal{O}(1)$ are considered for the Whitham approximation, some smallness condition is needed such that the energy used allows us to estimate the associated Sobolev norm.

For (1.1) a possible Whitham approximation result is formulated as follows.

THEOREM 1.3. *Let $s \geqslant 0$. There exists a $C_1 > 0$ such that the following holds. Let $U \in C([0, T_0], H^{3+s})$ be a solution of (1.7) with*

$$\sup_{T\in[0,T_0]} \|U(\cdot,T)\|_{H^{3+s}} \leqslant C_1.$$

*Then there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have solutions $u$ of (1.1) with*

$$\sup_{t\in[0,T_0/\varepsilon]} \|u(\cdot,t) - U(\varepsilon\cdot, \varepsilon t)\|_{H^{1+s}} \leqslant C\varepsilon^{3/2}.$$

The Whitham system for the water wave problem coincides with the shallow water wave equations that were justified for the water wave problem without surface tension in [18,20]. A Whitham approximation result that the periodic wave trains of the NLS equation are approximated by the Whitham system can be found in [14].

### 1.3. The spatially periodic situation

The last few years have seen some initial attempts to justify the KdV equation in periodic media. In [17] it has been justified for the water wave problem over a periodic bottom in the KdV scaling, i.e. with long wave oscillations of the bottom of magnitude $\mathcal{O}(\varepsilon^2)$, varying on a spatial scale of order $\mathcal{O}(\varepsilon^{-1})$. The same result can be found in [5], where general bottom topographies of small amplitude have been handled. The result is based on that in [4], where other amplitude systems have been justified. This situation can be handled as a perturbation of the spatially homogeneous case.

In the case of $\mathcal{O}(1)$ oscillations of the bottom varying on a spatial scale of order $\mathcal{O}(1)$, no approximation result can be found in the existing literature. As a first attempt to solve this case for the water wave problem, we consider a spatially periodic Boussinesq equation,

$$\partial_t^2 u(x,t) = \partial_x(a(x)\partial_x u(x,t)) - \partial_x^2(b(x)\partial_x^2 u(x,t)) + \partial_x(c(x)\partial_x(u(x,t)^2)), \tag{1.9}$$

with $x \in \mathbb{R}$, $t \geqslant 0$, $u(x,t) \in \mathbb{R}$ and smooth $x$-dependent $2\pi$-spatially periodic coefficients $a$, $b$ and $c$ satisfying

$$\inf_{x\in\mathbb{R}} a(x) > 0 \quad \text{and} \quad \inf_{x\in\mathbb{R}} b(x) > 0.$$

For this equation we derive the KdV equation by making a Bloch mode expansion of (1.9). The KdV approximation describes the modes that are contained in the circles in figure 1. We prove an approximation result formulated in theorem 5.1. This guarantees that the KdV equation makes correct predictions about the dynamics of the spatially periodic Boussinesq model (1.9) over the natural KdV timescale. Our result is the first justification of the KdV approximation for a dispersive nonlinear partial differential equation posed in a spatially periodic medium of non-small contrast. For linear systems this limit has been considered independently in [11,12].

In order to make the residual small, an improved approximation must be constructed. Since this construction is not the main purpose of this paper, we additionally assume the following.

(SYM) the coefficient functions

$$a = a(x), \quad b = b(x) \quad \text{and} \quad c = c(x) \quad \text{are even w.r.t. } x.$$

As in the spatially homogeneous situation, it turns out that the proof given for the KdV approximation transfers more or less line for line into proofs for the justification of the approximation via the inviscid Burgers equation and the justification of the Whitham system. The associated approximation results are formulated in theorems 5.2 and 5.3.

This paper was originally intended as the next step in generalizing a method developed in [9] for the justification of the KdV approximation in situations when

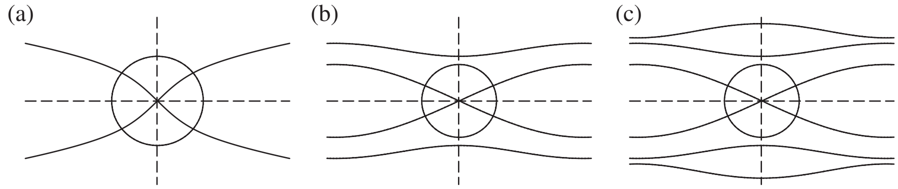


Figure 1. (a) Curves of eigenvalues over the Fourier wavenumbers as they appear for the water wave problem [1, 10, 13, 17, 23, 24]. (b) Finitely many curves of eigenvalues as they appear, for instance, for the poly-atomic FPU system [6, 9]. (c) Infinitely many curves of eigenvalues over the Bloch wavenumbers as they appear for the spatially periodic Boussinesq model (1.9), the water wave problem over a periodic bottom topography or for the linearization around a periodic wave in dispersive systems. Since the Fourier transform of $\varepsilon^2 A(\varepsilon x)$ is given by $\varepsilon^2 \varepsilon^{-1} \hat{A}(x/\varepsilon)$ the KdV equations describe the modes at the wavenumbers $k = 0$ with the vanishing eigenvalues that are contained in the circles. One of the two curves in the circle describes wave packets moving to the left, and the other the wave packets moving to the right.

the KdV modes are resonant with other long wave modes. The method had already been applied successfully in justifying the KdV approximation for the poly-atomic Fermi–Pasta–Ulam (FPU) problem in [6]. The qualitative difference in justifying the KdV equation for the spatially periodic Boussinesq equation in contrast to [9] and [6] is that, for fixed Bloch and Fourier wavenumbers, respectively, the problem presented is infinite dimensional. The cases in [6, 9] correspond to figure 1(b), whereas the spatially periodic Boussinesq equation corresponds to figure 1(c). As a consequence, the normal form transform that is a major part of the proofs in [6, 9] would be more demanding from an analytic point of view. In the justification of the Whitham system with the approach in [6,9], infinitely many normal form transforms have to be performed [15].

Interestingly, for the spatially periodic Boussinesq equation (1.9) there exists an energy in physical space that allows us to incorporate the normal form transforms into the energy estimates. This energy approach is presented below.

*Notation.* Constants that can be chosen independently of the small perturbation parameter $0 < \varepsilon \ll 1$ are denoted by the same symbol, $C$. We write $\int$ for $\int_{-\infty}^{\infty}$. The Fourier transform of a function $u$ is denoted by $\hat{u}$. The Bloch transform of a function $u$ is denoted by $\tilde{u}$, and this tool is recalled in Appendix C. We introduce the norm $\|\cdot\|_{L^2_s}$ by

$$\|\hat{u}\|_{L^2_s}^2 = \int |\hat{u}(k)|^2 (1 + k^2)^s \, \mathrm{d}k,$$

and define the Sobolev norm as $\|u\|_{H^s} = \|\hat{u}\|_{L^2_s}$, but use also equivalent versions.

## 2. The spatially homogeneous case

In this section we give a simple proof for theorems 1.1, 1.2 and 1.3 using the energy method. This proof will be the basis of the subsequent analysis. All three cases can be handled with the same approach.

The residual

$$\mathrm{Res}(u) = -\partial_t^2 u(x,t) + \partial_x^2 u(x,t) - \partial_x^4 u(x,t) + \partial_x^2 (u(x,t)^2)$$

quantifies how much a function $u$ fails to satisfy the Boussinesq model (1.1). For the KdV approximation (1.2), abbreviated as $\varepsilon^2\Psi$, if we choose $A$ to satisfy the KdV equation (1.3), we find

$$\begin{aligned}\mathrm{Res}(\varepsilon^2\Psi) &= -\varepsilon^4 c^2 \partial_X^2 A - 2\varepsilon^6 \partial_T \partial_X A - \varepsilon^8 \partial_T^2 A \\ &\quad + \varepsilon^4 \partial_X^2 A - \varepsilon^6 \partial_X^4 A + \varepsilon^6 \partial_X^2 (A^2) \\ &= -\varepsilon^8 \partial_T^2 A.\end{aligned}$$

Therefore, we have the following.

LEMMA 2.1. *Let $s \geqslant 0$ and let $A \in C([0,T_0], H^{5+s})$ be a solution of the KdV equation (1.3). Then there exist $\varepsilon_0 > 0$ and $C_{\mathrm{res}}$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon^3]} \|\partial_x^{s-1} \mathrm{Res}(\varepsilon^2\Psi(\cdot,t,\varepsilon))\|_{L^2} \leqslant C_{\mathrm{res}} \varepsilon^{(13+2s)/2}.$$

*Proof.* Using the KdV equation allows us to write

$$\begin{aligned}4\partial_T^2 A &= -2\partial_T(\partial_X^3 A + \partial_X(A^2)) \\ &= -2(\partial_X^3 \partial_T A + 2\partial_X(A\partial_T A)) \\ &= \partial_X^3(\partial_X^3 A + \partial_X(A^2)) + 2\partial_X(A(\partial_X^3 A + \partial_X(A^2))).\end{aligned}$$

This shows that $A(\cdot,T) \in H^6$ is necessary to estimate the residual in $L^2$. The formal error of order $\mathcal{O}(\varepsilon^8)$ is reduced by a factor $\varepsilon^{-1/2}$ due to the scaling properties of the $L^2$-norm. Moreover, due to the representation of $\partial_T^2 A$ as a spatial derivative, below, we can apply $\partial_x^{-1} = \varepsilon^{-1}\partial_X^{-1}$ to the residual terms, which, however, lose another factor $\varepsilon^{-1}$. ☐

Similarly, for the Whitham approximation (1.6), abbreviated to $\varepsilon^2\Psi$, we find $\mathrm{Res}(\Psi) = -\varepsilon^4 \partial_X^4 U$ if we choose $U$ to satisfy the Whitham system (1.7). Hence, for an estimate in $L^2$ we need $U \in H^4$. Exactly as above, we have the following.

LEMMA 2.2. *Let $s \geqslant 0$ and let $A \in C([0,T_0], H^{3+s})$ be a solution of the Whitham system (1.7). Then there exist $\varepsilon_0 > 0$ and $C_{\mathrm{res}}$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon]} \|\partial_x^{s-1} \mathrm{Res}(\Psi(\cdot,t,\varepsilon))\|_{L^2} \leqslant C_{\mathrm{res}} \varepsilon^{(5+2s)/2}.$$

REMARK 2.3. For the inviscid Burgers equation, the residual becomes too large with the simple ansatz (1.2). However, by adding higher-order terms to the approximation (1.2) (with a slight abuse of notation this approximation is again called $\varepsilon^\alpha\Psi$), one can always achieve

$$\sup_{t\in[0,T_0/\varepsilon^{1+\alpha}]} \|\mathrm{Res}(\varepsilon^\alpha\Psi(\cdot,t,\varepsilon))\|_{L^2} \leqslant C_{\mathrm{res}} \varepsilon^{(7+4\alpha)/2}$$

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

and

$$\sup_{t\in[0,T_0/\varepsilon^{1+\alpha}]} \|\partial_x^{-1}\operatorname{Res}(\varepsilon^\alpha\Psi(\cdot,t,\varepsilon))\|_{L^2} \leqslant C_{\mathrm{res}}\varepsilon^{(5+4\alpha)/2}.$$

(See Appendix A, where we prove these estimates for $\alpha = 1$ and explain that the number of additional terms goes to infinity for $\alpha \to 0$ and $\alpha \to 2$.)

From this point on, the remaining estimates can be handled in exactly the same way. The $\alpha = 0$ case corresponds to the Whitham approximation, and the $\alpha = 2$ case to the KdV approximation. The difference $\varepsilon^{(3+2\alpha)/2}R = u - \varepsilon^\alpha\Psi$ satisfies

$$\partial_t^2 R = \partial_x^2 R - \partial_x^4 R + 2\varepsilon^\alpha\partial_x^2(\Psi R) + \varepsilon^{(3+2\alpha)/2}\partial_x^2(R^2) + \varepsilon^{-(3+2\alpha)/2}\operatorname{Res}(\varepsilon^2\Psi). \quad (2.1)$$

We multiply the error equation (2.1) by $-\partial_t\partial_x^{-2}R$, which is defined via its Fourier transform w.r.t. $x$, namely via

$$\widehat{\partial_x^{-1}R}(k) = \frac{1}{\mathrm{i}k}\hat{R}(k),$$

integrate it w.r.t. $x$, and find

$$\begin{aligned}
-\int(\partial_t\partial_x^{-2}R)\partial_t^2 R\,\mathrm{d}x &= \tfrac{1}{2}\partial_t\int(\partial_t\partial_x^{-1}R)^2\,\mathrm{d}x,\\
-\int(\partial_t\partial_x^{-2}R)\partial_x^2 R\,\mathrm{d}x &= -\tfrac{1}{2}\partial_t\int R^2\,\mathrm{d}x,\\
\int(\partial_t\partial_x^{-2}R)\partial_x^4 R\,\mathrm{d}x &= -\tfrac{1}{2}\partial_t\int(\partial_x R)^2\,\mathrm{d}x,\\
-\int(\partial_t\partial_x^{-2}R)\partial_x^2(\Psi R)\,\mathrm{d}x &= -\int(\partial_t R)\Psi R\,\mathrm{d}x\\
&= -\tfrac{1}{2}\partial_t\int\Psi R^2\,\mathrm{d}x + \varepsilon\int(\partial_\tau\Psi)R^2\,\mathrm{d}x,\\
-\int(\partial_t\partial_x^{-2}R)\partial_x^2(R^2)\,\mathrm{d}x &= -\int(\partial_t R)R^2\,\mathrm{d}x = -\tfrac{1}{3}\partial_t\int R^3\,\mathrm{d}x,\\
-\int(\partial_t\partial_x^{-2}R)\operatorname{Res}(\varepsilon^2\Psi)\,\mathrm{d}x &= \int(\partial_t\partial_x^{-1}R)\partial_x^{-1}\operatorname{Res}(\varepsilon^2\Psi)\,\mathrm{d}x.
\end{aligned}$$

We can estimate

$$\begin{aligned}
\left|\int(\partial_t\partial_x^{-1}R)\partial_x^{-1}\operatorname{Res}(\varepsilon^2\Psi)\,\mathrm{d}x\right| &\leqslant \|\partial_t\partial_x^{-1}R\|_{L^2}\|\partial_x^{-1}\operatorname{Res}(\varepsilon^2\Psi)\|_{L^2},\\
\left|\int(\partial_\tau\Psi)R^2\,\mathrm{d}x\right| &\leqslant \|\partial_\tau\Psi\|_{L^\infty}\|R\|_{L^2}^2.
\end{aligned}$$

For the energy

$$E = \int(\partial_t\partial_x^{-1}R)^2 + R^2 + (\partial_x R)^2 + 2\varepsilon^\alpha\Psi R^2 + \tfrac{2}{3}\varepsilon^{(3+2\alpha)/2}R^3\,\mathrm{d}x,$$

the following hold: in the $\alpha > 0$ case we have that for all $M > 0$ there exist $C_1$ and $\varepsilon_1 > 0$ such that for all $\varepsilon \in (0, \varepsilon_1)$ we have

$$\|R\|_{H^1} \leqslant C_1 E^{1/2}$$

as long as $E \leqslant M$; in the $\alpha = 0$ case the energy $E$ is an upper bound for the squared $H^1$-norm for $\|\Psi\|_{L^\infty}$ sufficiently small but independent of $0 < \varepsilon \ll 1$. Therefore, $E$ satisfies the inequality

$$\begin{aligned}\frac{\mathrm{d}E}{\mathrm{d}t} &\leqslant C\varepsilon^{1+\alpha}E + C\varepsilon^{(3+2\alpha)/2}E^{3/2} + C\varepsilon^{1+\alpha}E^{1/2} \\ &\leqslant 2C\varepsilon^{1+\alpha}E + C\varepsilon^{(3+2\alpha)/2}E^{3/2} + C\varepsilon^{1+\alpha}, \end{aligned} \tag{2.2}$$

with a constant $C$ independent of $\varepsilon \in (0, \varepsilon_1)$. Under the assumption $C\varepsilon^{1/2}E^{1/2} \leqslant 1$ we obtain

$$\frac{\mathrm{d}E}{\mathrm{d}t} \leqslant (2C+1)\varepsilon^{1+\alpha}E + C\varepsilon^{1+\alpha}.$$

Gronwall's inequality immediately gives the bound

$$\sup_{t\in[0,T_0/\varepsilon^{1+\alpha}]} E(t) = CT_0\mathrm{e}^{(2C+1)T_0} =: M = \mathcal{O}(1).$$

Finally, choosing $\varepsilon_2 > 0$ so small that $C\varepsilon_2^{1/2}M^{1/2} \leqslant 1$ gives the required estimate for all $\varepsilon \in (0, \varepsilon_0)$ with $\varepsilon_0 = \min(\varepsilon_1, \varepsilon_2) > 0$ in all three cases.

Remark 2.4. The Boussinesq model (1.1) is a semilinear dispersive system, and so we have local existence and uniqueness of solutions. The variation-of-constant formula associated with the first-order system for the variables $u$ and $\partial_t(\partial_x^4 - \partial_x^2)^{-1/2}u$ is a contraction in the space $C([-T_*, T_*], H^\theta \times H^\theta)$ for every $\theta > \frac{1}{2}$ if $T_* > 0$ is sufficiently small. The local existence and uniqueness of solutions combined with the previous estimates, for instance, yield the existence and uniqueness of solutions for all $t \in [0, T_0/\varepsilon^3]$ in the KdV case and all $t \in [0, T_0/\varepsilon]$ in the Whitham case.

## 3. Derivation of the amplitude equations

In this section we return to the spatially periodic situation. The derivation of the amplitude equations is less obvious than in the spatially homogeneous case. In order to derive the amplitude equations, we expand (1.9) into the eigenfunctions of the linear problem. As in [2], after this expansion we are back in the spatially homogeneous set-up, except that the Fourier transform has been replaced by the Bloch transform.

### 3.1. Spectral properties

The linearized problem

$$\partial_t^2 u(x,t) = \partial_x(a(x)\partial_x u(x,t)) - \partial_x^2(b(x)\partial_x^2 u(x,t)) \tag{3.1}$$

is solved by so-called Bloch modes

$$u(x,t) = w(x)\mathrm{e}^{\mathrm{i}lx}\mathrm{e}^{\mathrm{i}\omega t},$$

with $w$ being $2\pi$-periodic w.r.t. $x$ satisfying

$$-(\partial_x + \mathrm{i}l)(a(x)(\partial_x + \mathrm{i}l)w(x)) + (\partial_x + \mathrm{i}l)^2(b(x)(\partial_x + \mathrm{i}l)^2 w(x)) = \omega^2 w(x).$$

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

The left-hand side defines a self-adjoint elliptic operator $L_l(\partial_x)\colon H^{\theta+4} \to H^{\theta}$. Hence, for fixed $l$, there exists a countable set of eigenvalues $\lambda_n(l)$, with $n \in \mathbb{N}$, ordered such that $\lambda_{n+1}(l) \geqslant \lambda_n(l)$, with associated eigenfunctions $w_n(x,l)$.

LEMMA 3.1. *For $l = 0$, the operator $L_0(\partial_x)$ possesses the simple eigenvalue $\lambda_1(0) = 0$ associated with the eigenfunction $\tilde{w}_1(0,x) = 1$.*

*Proof.* Obviously, we have $L_0(\partial_x)1 = 0$. Moreover, we have

$$(w, L_0(\partial_x)w)_{L^2} = \int_{-1/2}^{1/2} a(x)(\partial_x w(x))^2 \, \mathrm{d}x + \int_{-1/2}^{1/2} b(x)(\partial_x^2 w(x))^2 \, \mathrm{d}x \geqslant 0.$$

Hence, $L_0(\partial_x)w = 0$ implies $\partial_x w = 0$. From the $2\pi$-periodicity it follows that $w = \text{const}$. Hence, $\lambda_1(0) = 0$ is a simple eigenvalue. □

It is well known that the curves $l \mapsto \lambda_n(l)$ and $l \mapsto \tilde{w}_n(l,\cdot)$ are smooth w.r.t. $l$ for simple eigenvalues. Hence, there exists a $\delta_0 > 0$ such that for $l \in [-\delta_0, \delta_0]$ the smallest eigenvalue $\lambda_1(l)$ is separated from the rest of the spectrum. Since $L_l(\partial_x)$ is self-adjoint and positive definite for all $l$ we have $\lambda_1(l) \geqslant 0$ for all $l$. In the KdV equation only odd spatial derivatives occur, and in the Whitham system only even spatial derivatives occur. This is a consequence of the following lemma.

LEMMA 3.2. *The curve $l \mapsto \lambda_1(l)$ for $l \in [-\delta_0, \delta_0]$ is an even real-valued function. The associated eigenfunctions satisfy $\tilde{w}_1(l,x) = \overline{\tilde{w}_1(-l,x)}$. Under the assumption that the coefficient functions $a$ and $b$ are even, the eigenfunctions possess an expansion*

$$\tilde{w}_1(l,x) = \sum_{j=0}^{\infty} (\mathrm{i}l)^j g_j(x),$$

*with $g_0(x) = 1$, $\int_0^{2\pi} g_j(x) \, \mathrm{d}x = 0$ for $j \geqslant 1$,*

$$g_{2j}(x) = g_{2j}(-x) \in \mathbb{R} \quad \text{and} \quad g_{2j+1}(x) = -g_{2j+1}(-x) \in \mathbb{R}.$$

*Proof.* The first two statements follow from the fact that for fixed $l$ the operator $L_l(\partial_x)$ is self-adjoint and from the fact that (1.9) is a real problem. For $(\mathrm{i}l)^0$ we obtain

$$-\partial_x(a(x)\partial_x g_0(x)) + \partial_x^2(b(x)\partial_x^2 g_0(x)) = 0,$$

which is, as we already know, solved uniquely by $g_0(x) = 1$. For $(\mathrm{i}l)^1$ we obtain

$$-\partial_x(a(x)\partial_x g_1(x)) + \partial_x^2(b(x)\partial_x^2 g_1(x)) - \partial_x a(x) = 0.$$

The term $\partial_x a(x)$ is odd. The subspace of odd functions is invariant for the differential operator $L_0(\partial_x) = -\partial_x(a(x)(\partial_x \cdot)) + \partial_x^2(b(x)\partial_x^2 \cdot)$. Moreover, in this subspace its spectrum is bounded away from zero such that this equation possesses a unique odd solution, $g_1 = g_1(x)$. For $(\mathrm{i}l)^2$ we obtain

$$-\partial_x(a(x)\partial_x g_2(x)) + \partial_x^2(b(x)\partial_x^2 g_2(x)) + 1 + f_2(x) = 1$$

with $f_2(x)$ an even function depending on $a$, $b$, $g_0$ and $g_1$ and possessing vanishing mean value. In the subspace of vanishing mean value the differential operator $L_0(\partial_x)$

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

possesses a spectrum bounded away from zero such that this equation possesses a unique even solution $g_2 = g_2(x)$. With the same arguments the next orders with the stated properties can be computed. The convergence of the series in a neighbourhood of $l = 0$ in $H^\theta$ for every $\theta \geqslant 0$ follows from the smoothness of the curve of simple eigenfunctions w.r.t. $l$ and the smoothness of the coefficient functions $a$, $b$ and $c$ w.r.t. $x$. □

The KdV equation, the inviscid Burgers equation and the Whitham system describe the modes associated with the curve $\lambda_1$ close to $l = 0$. Therefore, in order to derive these amplitude equations we consider the Bloch transform

$$u(x,t) = \int_{-1/2}^{1/2} \tilde{u}(l,x,t)\mathrm{e}^{\mathrm{i}lx}\,\mathrm{d}x$$

of (1.9), namely

$$\partial_t^2 \tilde{u}(l,x,t) = -L_l(\partial_x)\tilde{u}(l,x,t) + N_l(\partial_x)(\tilde{u})(l,x,t), \tag{3.2}$$

where

$$N_l(\partial_x)(\tilde{u})(l,x,t) = (\partial_x + \mathrm{i}l)\left(c(x)(\partial_x + \mathrm{i}l)\int_{-1/2}^{1/2} \tilde{u}(l-m,x,t)\tilde{u}(m,x,t)\,\mathrm{d}m\right).$$

Then we make the ansatz

$$\tilde{u}(l,x,t) = \chi_{[-\delta_0/2,\delta_0/2]}(l)\tilde{u}_1(l,t)\tilde{w}_1(l,x) + \tilde{v}(l,x,t)$$

with

$$\int_0^{2\pi} \overline{\tilde{w}_1(l,x)}\tilde{v}(l,x,t)\,\mathrm{d}x = 0$$

for $l \in [-\frac{1}{2}\delta_0, \frac{1}{2}\delta_0]$ and find

$$\begin{aligned}\partial_t^2 \tilde{u}_1(l,t) &= -\lambda_1(l)\tilde{u}_1(l,t) + P_c(l)N_l(\partial_x)(\tilde{u})(l,t),\\ \partial_t^2 \tilde{v}(l,x,t) &= -L_l(\partial_x)\tilde{v}(l,x,t) + P_s(l)N_l(\partial_x)(\tilde{u})(l,x,t),\end{aligned}$$

where

$$\begin{aligned}(P_c\tilde{u})(l,t) &= \frac{1}{2\pi}\chi_{[-\delta_0/2,\delta_0/2]}(l)\int_0^{2\pi} \overline{\tilde{w}_1(l,x)}\tilde{u}(l,x,t)\,\mathrm{d}x,\\ (P_s\tilde{u})(l,x,t) &= \tilde{u}(l,x,t) - (P_c\tilde{u})(l,t)w_1(l,x).\end{aligned}$$

All amplitude equations we have in mind can be derived in a very similar way. They describe the evolution of the $\tilde{u}_1$ modes that are concentrated in an $\mathcal{O}(\varepsilon)$ neighbourhood of the Bloch wavenumber $l = 0$. In all three cases we make an ansatz

$$\tilde{u}_1(l,t) = \varepsilon^{-1}\varepsilon^{\alpha}\chi_{[-\delta_0/4,\delta_0/4]}\left(\frac{l}{\varepsilon}\right)\hat{A}\left(\frac{l}{\varepsilon},\varepsilon^{1+\alpha}t\right)\mathrm{e}^{\mathrm{i}lct} \tag{3.3}$$

with $\alpha = 2$ and $c > 0$ for the KdV approximation, $\alpha \in (0,2)$ and $c > 0$ for the inviscid Burgers approximation and $\alpha = 0$ and $c = 0$ for the Whitham approximation (see the caption of figure 1). The amplitude $\hat{A}$ will be defined in Fourier space, and the cut-off function $\chi_{[-\delta_0/4,\delta_0/4]}(l/\varepsilon)$ allows us to transfer $\hat{A}$ into Bloch space. In the following we use the abbreviation

$$\tilde{A}\left(\frac{l}{\varepsilon}, \varepsilon^{1+\alpha} t\right) = \chi_{[-\delta_0/4,\delta_0/4]}\left(\frac{l}{\varepsilon}\right) \hat{A}\left(\frac{l}{\varepsilon}, \varepsilon^{1+\alpha} t\right). \tag{3.4}$$

For each of the three approximations we have to derive the associated amplitude equation and to compute and estimate the residual terms

$$\mathrm{Res}(\tilde{u})(l,x,t) = -\partial_t^2 \tilde{u}(l,x,t) - L_l(\partial_x)\tilde{u}(l,x,t) + N_l(\partial_x)(\tilde{u})(l,x,t).$$

### 3.2. Derivation of the KdV and the inviscid Burgers equations

The amplitude equations we have in mind have derivatives in front of the nonlinear terms. Hence, before deriving these equations we need to prove a number of properties about the nonlinear terms. We introduce the kernels $s_{11}^1(l, l-m, m), \ldots, s_{vv}^v(l, l-m, m)$ by

$$\begin{aligned}
(P_c N_l(\partial_x)(\tilde{u}))(l,t) &= \int_{-1/2}^{1/2} s_{11}^1(l, l-m, m)\tilde{u}_1(l-m,t)\tilde{u}_1(m,t)\,\mathrm{d}m \\
&\quad + \int_{-1/2}^{1/2} s_{1v}^1(l, l-m, m)\tilde{u}_1(l-m,t)\tilde{v}(m,x,t)\,\mathrm{d}m \\
&\quad + \int_{-1/2}^{1/2} s_{v1}^1(l, l-m, m)\tilde{v}(l-m,x,t)\tilde{u}_1(m,t)\,\mathrm{d}m \\
&\quad + \int_{-1/2}^{1/2} s_{vv}^1(l, l-m, m)\tilde{v}(l-m,x,t)\tilde{v}(m,x,t)\,\mathrm{d}m
\end{aligned}$$

and

$$\begin{aligned}
(P_s N_l(\partial_x)(\tilde{u}))(l,x,t) &= \int_{-1/2}^{1/2} s_{11}^v(l, l-m, m)\tilde{u}_1(l-m,t)\tilde{u}_1(m,t)\,\mathrm{d}m \\
&\quad + \int_{-1/2}^{1/2} s_{1v}^v(l, l-m, m)\tilde{u}_1(l-m,t)\tilde{v}(m,x,t)\,\mathrm{d}m \\
&\quad + \int_{-1/2}^{1/2} s_{v1}^v(l, l-m, m)\tilde{v}(l-m,x,t)\tilde{u}_1(m,t)\,\mathrm{d}m \\
&\quad + \int_{-1/2}^{1/2} s_{vv}^v(l, l-m, m)\tilde{v}(l-m,x,t)\tilde{v}(m,x,t)\,\mathrm{d}m.
\end{aligned}$$

For the derivation of the KdV and Burgers equations we need the following.

LEMMA 3.3. *We have*

$$|s_{11}^1(l, l-m, m) - \nu_2 l^2| \leqslant C|l|(l^2 + (l-m)^2 + m^2),$$

*where*

$$\nu_2 = -\frac{1}{2\pi}\int_0^{2\pi} c(x)(1+\partial_x g_1(x))^2\,\mathrm{d}x. \tag{3.5}$$

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

*Proof.* Due to lemma 3.2 we have

$$\tilde{w}_1(l,x) = 1 + \mathrm{i}lg_1(x) + \mathcal{O}(l^2), \tag{3.6}$$

where $g_1(x) \in \mathbb{R}$ with $\int_0^{2\pi} g_1(x)\,\mathrm{d}x = 0$. This expansion yields

$$\begin{aligned}
&2\pi s_{11}^1(l, l-m, m)\\
&\quad = \int_0^{2\pi} \overline{\tilde{w}_1(l,x)}(\partial_x + \mathrm{i}l)(c(x)(\partial_x + \mathrm{i}l)(\tilde{w}_1(l-m,x)\tilde{w}_1(m,x)))\,\mathrm{d}x\\
&\quad = \int_0^{2\pi} (1 - \mathrm{i}lg_1(x) + \mathcal{O}(l^2))(\partial_x + \mathrm{i}l)\\
&\qquad\qquad \times (c(x)(\partial_x + \mathrm{i}l)((1 + \mathrm{i}(l-m)g_1(x) + \mathcal{O}((l-m)^2))\\
&\qquad\qquad\qquad \times (1 + \mathrm{i}mg_1(x) + \mathcal{O}(m^2))))\,\mathrm{d}x\\
&\quad = -\int_0^{2\pi} c(x)((\partial_x - \mathrm{i}l)(1 - \mathrm{i}lg_1(x) + \mathcal{O}(l^2)))\\
&\qquad\qquad \times ((\partial_x + \mathrm{i}l)((1 + \mathrm{i}(l-m)g_1(x) + \mathcal{O}((l-m)^2))\\
&\qquad\qquad\qquad \times (1 + \mathrm{i}mg_1(x) + \mathcal{O}(m^2))))\,\mathrm{d}x\\
&\quad = -\int_0^{2\pi} c(x)(-\mathrm{i}l - \mathrm{i}l\partial_x g_1(x) + \mathcal{O}(l^2))\\
&\qquad\qquad \times ((\partial_x + \mathrm{i}l)((1 + \mathrm{i}lg_1(x) + \mathcal{O}((l-m)^2 + m^2)))\,\mathrm{d}x\\
&\quad = -\int_0^{2\pi} c(x)(-\mathrm{i}l - \mathrm{i}l\partial_x g_1(x) + \mathcal{O}(l^2))\\
&\qquad\qquad \times (\mathrm{i}l + \mathrm{i}l\partial_x g_1(x) + \mathcal{O}((l-m)^2 + m^2)))\,\mathrm{d}x\\
&\quad = \nu_2 l^2 + \mathcal{O}(|l|(l^2 + (l-m)^2 + m^2)).
\end{aligned}$$

We remark at this point that, due to the fact that $a$, $b$ and $c$ are assumed to be even, we have for reasons of symmetry that the higher-order terms are not only $\mathcal{O}(|l|(l^2 + (l-m)^2 + m^2))$ but also $\mathcal{O}(l^4 + (l-m)^4 + m^4)$ (see below). □

The following derivation of amplitude equations in Fourier or Bloch space is straightforward and documented in various papers. We refer the reader to [22, ch. 5] for an introduction.

### 3.2.1. *The KdV equation*

We start with the KdV approximation $\varepsilon^2\Psi$, which is defined via (3.3) for $\alpha = 2$ and is inserted into $\mathrm{Res}(\tilde{u})$. We find with $\tilde{u}_1(l,t) = \varepsilon\tilde{A}(K,T)\boldsymbol{E}$, $\boldsymbol{E} = \mathrm{e}^{\mathrm{i}\varepsilon Kct}$, $T = \varepsilon^3 t$ and $l = \varepsilon K$ that

$$\begin{aligned}
P_c(\mathrm{Res}(\tilde{u}))(l,t) = &-\partial_t^2\tilde{u}_1(l,t) - \lambda_1(l)\tilde{u}_1(l,t)\\
&+ \int_{-1/2}^{1/2} s_{11}^1(l, l-m, m)\tilde{u}_1(l-m,t)\tilde{u}_1(m,t)\,\mathrm{d}m
\end{aligned}$$

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

$$
\begin{aligned}
&= \varepsilon^3 c^2 K^2 \tilde{A}(K,T)\boldsymbol{E} - 2\varepsilon^5 \mathrm{i} cK(\partial_T \tilde{A}(K,T))\boldsymbol{E} - \varepsilon^7(\partial_T^2 \tilde{A}(K,T))\boldsymbol{E} \\
&\quad - \tfrac{1}{2}\varepsilon^3 \lambda_1''(0) K^2 \tilde{A}(K,T)\boldsymbol{E} - \tfrac{1}{24}\varepsilon^5 \lambda_1''''(0) K^4 \tilde{A}(K,T)\boldsymbol{E} + \mathcal{O}(\varepsilon^7) \\
&\quad + \varepsilon^5 \int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \nu_2 K^2 \tilde{A}(K-M,T)\tilde{A}(M,T)\,\mathrm{d}M\boldsymbol{E} + \mathcal{O}(\varepsilon^6).
\end{aligned}
$$

If $\hat{A}(\cdot,T) \in L_s^2$, then the error made by replacing $\int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \cdots \,\mathrm{d}M$ by $\int_{-\infty}^{\infty} \cdots \,\mathrm{d}M$ is $\mathcal{O}(\varepsilon^{s-1/2})$. Hence, by equating the coefficients of $\varepsilon^3$ and $\varepsilon^5$ to zero we find $c^2 = \frac{1}{2}\lambda_1''(0)$ and $\hat{A}$ satisfying

$$
-2\mathrm{i}c\partial_T \hat{A}(K,T) - \tfrac{1}{24}\lambda_1''''(0)K^3 \hat{A}(K,T) + \int_{-\infty}^{\infty} \nu_2 K \hat{A}(K-M,T)\hat{A}(M,T)\,\mathrm{d}M = 0,
$$

and $A$ satisfying the KdV equation

$$
2c\partial_T A(X,T) + \tfrac{1}{24}\lambda_1''''(0)\partial_X^3 A(X,T) + \nu_2 \partial_X (A(X,T)^2) = 0, \tag{3.7}
$$

respectively.

### 3.2.2. *The inviscid Burgers equation*

Due to the explanations given in Appendix A we restrict our analysis to the $\alpha = 1$ case. We insert the inviscid Burgers approximation $\varepsilon^\alpha \Psi$, which is defined via (3.3) for $\alpha = 1$, into $\mathrm{Res}(\tilde{u})$. We find with $\tilde{u}_1(l,t) = \tilde{A}(K,T)\boldsymbol{E}$, $\boldsymbol{E} = \mathrm{e}^{\mathrm{i}\varepsilon Kct}$, $T = \varepsilon^2 t$ and $l = \varepsilon K$ that

$$
\begin{aligned}
P_c(\mathrm{Res}(\tilde{u}))(l,t) &= -\partial_t^2 \tilde{u}_1(l,t) - \lambda_1(l)\tilde{u}_1(l,t) \\
&\quad + \int_{-1/2}^{1/2} s_{11}^1(l,l-m,m)\tilde{u}_1(l-m,t)\tilde{u}_1(m,t)\,\mathrm{d}m \\
&= \varepsilon^2 c^2 K^2 \tilde{A}(K,T)\boldsymbol{E} - 2\varepsilon^3 \mathrm{i}cK(\partial_T \tilde{A}(K,T))\boldsymbol{E} - \varepsilon^4(\partial_T^2 \tilde{A}(K,T))\boldsymbol{E} \\
&\quad - \tfrac{1}{2}\varepsilon^2 \lambda_1''(0)K^2 \tilde{A}(K,T)\boldsymbol{E} - \tfrac{1}{24}\varepsilon^4 \lambda_1''''(0)K^4 \tilde{A}(K,T)\boldsymbol{E} + \mathcal{O}(\varepsilon^4) \\
&\quad + \varepsilon^3 \int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \nu_2 K^2 \tilde{A}(K-M,T)\tilde{A}(M,T)\,\mathrm{d}M\boldsymbol{E} + \mathcal{O}(\varepsilon^4).
\end{aligned}
$$

We proceed as above and equate the coefficients of $\varepsilon^2$ and $\varepsilon^3$ to zero. We find $c^2 = \frac{1}{2}\lambda_1''(0)$ and $\hat{A}$ satisfying

$$
-2\mathrm{i}c\partial_T \hat{A}(K,T) + \int_{-\infty}^{\infty} \nu_2 K \hat{A}(K-M,T)\hat{A}(M,T)\,\mathrm{d}M = 0
$$

and $A$ satisfying the inviscid Burgers equation

$$
2c\partial_T A(X,T) + \nu_2 \partial_X (A(X,T)^2) = 0, \tag{3.8}
$$

respectively.

## 3.3. Derivation of the Whitham system

The derivation of the Whitham system is much more involved, since it must include the $\tilde{v}$ part. Due to the symmetry assumption (SYM) with $u = u(x,t)$,

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

$u = u(-x,t)$ is also a solution of (1.9). As a consequence, in (1.9) all terms must contain an even number of $\partial_x$-derivatives. Since in Bloch space

$$\begin{aligned} u(-x,t) &= \int_{-1/2}^{1/2} \tilde{u}(-x,l)\mathrm{e}^{-\mathrm{i}lx}\,\mathrm{d}l \\ &= -\int_{1/2}^{-1/2} \tilde{u}(-x,-l)\mathrm{e}^{\mathrm{i}lx}\,\mathrm{d}l \\ &= \int_{-1/2}^{1/2} \tilde{u}(-x,-l)\mathrm{e}^{\mathrm{i}lx}\,\mathrm{d}l \end{aligned}$$

with $\tilde{u} = \tilde{u}(l,x,t)$, $\tilde{u} = \tilde{u}(-l,-x,t)$ is also a solution of the Bloch-wave-transformed system (3.2). As a consequence, in (3.2) all terms must contain an even number of $\partial_x$-derivatives or $\mathrm{i}l$, $\mathrm{i}(l-m)$, or $\mathrm{i}m$ factors, i.e. for instance $\mathrm{i}l\partial_x$ can occur, but $-l^2\partial_x$ cannot. Before we start the derivation of the Whitham system, we additionally need that at least one $l$ factor occurs in some of the kernel functions $s^j_{j_1 j_2}$.

LEMMA 3.4. *We have*

(a) $|s^1_{vv}(l,l-m,m)| \leqslant C|l|$,

(b) $|s^v_{11}(l,l-m,m)| \leqslant C(|l| + (l-m)^2 + m^2)$.

*Proof.*
(a) Using the expansion (3.6) yields, after some integration by parts, that

$$\begin{aligned} &\int_0^{2\pi} \overline{\tilde{w}_1(l,x)}(\partial_x + \mathrm{i}l)\left(c(x)(\partial_x + \mathrm{i}l)\int_{-1/2}^{1/2} \tilde{v}(l-m,x,t)\tilde{v}(m,x,t)\,\mathrm{d}m\right)\mathrm{d}x \\ &= \int_{-1/2}^{1/2}\int_0^{2\pi} c(x)(-\mathrm{i}l + \mathrm{i}l\partial_x g_1(x) + \mathcal{O}(l^2))(\partial_x + \mathrm{i}l)(\tilde{v}(l-m,x,t)\tilde{v}(m,x,t))\,\mathrm{d}x\,\mathrm{d}m \\ &= \mathcal{O}(l). \end{aligned}$$

(b) As above we obtain

$$\begin{aligned} &s^v_{11}(l,l-m,m) \\ &\quad= \int_0^{2\pi} \overline{\tilde{v}(l,x)}(\partial_x + \mathrm{i}l)(c(x)(\partial_x + \mathrm{i}l)(\tilde{w}_1(l-m,x)\tilde{w}_1(m,x)))\,\mathrm{d}x \\ &\quad= \int_0^{2\pi} \overline{\tilde{v}(l,x)}(\partial_x + \mathrm{i}l) \\ &\qquad\qquad \times (c(x)(\partial_x + \mathrm{i}l)((1 + \mathrm{i}(l-m)g_1(x) + \mathcal{O}((l-m)^2)) \\ &\qquad\qquad\qquad\qquad \times (1 + \mathrm{i}mg_1(x) + \mathcal{O}(m^2))))\,\mathrm{d}x \\ &\quad= \int_0^{2\pi} \overline{\tilde{v}(l,x)}(\partial_x + \mathrm{i}l)(c(x)(\mathrm{i}l + \mathrm{i}l\partial_x g_1(x) + \mathcal{O}((l-m)^2 + m^2)))\,\mathrm{d}x \\ &\quad= \mathcal{O}(|l| + (l-m)^2 + m^2). \end{aligned}$$

□

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

For the derivation of the Whitham system we make the following ansatz:

$$\tilde{u}_1(l,t) = \varepsilon^{-1}\tilde{A}(K,T) \quad \text{and} \quad \tilde{v}(l,x,t) = \tilde{B}(K,x,T). \tag{3.9}$$

where $T = \varepsilon t$ and $l = \varepsilon K$. With $\tilde{u}(l,x,t) = \tilde{u}_1(l,t)\tilde{w}_1(l,x) + \tilde{v}(l,x,t)$ we find that

$$\begin{aligned} P_c(\mathrm{Res}(\tilde{u}))(l,t) &= -\partial_t^2\tilde{u}_1(l,t) - \lambda_1(l)\tilde{u}_1(l,t) + P_c(l)N_l(\partial_x)(\tilde{u})(l,t) \\ &= -\varepsilon\partial_T^2\tilde{A}(K,T) - \tfrac{1}{2}\varepsilon\lambda_1''(0)K^2\tilde{A}(K,T) + \mathcal{O}(\varepsilon^3) \\ &\quad + P_c(\varepsilon K)N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K, T/\varepsilon) \end{aligned}$$

and

$$\begin{aligned} P_s(\mathrm{Res}(\tilde{u}))(l,x,t) &= -\partial_t^2\tilde{v}(l,x,t) - L_l(\partial_x)\tilde{v}(l,x,t) + P_s(l)N_l(\partial_x)(\tilde{u})(l,x,t) \\ &= -\varepsilon^2\partial_T^2\tilde{B}(K,x,T) - \tilde{L}_{\varepsilon K}(\partial_x)\tilde{B}(K,x,T) \\ &\quad + P_s(l)N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K, x, T/\varepsilon). \end{aligned}$$

Since $P_s(\varepsilon K)N_{\varepsilon K}(\partial_x)(\tilde{u})$ is quadratic w.r.t. $\tilde{u}$ and since $\tilde{L}_{\varepsilon K}$ is invertible on the range of $P_s(\varepsilon K)$, we can use the implicit function theorem to solve

$$-\tilde{L}_{\varepsilon K}(\partial_x)\tilde{B}(K,x,T) + P_s(\varepsilon K)N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K, x, T/\varepsilon) = 0$$

w.r.t. $\tilde{B} = H(\tilde{A})(K,x,T)$ for sufficiently small $\tilde{A}$. Note that we have kept our notation and wrote $T/\varepsilon$ in the arguments of $N$, although in fact it depends only on $T$. We insert $\tilde{B} = H(\tilde{A})(K,x,T)$ into the first equation and obtain

$$\begin{aligned} P_c(\mathrm{Res}(\tilde{u}))(l,t) &= -\varepsilon\partial_T^2\tilde{A}(K,T) - \tfrac{1}{2}\varepsilon\lambda_1''(0)K^2\tilde{A}(K,T) + \mathcal{O}(\varepsilon^3) \\ &\quad + P_c(\varepsilon K)N_{\varepsilon K}(\partial_x)(\varepsilon^{-1}\tilde{A}(K,T)\tilde{w}_1(\varepsilon K, x) \\ &\quad + H(\tilde{A})(K,x,T))(\varepsilon K, T/\varepsilon). \end{aligned}$$

The Whitham system occurs by expanding the right-hand side w.r.t. $\varepsilon$ and by equating the coefficient in front of $\varepsilon^1$ to zero. We obtain in a first step

$$\partial_T^2\tilde{A}(K,T) + \tfrac{1}{2}\lambda_1''(0)K^2\tilde{A}(K,T) + \tilde{G}(\tilde{A})(K,T) = 0,$$

where $\tilde{G}$ is a nonlinear function that can be written as

$$\tilde{G}(\tilde{A})(K,T) = -\chi_{[-\delta_0/4,\delta_0/4]}(\varepsilon K)\sum_{j=2}^{\infty} s_j \mathrm{i}K \int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \tilde{A}^{*(j-1)}(K-M)\mathrm{i}M\tilde{A}(M)\,\mathrm{d}M$$

with coefficients $s_j$. The factor $\mathrm{i}K$ comes from lemmas 3.3 and 3.4(a); the factor $\mathrm{i}M$ comes from the fact that we need an even number of such factors due to the reflection symmetry. Due to the long-wave character of the approximation we have exactly two such factors at $\varepsilon$. Replacing the Bloch transform $\tilde{A}(K,T)$ by the Fourier transform $\tilde{A}(K,T)$ via (3.4) finally gives Whitham's system,

$$\partial_T^2\hat{A}(K,T) + \tfrac{1}{2}\lambda_1''(0)K^2\hat{A}(K,T) + \hat{G}(\hat{A})(K,T) = 0, \tag{3.10}$$

in Fourier space, where $\hat{G}$ is a nonlinear function that can be written as

$$\hat{G}(\hat{A})(K,T) = -\sum_{j=2}^{\infty} s_j \mathrm{i}K \int_{-\infty}^{\infty} \hat{A}^{*(j-1)}(K-M)\mathrm{i}M\hat{A}(M)\,\mathrm{d}M.$$

In physical space we have

$$G(A)(X,T) = -\sum_{j=2}^{\infty} s_j \partial_X (A^{j-1}\partial_X A) = -\partial_X^2 \sum_{j=2}^{\infty} s_j A^j/j$$

such that Whitham's system can finally be written as

$$\partial_T^2 A = \partial_X^2 \mathcal{H}(A) \quad \text{with } \mathcal{H}(A) = -\tfrac{1}{2}\lambda_1''(0)A - \sum_{j=2}^{\infty} s_j A^j/j. \tag{3.11}$$

## 4. Estimates for the residual

After the derivation of the amplitude equations, we estimate the so-called residual, i.e. the terms that do not cancel after inserting the approximation into (1.9). In order to obtain estimates as in the spatially homogeneous case for the residual terms in terms of $\varepsilon$ we have to modify our approximations with higher-order terms.

*The improved KdV approximation.* For the construction of the improved KdV approximation we proceed as for the derivation of the Whitham system. With $\boldsymbol{E} = \mathrm{e}^{\mathrm{i}\varepsilon Kct}$, $T = \varepsilon^3 t$ and $l = \varepsilon K$ we make the ansatz

$$\begin{aligned}\tilde{u}_1(l,t) &= \varepsilon \tilde{A}(K,T)\boldsymbol{E},\\ \tilde{v}(l,x,t) &= \varepsilon^4 \tilde{B}(K,x,T)\boldsymbol{E} + \varepsilon^5 \tilde{B}_2(K,x,T)\boldsymbol{E} + \varepsilon^3 \tilde{B}_3(K,x,T)\boldsymbol{E}.\end{aligned}$$

With $\tilde{u}(l,x,t) = \tilde{u}_1(l,t)\tilde{w}_1(l,x) + \tilde{v}(l,x,t)$, $T = \varepsilon t$ and $l = \varepsilon K$ we find that

$$\begin{aligned}P_c(\mathrm{Res}(\tilde{u}))(l,t) &= -\partial_t^2 \tilde{u}_1(l,t) - \lambda_1(l)\tilde{u}_1(l,t) + P_c(l)N_l(\partial_x)(\tilde{u})(l,t)\\ &= \varepsilon^3 c^2 K^2 \tilde{A}(K,T)\boldsymbol{E} - 2\varepsilon^5 \mathrm{i}cK(\partial_T \tilde{A}(K,T))\boldsymbol{E} - \varepsilon^7(\partial_T^2 \tilde{A}(K,T))\boldsymbol{E}\\ &\quad - \tfrac{1}{2}\varepsilon^3 \lambda_1''(0)K^2 \tilde{A}(K,T)\boldsymbol{E} - \varepsilon^5 \lambda_1''''(0)K^4 \tilde{A}(K,T)\boldsymbol{E}/24 + \mathcal{O}(\varepsilon^7)\\ &\quad + \varepsilon^5 \int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \nu_2 K^2 \tilde{A}(K-M,T)\tilde{A}(M,T)\,\mathrm{d}M\,\boldsymbol{E} + \mathcal{O}(\varepsilon^7)\\ &= \mathcal{O}(\varepsilon^7)\end{aligned}$$

if we choose $c$ and $\tilde{A}$ as above. We have $\mathcal{O}(\varepsilon^7)$ and not $\mathcal{O}(\varepsilon^6)$ since $P_c(\mathrm{Res}(\tilde{u}))(l,t)$ does not depend on $x$ and has to be even w.r.t. factors in $l$, i.e. $\varepsilon^5 K^4 \tilde{A}(K,T)$ is allowed, but $\varepsilon^6 K^5 \tilde{A}(K,T)$ is not. Next we have

$$\begin{aligned}&P_s(\mathrm{Res}(\tilde{u}))(l,x,t)\\ &\quad = -\partial_t^2 \tilde{v}(l,x,t) - L_l(\partial_x)\tilde{v}(l,x,t) + P_s(l)N_l(\partial_x)(\tilde{u})(l,x,t)\\ &\quad = c^2K^2(\varepsilon^6 \tilde{B}(K,x,T) + \varepsilon^7 \tilde{B}_2(K,x,T) + \varepsilon^8 \tilde{B}_3(K,x,T))\boldsymbol{E}\\ &\qquad - 2\mathrm{i}cK(\varepsilon^8 \partial_T \tilde{B}(K,x,T) + \varepsilon^9 \partial_T \tilde{B}_2(K,x,T) + \varepsilon^{10}\partial_T \tilde{B}_3(K,x,T))\boldsymbol{E}\\ &\qquad - (\varepsilon^{10}\partial_T^2 \tilde{B}(K,x,T) + \varepsilon^{11}\partial_T^2 \tilde{B}_2(K,x,T) + \varepsilon^{12}\partial_T^2 \tilde{B}_3(K,x,T))\boldsymbol{E}\\ &\qquad - (\varepsilon^4 \tilde{L}_{\varepsilon K}(\partial_x)\tilde{B}(K,x,T) + \varepsilon^5 \tilde{L}_{\varepsilon K}(\partial_x)\tilde{B}_2(K,x,T)\\ &\qquad + \varepsilon^6 \tilde{L}_{\varepsilon K}(\partial_x)\tilde{B}_3(K,x,T))\boldsymbol{E} + P_s(\varepsilon K)N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K,x,T/\varepsilon),\end{aligned}$$

where we expand

$$P_s(l)N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K, x, T/\varepsilon) \\ = (\varepsilon^4 F_4(\tilde{A}) + \varepsilon^5 F_5(\tilde{A}) + \varepsilon^6 F_6(\tilde{A}, \tilde{B}) + \mathcal{O}(\varepsilon^7))(K, x, T)\boldsymbol{E}.$$

If we set

$$\begin{aligned} 0 &= -\tilde{L}_{\varepsilon K}(\partial_x)\tilde{B}(K, x, T) + \varepsilon^4 F_4(\tilde{A})(K, x, T), \\ 0 &= -\tilde{L}_{\varepsilon K}(\partial_x)\tilde{B}_2(K, x, T) + \varepsilon^4 F_5(\tilde{A})(K, x, T), \\ 0 &= -\tilde{L}_{\varepsilon K}(\partial_x)\tilde{B}_3(K, x, T) + \varepsilon^4 F_6(\tilde{A}, \tilde{B})(K, x, T) + c^2 K^2 \tilde{B}(K, x, T), \end{aligned}$$

we finally have

$$P_s(\mathrm{Res}(\tilde{u}))(l, x, t) = \mathcal{O}(\varepsilon^7).$$

The functions $\tilde{B}$, $\tilde{B}_2$ and $\tilde{B}_3$ are well defined since $\tilde{L}_{\varepsilon K}$ can be inverted on the range of $P_s(\varepsilon K)$.

*The improved inviscid Burgers approximation.* We leave this part to the reader; we refer to Appendix A, where the modified approximation is discussed for the spatially homogeneous situation.

*The improved Whitham approximation.* We need the residual formally to be of order $\mathcal{O}(\varepsilon^3)$. With the previous approximation we have $\mathcal{O}(\varepsilon^3)$ for the $P_c$-part of the residual again due to symmetry, but we only have $\mathcal{O}(\varepsilon^2)$ for the $P_s$-part. As above, we modify our ansatz to

$$\tilde{u}_1(l, t) = \varepsilon^{-1}\tilde{A}(K, T) \quad \text{and} \quad \tilde{v}(l, x, t) = \tilde{B}(K, x, T) + \varepsilon^2 \tilde{B}_2(K, x, T).$$

We define $\tilde{A}$ and $\tilde{B}$ exactly as above, and $\tilde{B}_2$ as a solution of

$$-\partial_T^2 \tilde{B}(K, x, T) - \tilde{L}_{\varepsilon K}(\partial_x)\tilde{B}_2(K, x, T) = 0,$$

which is again well defined due the fact that $\tilde{L}_{\varepsilon K}$ can be inverted on the range of $P_s(\varepsilon K)$.

For all three approximations we gain a factor $\varepsilon^{1/2}$ when we estimate the error in $L^2$-based spaces, due to the scaling properties of the $L^2$-norm. Since the error made by the various approximations will be estimated in physical space via energy estimates, we conclude the following for the KdV approximation, the inviscid Burgers approximation and the Whitham approximation, respectively.

LEMMA 4.1. *Let $A \in C([0, T_0], H^6)$ be a solution of the KdV equation (3.7). Then there exist $\varepsilon_0 > 0$ and $C_{\mathrm{res}}$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have*

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|\mathrm{Res}(\varepsilon^2 \varPsi(\cdot, t, \varepsilon))\|_{H^1} \leqslant C_{\mathrm{res}} \varepsilon^{15/2}$$

*and*

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|\partial_x^{-1} \mathrm{Res}(\varepsilon^2 \varPsi(\cdot, t, \varepsilon))\|_{H^1} \leqslant C_{\mathrm{res}} \varepsilon^{13/2}.$$

LEMMA 4.2. *Let $\alpha = 1$ and let $A \in C([0,T_0],H^4)$ be a solution of the inviscid Burgers equation (3.8). Then there exist $\varepsilon_0 > 0$ and $C_{\mathrm{res}}$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon^{1+\alpha}]} \|\mathrm{Res}(\varepsilon^{\alpha}\Psi(\cdot,t,\varepsilon))\|_{H^1} \leqslant C_{\mathrm{res}}\varepsilon^{(7+4\alpha)/2}$$

*and*

$$\sup_{t\in[0,T_0/\varepsilon^{1+\alpha}]} \|\partial_x^{-1}\,\mathrm{Res}(\varepsilon^{\alpha}\Psi(\cdot,t,\varepsilon))\|_{H^1} \leqslant C_{\mathrm{res}}\varepsilon^{(5+4\alpha)/2}.$$

LEMMA 4.3. *Let $A \in C([0,T_0],H^4)$ be a solution of the Whitham equation (3.11). Then there exist $\varepsilon_0 > 0$ and $C_{\mathrm{res}}$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon]} \|\mathrm{Res}(\Psi(\cdot,t,\varepsilon))\|_{H^1} \leqslant C_{\mathrm{res}}\varepsilon^{7/2}$$

*and*

$$\sup_{t\in[0,T_0/\varepsilon]} \|\partial_x^{-1}\,\mathrm{Res}(\Psi(\cdot,t,\varepsilon))\|_{H^1} \leqslant C_{\mathrm{res}}\varepsilon^{5/2}.$$

## 5. The error estimates

As for the spatially homogeneous case, the proofs for the KdV approximations follow more or less line for line the proofs for the justification of the inviscid Burgers equation and of the Whitham system. Our approximation results are as follows.

THEOREM 5.1. *Let $A \in C([0,T_0],H^6(\mathbb{R}))$ be a solution of the KdV equation (3.7). Then there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have solutions $u \in C([0,T_0/\varepsilon^3],H^2)$ of the spatially periodic Boussinesq model (1.9) with*

$$\sup_{t\in[0,T_0/\varepsilon^3]} \|u(\cdot,t) - \varepsilon^2 A(\varepsilon(\cdot - t),\varepsilon^3 t)\|_{H^2} \leqslant C\varepsilon^{5/2}.$$

THEOREM 5.2. *Let $\alpha = 1$ and let $A \in C([0,T_0],H^4(\mathbb{R}))$ be a solution of the inviscid Burgers equation (3.8). Then there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have solutions $u \in C([0,T_0/\varepsilon^3],H^2)$ of the spatially periodic Boussinesq model (1.9) with*

$$\sup_{t\in[0,T_0/\varepsilon^{1+\alpha}]} \|u(\cdot,t) - \varepsilon^{\alpha} A(\varepsilon(\cdot - t),\varepsilon^{1+\alpha} t)\|_{H^2} \leqslant C\varepsilon^{(1+2\alpha)/2}.$$

THEOREM 5.3. *There exists a $C_1 > 0$ such that the following holds. Let $U \in C([0,T_0],H^4)$ be a solution of the Whitham system (3.11) with*

$$\sup_{T\in[0,T_0]} \|U(\cdot,T)\|_{H^4} \leqslant C_1.$$

*Then there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have solutions $u \in C^0([0,T_0/\varepsilon^3],H^2)$ of our spatially periodic Boussinesq model (1.9), such that*

$$\sup_{t\in[0,T_0/\varepsilon^3]} \|u(\cdot,t) - U(\varepsilon\cdot,\varepsilon t)\|_{H^2} \leqslant C_2\varepsilon^{1/2}.$$

*Proof of theorems 5.1–5.3.* Since we already have the estimates for the residuals in lemmas 4.1–4.3, from this point onwards the remaining estimates can be handled in exactly the same way. The $\alpha = 0$ case corresponds to the Whitham approximation, and the $\alpha = 2$ case to the KdV approximation.

The difference $\varepsilon^{(3+2\alpha)/2} R = u - \varepsilon^{\alpha}\Psi$ satisfies

$$\partial_t^2 R = \partial_x(a\partial_x R) - \partial_x^2(b\partial_x^2 R) + 2\partial_x(c\partial_x(\varepsilon^{\alpha}\Psi R)) \\ + \varepsilon^{(3+2\alpha)/2}\partial_x(c\partial_x(R^2)) + \varepsilon^{-(3+2\alpha)/2}\operatorname{Res}(\varepsilon^{\alpha}\Psi). \quad (5.1)$$

The first three terms on the right-hand side can be written as

$$\partial_x(a\partial_x R) - \partial_x^2(b\partial_x^2 R) + 2\partial_x(c\varepsilon^{\alpha}\Psi\partial_x R) + 2\partial_x(c(\partial_x\varepsilon^{\alpha}\Psi)R).$$

The last term is of order $\mathcal{O}(\varepsilon^{1+\alpha})$ due to the long-wave character of the approximation $\varepsilon^{\alpha}\Psi$. More essentially, the first three terms can be written as $\partial_x(B(\partial_x R))$, where $B$ is the self-adjoint operator

$$B = (a + 2c\varepsilon^{\alpha}\Psi) - \partial_x(b\partial_x).$$

In the $\alpha > 0$ case for sufficiently small $\varepsilon > 0$ and in the $\alpha = 0$ case for sufficiently small $\|\Psi\|_{C_b^0}$, the linear operator $B$ is positive definite. Hence, there exists a positive-definite self-adjoint operator $\mathcal{A}$ with $\mathcal{A}^2 = B$. The associated operator norm $\|\cdot\|_{\mathcal{A}} = \|\mathcal{A}\cdot\|_{L^2}$ is then equivalent to the $H^1$-norm, and $\mathcal{A}^{-1}$ is a bounded operator from $L^2$ to $H^1$. Hence, the equation for the error can be written as

$$\partial_t^2 R = \partial_x(\mathcal{A}^2(\partial_x R)) + 2\partial_x(c(\partial_x\varepsilon^{\alpha}\Psi)R) \\ + \varepsilon^{(3+2\alpha)/2}\partial_x(c\partial_x(R^2)) + \varepsilon^{-(3+2\alpha)/2}\operatorname{Res}(\varepsilon^{\alpha}\Psi). \quad (5.2)$$

In order to bound the solutions of (5.2), we use energy estimates. Therefore, we first multiply (5.2) by $\partial_t R$ and integrate the expression obtained w.r.t. $x$. We obtain

$$\int(\partial_t R)\partial_t^2 R\,\mathrm{d}x = \tfrac{1}{2}\partial_t\int(\partial_t R)^2\,\mathrm{d}x$$

and

$$\begin{aligned}\int(\partial_t R)\partial_x(\mathcal{A}^2(\partial_x R))\,\mathrm{d}x &= -\int(\partial_t\partial_x R)(\mathcal{A}^2(\partial_x R))\,\mathrm{d}x \\ &= -\int(\mathcal{A}\partial_t\partial_x R)(\mathcal{A}\partial_x R)\,\mathrm{d}x \\ &= -\int(\partial_t(\mathcal{A}\partial_x R))(\mathcal{A}\partial_x R)\,\mathrm{d}x - \int([\partial_t,\mathcal{A}]\partial_x R)(\mathcal{A}\partial_x R)\,\mathrm{d}x \\ &= -\tfrac{1}{2}\partial_t\int(\mathcal{A}\partial_x R)^2\,\mathrm{d}x - \int([\partial_t,\mathcal{A}]\partial_x R)(\mathcal{A}\partial_x R)\,\mathrm{d}x,\end{aligned}$$

where

$$[\partial_t,\mathcal{A}]\cdot = \partial_t(\mathcal{A}\cdot) - \mathcal{A}\partial_t\cdot$$

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

is the commutator of the operators $\mathcal{A}$ and $\partial_t$. Moreover, we estimate

$$\left|\int (\partial_t R) 2\partial_x (c(\partial_x \varepsilon^\alpha \Psi) R)\,\mathrm{d}x\right| \leqslant C\varepsilon^{1+\alpha} \|\partial_t R\|_{L^2} \|R\|_{H^1},$$

$$\left|\int (\partial_t R) \varepsilon^{(3+2\alpha)/2} \partial_x (c\partial_x (R^2))\,\mathrm{d}x\right| \leqslant C\varepsilon^{(3+2\alpha)/2} \|\partial_t R\|_{L^2} \|R\|_{H^2}^2,$$

$$\left|\int (\partial_t R) \varepsilon^{-(3+2\alpha)/2} \operatorname{Res}(\varepsilon^\alpha \Psi)\,\mathrm{d}x\right| \leqslant C\varepsilon^{2+\alpha} \|\partial_t R\|_{L^2},$$

using lemmas 4.1–4.3. Finally, we have

$$[\partial_t, \mathcal{A}]\partial_x R = (\partial_t \mathcal{A})\partial_x R$$

such that

$$\int ([\partial_t, \mathcal{A}]\partial_x R)(\mathcal{A}\partial_x R)\,\mathrm{d}x = \int ((\partial_t \mathcal{A})\partial_x R)(\mathcal{A}\partial_x R)\,\mathrm{d}x.$$

In order to control this term we first note that

$$(\partial_t \mathcal{A})\mathcal{A} + \mathcal{A}\partial_t \mathcal{A} = \partial_t(\mathcal{A}^2) = 2c\partial_t(\varepsilon^\alpha \Psi)$$

and

$$((\partial_t \mathcal{A})u, v)_{L^2} = (u, (\partial_t \mathcal{A})v)_{L^2},$$

which follows from differentiating the associated formula for $\mathcal{A}$ w.r.t. $t$ such that

$$\begin{aligned}\left|2\int ((\partial_t \mathcal{A})\partial_x R)(\mathcal{A}\partial_x R)\,\mathrm{d}x\right| &= \left|\int (\mathcal{A}(\partial_t \mathcal{A})\partial_x R)\partial_x R + \partial_x R(\partial_t \mathcal{A}(\mathcal{A}\partial_x R))\,\mathrm{d}x\right| \\ &= \left|\int \partial_x R(\mathcal{A}\partial_t \mathcal{A} + (\partial_t \mathcal{A})\mathcal{A})\partial_x R\,\mathrm{d}x\right| \\ &= \left|\int 2c(\partial_t(\varepsilon^\alpha \Psi))(\partial_x R)^2\,\mathrm{d}x\right| \\ &\leqslant 2\sup_{x\in\mathbb{R}} |c(x)\partial_t(\varepsilon^\alpha \Psi(x,t))| \|\partial_x R\|_{L^2}^2 \\ &= \mathcal{O}(\varepsilon^{1+\alpha})\|\partial_x R\|_{L^2}^2.\end{aligned}$$

In order to obtain a bound for the $L^2$-norm of $R$ and not just of its derivatives, we next multiply '$\partial_x^{-1}(5.2)$' by $\mathcal{A}^{-2}\partial_x^{-1}\partial_t R$ and integrate the expression obtained in this way w.r.t. $x$. We find

$$\begin{aligned}\int (\mathcal{A}^{-2}\partial_x^{-1}\partial_t R)\partial_x^{-1}\partial_t^2 R\,\mathrm{d}x &= \int (\mathcal{A}^{-1}\partial_x^{-1}\partial_t R)\mathcal{A}^{-1}\partial_t\partial_x^{-1}\partial_t R\,\mathrm{d}x \\ &= \tfrac{1}{2}\partial_t \int (\mathcal{A}^{-1}\partial_x^{-1}\partial_t R)^2\,\mathrm{d}x \\ &\quad - \int (\mathcal{A}^{-1}\partial_x^{-1}\partial_t R)[\partial_t, \mathcal{A}^{-1}]\partial_x^{-1}\partial_t R\,\mathrm{d}x,\end{aligned}$$

$$\int (\mathcal{A}^{-2}\partial_x^{-1}\partial_t R)\partial_x^{-1}\partial_x \mathcal{A}^2 \partial_x R\,\mathrm{d}x = -\tfrac{1}{2}\partial_t \int R^2\,\mathrm{d}x.$$

Moreover, using $\mathcal{A}^{-1}\colon L^2 \to H^1$ and the self-adjointness of $\mathcal{A}^{-1}$ we estimate

$$\begin{aligned}
&\left| \int (\mathcal{A}^{-2}\partial_x^{-1}\partial_t R) 2\partial_x^{-1}\partial_x (c(\partial_x \varepsilon^\alpha \Psi)R)\,\mathrm{d}x \right| \\
&\qquad = \left| \int (\mathcal{A}^{-1}\partial_x^{-1}\partial_t R) 2\mathcal{A}^{-1} (c(\partial_x \varepsilon^\alpha \Psi)R)\,\mathrm{d}x \right| \\
&\qquad \leqslant C\varepsilon^{1+\alpha} \|\partial_x^{-1}\partial_t R\|_{L^2} \|R\|_{L^2}, \\
&\left| \int (\mathcal{A}^{-2}\partial_x^{-1}\partial_t R) \varepsilon^{(3+2\alpha)/2} \partial_x^{-1}\partial_x (c\partial_x(R^2))\,\mathrm{d}x \right| \\
&\qquad = \left| \int (\mathcal{A}^{-1}\partial_x^{-1}\partial_t R) \varepsilon^{(3+2\alpha)/2} \mathcal{A}^{-1} (c\partial_x(R^2))\,\mathrm{d}x \right| \\
&\qquad \leqslant C\varepsilon^{(3+2\alpha)/2} \|\partial_x^{-1}\partial_t R\|_{L^2} \|R\|_{H^1}^2, \\
&\left| \int (\mathcal{A}^{-2}\partial_x^{-1}\partial_t R) \varepsilon^{-(3+2\alpha)/2} \partial_x^{-1} \operatorname{Res}(\varepsilon^\alpha \Psi)\,\mathrm{d}x \right| \\
&\qquad \leqslant C\varepsilon^{1+\alpha} \|\partial_x^{-1}\partial_t R\|_{L^2},
\end{aligned}$$

again using lemmas 4.1–4.3. Finally, we have

$$[\partial_t, \mathcal{A}^{-1}]\partial_x R = (\partial_t \mathcal{A}^{-1})\partial_x R$$

such that

$$\int (\mathcal{A}^{-1}\partial_x^{-1}\partial_t R)[\partial_t, \mathcal{A}^{-1}]\partial_x^{-1}\partial_t R\,\mathrm{d}x = \int (\mathcal{A}^{-1}\partial_x^{-1}\partial_t R)(\partial_t \mathcal{A}^{-1})\partial_x^{-1}\partial_t R\,\mathrm{d}x.$$

We write this as half of

$$\begin{aligned}
&\int ((\partial_t \mathcal{A}^{-1})\partial_x^{-1}\partial_t R)(\mathcal{A}^{-1}\partial_x^{-1}\partial_t R)\,\mathrm{d}x + \int (\mathcal{A}^{-1}\partial_x^{-1}\partial_t R)(\partial_t \mathcal{A}^{-1})\partial_x^{-1}\partial_t R\,\mathrm{d}x \\
&\qquad = \int (\partial_x^{-1}\partial_t R)((\partial_t \mathcal{A}^{-1})\mathcal{A}^{-1} + \mathcal{A}^{-1}\partial_t \mathcal{A}^{-1})\partial_x^{-1}\partial_t R\,\mathrm{d}x \\
&\qquad = \int (\partial_x^{-1}\partial_t R)(\partial_t(\mathcal{A}^{-2}))\partial_x^{-1}\partial_t R\,\mathrm{d}x \\
&\qquad =: s_1.
\end{aligned}$$

From

$$\partial_t(\mathcal{A}^2\mathcal{A}^{-2}) = (\partial_t(\mathcal{A}^2))\mathcal{A}^{-2} + \mathcal{A}^2\partial_t(\mathcal{A}^{-2}) = 0$$

it follows that

$$\partial_t(\mathcal{A}^{-2}) = -\mathcal{A}^{-2}(\partial_t(\mathcal{A}^2))\mathcal{A}^{-2} = -\mathcal{A}^{-2}(\partial_t(\varepsilon^\alpha \Psi))\mathcal{A}^{-2}$$

such that

$$s_1 = \int (\mathcal{A}^{-2}\partial_x^{-1}\partial_t R)(\partial_t(\varepsilon^\alpha \Psi))(\mathcal{A}^{-2}\partial_x^{-1}\partial_t R)\,\mathrm{d}x = \mathcal{O}(\varepsilon^{1+\alpha})\|\mathcal{A}^{-2}\partial_x^{-1}\partial_t R\|_{L^2}^2,$$

which can be bounded by $\mathcal{O}(\varepsilon^{1+\alpha})\|\mathcal{A}^{-1}\partial_x^{-1}\partial_t R\|_{L^2}^2$. If we define

$$E(t) = \tfrac{1}{2}(\|\partial_t R\|_{L^2}^2 + \|\mathcal{A}^{-1}\partial_x^{-1}\partial_t R\|_{L^2}^2 + \|R\|_{L^2}^2 + \|\mathcal{A}\partial_x R\|_{L^2}^2),$$

we find

$$\frac{\mathrm{d}}{\mathrm{d}t}E \leqslant C_1\varepsilon^{1+\alpha}E + C_2\varepsilon^{(3+2\alpha)/2}E^{3/2} + C_3\varepsilon^{1+\alpha}E^{1/2}$$
$$\leqslant C_1\varepsilon^{1+\alpha}E + C_2\varepsilon^{(3+2\alpha)/2}E^{3/2} + C_3\varepsilon^{1+\alpha} + C_3\varepsilon^{1+\alpha}E,$$

with constants $C_1$, $C_2$ and $C_3$ independent of $0 < \varepsilon \ll 1$ since all the $\|\partial_t R\|_{L^2}$, $\|\mathcal{A}^{-1}\partial_x^{-1}\partial_t R\|_{L^2}$, and so on, appearing above can be estimated by $E^{1/2}$. Choosing $\varepsilon^{1/2}E^{1/2} \leqslant 1$ gives

$$\frac{\mathrm{d}}{\mathrm{d}t}E(t) \leqslant (C_1 + C_2 + C_3)\varepsilon^{1+\alpha}E + C_3\varepsilon^{1+\alpha},$$

which can be estimated with Gronwall's inequality and yields

$$E(t) \leqslant C_3 T_0 \mathrm{e}^{(C_1+C_2+C_3)T_0} =: M$$

for all $0 \leqslant \varepsilon^{1+\alpha}t \leqslant T_0$. Choosing $\varepsilon_0 > 0$ so small that $\varepsilon_0^{1/2}M^{1/2} \leqslant 1$ gives the required estimate first for $E(t)$. Since in the $\alpha > 0$ case for sufficiently small $\varepsilon > 0$ and the $\alpha = 0$ case for sufficiently small $\|\Psi\|_{C_b^2}$ the quantity $E^{1/2}$ is equivalent to the $H^2$-norm of $R$, the proof of theorems 5.1–5.3 is complete. □

## 6. Discussion

This section gives some heuristic arguments for why the previous approach works and demonstrates the approach in a larger framework.

The error equation (2.1) of the spatially homogeneous Boussinesq equation (1.1) can be written at lowest order in the form of a Hamiltonian system:

$$\partial_t \begin{pmatrix} R \\ w \end{pmatrix} = \begin{pmatrix} w \\ \partial_x^2 R - \partial_x^4 R + \varepsilon^\alpha \Psi \partial_x^2 R + \mathcal{O}(\varepsilon^{1+\alpha}) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \partial_R H \\ \partial_w H \end{pmatrix},$$

with the Hamiltonian

$$H = \frac{1}{2}\int w^2 + (\partial_x R)^2 + (\partial_x^2 R)^2 + \varepsilon^\alpha \Psi (\partial_x R)^2 \,\mathrm{d}x,$$

where for this presentation we have used $\partial_x \Psi = \mathcal{O}(\varepsilon)$. This Hamiltonian is a part of our energy, and it can be used to estimate parts of the $H^2$-norm. Since $\Psi$ depends on $t$, the Hamiltonian is not conserved, but we have

$$\frac{\mathrm{d}}{\mathrm{d}t}H = \nabla H \cdot \partial_t \begin{pmatrix} R \\ w \end{pmatrix} + \partial_t H = 0 + \mathcal{O}(\varepsilon^{1+\alpha}) \tag{6.1}$$

since $\partial_t \Psi = \mathcal{O}(\varepsilon)$ due to the long-wave character of the approximation.

The spatially periodic case can be understood in a similar way. The error equation (5.2) of the spatially homogeneous Boussinesq equation (1.9) can be written at lowest order in the form of a Hamiltonian system:

$$\partial_t \begin{pmatrix} R \\ w \end{pmatrix} = \begin{pmatrix} w \\ \partial_x(\mathcal{A}^2(\partial_x R)) + \mathcal{O}(\varepsilon^{1+\alpha}) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \partial_R H \\ \partial_w H \end{pmatrix} + \mathcal{O}(\varepsilon^{1+\alpha}),$$

with the Hamiltonian

$$H = \frac{1}{2}\int w^2 + (\mathcal{A}\partial_x R)^2 \, \mathrm{d}x,$$

where for this presentation we have used $\partial_x \Psi = \mathcal{O}(\varepsilon)$. This Hamiltonian is also a part of our energy, and it can be used to estimate parts of the $H^2$-norm. Since $\mathcal{A}$ depends via $\Psi$ on $t$ the Hamiltonian is not conserved, but again we have (6.1) since $\partial_t \Psi = \mathcal{O}(\varepsilon)$ due to the long-wave character of the approximation.

As stated earlier, this paper was originally intended as the next step in generalizing a method developed in [9] for the justification of the KdV approximation in situations when the KdV modes are resonant with other long wave modes in [15] and for the justification of the Whitham approximation, respectively. The normal form transforms used in the proofs of [9, 15] leave the energy surfaces invariant and can therefore be avoided by our 'good' choice of energy. Hence, the toy problem considered in [9, 15] can also be handled with the approach presented here if the nonlinear terms are modified in such a way that a Hamiltonian structure is observed.

**Acknowledgements**

We are grateful to Florent Chazel for helping us to understand the existing literature. Moreover, we thank Martina Chirilus-Bruckner for a number of helpful discussions. The paper is partly supported by the Deutsche Forschungsgemeinschaft (DFG) under Grant no. Schn520/9-1.

## Appendix A. The inviscid Burgers approximation

This appendix provides more details about the derivation and the justification via error estimates for the inviscid Burgers approximation. Inserting the ansatz

$$\varepsilon^\alpha \Psi(x,t) = \varepsilon^\alpha A(\varepsilon(x-t), \varepsilon^{1+\alpha} t)$$

with $\alpha \in (0,2)$ into the homogeneous Boussinesq equation (1.9) gives the residual

$$\begin{aligned}\mathrm{Res}(u)(x,t) &= -\partial_t^2 u(x,t) + \partial_x^2 u(x,t) - \partial_x^4 u(x,t) + \partial_x^2 (u(x,t)^2) \\ &= \varepsilon^{\alpha+4}\partial_X^4 A + \varepsilon^{3\alpha+2}\partial_T^2 A\end{aligned}$$

and $A$ satisfying the inviscid Burgers equation

$$\partial_T A = -\tfrac{1}{2}\partial_X (A^2)$$

if the coefficient of $\varepsilon^{2\alpha+2}$ is set to zero. However, the residual is too large for the analysis in § 2. By adding higher-order terms to the approximation we obtain the estimates stated in remark 2.3:

$$\|\mathrm{Res}(\varepsilon^\alpha \Psi(\cdot,t,\varepsilon))\|_{L^2} = \mathcal{O}(\varepsilon^{(7+4\alpha)/2}), \qquad \|\partial_x^{-1}\,\mathrm{Res}(\varepsilon^\alpha \Psi(\cdot,t,\varepsilon))\|_{L^2} = \mathcal{O}(\varepsilon^{(5+4\alpha)/2}).$$

We consider the improved approximation:

$$\varepsilon^\alpha \Psi(x,t) = \varepsilon^\alpha A(\varepsilon(x-t), \varepsilon^{1+\alpha} t) + \varepsilon^\beta B(\varepsilon(x-t), \varepsilon^{1+\alpha} t)$$

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

with $\beta = \min\{2\alpha, 2\}$. For the residual we find

$$\operatorname{Res}(\varepsilon^2\Psi) = -2\varepsilon^{2+\alpha+\beta}\partial_T\partial_X B - \varepsilon^{2+2\alpha+\beta}\partial_T^2 B - \varepsilon^{4+\beta}\partial_X^4 B + 2\varepsilon^{2+\alpha+\beta}\partial_X^2(AB) + \varepsilon^{2+2\beta}\partial_X^2(B^2) + \varepsilon^{\alpha+4}\partial_X^4 A + \varepsilon^{3\alpha+2}\partial_T^2 A.$$

We choose $B$ satisfying

$$2\partial_T B = 2\partial_X(AB) + g$$

where

$$g = \begin{cases} \partial_X^{-1}\partial_T^2 A & \text{for } \alpha \in (0,1), \\ \partial_X^{-1}\partial_T^2 A + \partial_X^3 A & \text{for } \alpha = 1, \\ \partial_X^3 A & \text{for } \alpha \in (1,2). \end{cases}$$

By this choice we obtain

$$|\operatorname{Res}(\varepsilon^2\Psi)| = \mathcal{O}(\max\{\chi_{\alpha\neq 1}(\alpha)\min\{\varepsilon^{\alpha+4}, \varepsilon^{3\alpha+2}\}, \varepsilon^{2+2\alpha+\beta}, \varepsilon^{4+\beta}, \varepsilon^{2+2\beta}\}).$$

Hence, for $\alpha = 1$ only, where $\beta = 2$, this is of order $\mathcal{O}(\varepsilon^{4+2\alpha})$, which is the formal order necessary to obtain the $L^2$ bound. For all other values of $\alpha$, additional terms are necessary. For $\alpha \to 0$ and $\alpha \to 2$ the number of such terms goes to infinity and increasing regularity is necessary. We refrain from discussing the solvability of this system of amplitude equations. This question is non-trivial since for $\alpha = 1$ the term $\partial_X^{-1}\partial_T^2 A$ has to be computed, which is possible as the temporal derivatives can be expressed as spatial derivatives via the inviscid Burgers equation:

$$\partial_T^2 A = -\tfrac{1}{2}\partial_T\partial_X(A^2) = -\partial_X(A\partial_T A) = \tfrac{1}{2}\partial_X(A\partial_X(A^2)) = \tfrac{1}{3}\partial_X^2(A^3).$$

The estimate for $\partial_x^{-1}\operatorname{Res}(\varepsilon^\alpha\Psi)$ can also be obtained using this expression, since $\partial_T^2 B$ can now also be expressed as spatial derivatives.

## Appendix B. Higher regularity results

This appendix explains how the approximation results can be transferred from $H^2$ to $H^m$ with $m \geqslant 2$. Due to the $x$-dependent coefficients, energy estimates for the spatial derivatives turn out to be rather complicated. However, by considering time derivatives the previous ideas and energies still can be used. The spatial derivatives can then be estimated via the equation for the error, i.e.

$$LR = \partial_t^2 R - 2\partial_x(c\partial_x(\varepsilon^\alpha\Psi R)) - R - \varepsilon^{(3+2\alpha)/2}\partial_x(c\partial_x(R^2)) - \varepsilon^{-(3+2\alpha)/2}\operatorname{Res}(\varepsilon^2\Psi), \tag{B 1}$$

where

$$LR = \partial_x(a\partial_x R) - \partial_x^2(b\partial_x^2 R) - R.$$

The operator $L$ is invertible and maps $H^s$ into $H^{s+4}$, and $C^m([0, T_0/\varepsilon^{1+\alpha}], H^s)$ into $C^m([0, T_0/\varepsilon^{1+\alpha}], H^{s+4})$, respectively. For $R \in C^m([0, T_0/\varepsilon^{1+\alpha}], H^s)$, the right-hand side of (B 1) is in

$$C^{m-2}([0, T_0/\varepsilon^{1+\alpha}], H^s) \cap C^m([0, T_0/\varepsilon^{1+\alpha}], H^{s-2}).$$

An application of $L^{-1}$ to (B 1) shows that

$$R \in C^{m-2}([0, T_0/\varepsilon^{1+\alpha}], H^{s+4}) \cap C^m([0, T_0/\varepsilon^{1+\alpha}], H^{s+2}).$$

Iterating this process shows that temporal derivatives can be transformed into spatial derivatives.

It remains to obtain the estimates for the temporal derivatives. In order to do so we differentiate the equation for the error $m$ times w.r.t. $t$. We obtain an equation of the form

$$\partial_t^2(\partial_t^m R) = \partial_x(\mathcal{A}^2(\partial_x(\partial_t^m R))) + 2\partial_x(c(\partial_x \varepsilon^\alpha \Psi)(\partial_t^m R)) + \mathcal{O}(\varepsilon^{1+\alpha}) \qquad \text{(B 2)}$$

due to the fact that whenever a time derivative falls on $\mathcal{A}$ or $\Psi$ another $\varepsilon$ is gained. To bound the solutions of (B 2), we use energy estimates. Therefore, we first multiply (B 2) by $\partial_t^{m+1}R$ and integrate the expression obtained w.r.t. $x$. Next, as above, we multiply '$\partial_x^{-1}$(B 2)' by $\mathcal{A}^{-2}\partial_x^{-1}\partial_t R$ and integrate the expression obtained in this way w.r.t. $x$.

If we define

$$E_m(t) = \tfrac{1}{2}(\|\partial_t^{m+1}R\|_{L^2}^2 + \|\mathcal{A}^{-1}\partial_x^{-1}\partial_t^{m+1}R\|_{L^2}^2 + \|\partial_t^m R\|_{L^2}^2 + \|\mathcal{A}\partial_x\partial_t^m R\|_{L^2}^2).$$

we find

$$\frac{\mathrm{d}}{\mathrm{d}t}E_m \leqslant C_1\varepsilon^{1+\alpha}E_m + C_2\varepsilon^{(3+2\alpha)/2}\mathcal{E}_m^{3/2} + C_3\varepsilon^{1+\alpha},$$

with constants $C_1$, $C_2$ and $C_3$ independent of $0 < \varepsilon \ll 1$ and $\mathcal{E}_m = E + \cdots + E_m$. Summing all estimates for the $E_j$ for $j = 0, \ldots, m$ yields a similar inequality for $\mathcal{E}_m$. Applying Gronwall's inequality to this inequality gives, for example, the following result.

Theorem B.1. *Fix $s \in \mathbb{N}$ and let $A \in C([0, T_0], H^{6+s}(\mathbb{R}))$ be a solution of the KdV equation (3.7). Then there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have solutions $u \in C([0, T_0/\varepsilon^3], H^{2+s})$ of the spatially periodic Boussinesq model (1.9) with*

$$\sup_{t\in[0,T_0/\varepsilon^3]} \|u(\cdot,t) - \varepsilon^2 A(\varepsilon(\cdot - t), \varepsilon^3 t)\tilde{w}_1(0)(\cdot)\|_{H^{2+s}} \leqslant C\varepsilon^{5/2}.$$

Theorems 5.2 and 5.3 can be reformulated similarly.

## Appendix C. Bloch transform on the real line

In this appendix we recall some basic properties of the Bloch transform. Our presentation follows [16]. The Bloch transform $\mathcal{T}$ generalizes the Fourier transform $\mathcal{F}$ from spatially homogeneous problems to spatially periodic problems. The Bloch transform is (formally) defined by

$$\tilde{u}(l, x) = (\mathcal{T}u)(l, x) = \sum_{j\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}jx}\hat{u}(l + j), \qquad \text{(C 1)}$$

where $\hat{u}(\xi) = (\mathcal{F}u)(\xi)$, $\xi \in \mathbb{R}$, is the Fourier transform of $u$. The inverse of the Bloch transform is given by

$$u(x) = (\mathcal{T}^{-1}\tilde{u})(x) = \int_{-1/2}^{1/2} \mathrm{e}^{\mathrm{i}lx}\tilde{u}(l, x)\,\mathrm{d}l. \qquad \text{(C 2)}$$

https://doi.org/10.1017/S0308210518000227 Published online by Cambridge University Press

By construction, $\tilde{u}(l,x)$ is extended from $(l,x) \in \mathbb{T}_1 \times \mathbb{T}_{2\pi}$ to $(l,x) \in \mathbb{R} \times \mathbb{R}$ according to the continuation conditions:

$$\tilde{u}(l,x) = \tilde{u}(l, x+2\pi) \quad \text{and} \quad \tilde{u}(l,x) = \tilde{u}(l+1,x)\mathrm{e}^{\mathrm{i}x}. \tag{C 3}$$

The following lemma specifies the well-known property of the Bloch transform acting on Sobolev function spaces.

LEMMA C.1. *The Bloch transform $\mathcal{T}$ is an isomorphism between*

$$H^s(\mathbb{R}) \quad \textit{and} \quad L^2(\mathbb{T}_1, H^s(\mathbb{T}_{2\pi})),$$

*where $L^2(\mathbb{T}_1, H^s(\mathbb{T}_{2\pi}))$ is equipped with the norm*

$$\|\tilde{u}\|_{L^2(\mathbb{T}_1,H^s(\mathbb{T}_{2\pi}))} = \left( \int_{-1/2}^{1/2} \|\tilde{u}(l,\cdot)\|^2_{H^s(\mathbb{T}_{2\pi})} dl \right)^{1/2}.$$

Multiplication of two functions $u(x)$ and $v(x)$ in $x$-space corresponds to some convolution in Bloch space:

$$(\tilde{u} \star \tilde{v})(l,x) = \int_{-1/2}^{1/2} \tilde{u}(l-m,x)\tilde{v}(m,x)\,\mathrm{d}m, \tag{C 4}$$

where the continuation conditions (C 3) have to be used for $|l-m| > \frac{1}{2}$. If $\chi\colon \mathbb{R} \to \mathbb{R}$ is $2\pi$ periodic, then

$$\mathcal{T}(\chi u)(l,x) = \chi(x)(\mathcal{T}u)(l,x). \tag{C 5}$$

The relations (C 4) and (C 5) are well known and can be proved with the definition (C 1).

## References


1 J. L. Bona, T. Colin and D. Lannes. Long wave approximations for water waves. *Arch. Ration. Mech. Analysis* **178** (2005), 373–410.

2 K. Busch, G. Schneider, L. Tkeshelashvili and H. Uecker. Justification of the nonlinear Schrödinger equation in spatially periodic media. *Z. Angew. Math. Phys.* **57** (2006), 1–35.

3 C. Cercignani and D. H. Sattinger. *Scaling limits and models in physical processes.* Seminare der Deutschen Mathematiker-Vereinung, vol. 28 (Basel: Birkhäuser, 1998).

4 F. Chazel. Influence of bottom topography on long water waves. *Math. Model. Numer. Analysis* **41** (2007), 771–799.

5 F. Chazel. On the Korteweg–de Vries approximation for uneven bottoms. *Eur. J. Mech.* B **28** (2009), 234–252.

6 M. Chirilus-Bruckner, C. Chong, O. Prill and G. Schneider. Rigorous description of macroscopic wave packets in infinite periodic chains of coupled oscillators by modulation equations. *Discrete Contin. Dynam. Syst.* S **5** (2012), 879–901.

7 M. Chirilus-Bruckner, W.-P. Düll and G. Schneider. Validity of the KdV equation for the modulation of periodic traveling waves in the NLS equation. *J. Math. Analysis Applic.* **414** (2014), 166–175.

8 D. Chiron and F. Rousset. The KdV/KP-I limit of the nonlinear Schrödinger equation. *SIAM J. Math. Analysis* **42** (2010), 64–96.

9 C. Chong and G. Schneider. The validity of the KdV approximation in case of resonances arising from periodic media. *J. Math. Analysis Applic.* **383** (2011), 330–336.

10 W. Craig. An existence theory for water waves and the Boussinesq and Korteweg–de Vries scaling limits. *Commun. PDEs* **10** (1985), 787–1003.

11 T. Dohnal, A. Lamacz and B. Schweizer. Bloch-wave homogenization on large time scales and dispersive effective wave equations. *Multiscale Model. Simul.* **12** (2014), 488–513.

12 T. Dohnal, A. Lamacz and B. Schweizer. Dispersive homogenized models and coefficient formulas for waves in general periodic media. *Asymp. Analysis* **93** (2015), 21–49.

13 W.-P. Düll. Validity of the Korteweg–de Vries approximation for the two-dimensional water wave problem in the arc length formulation. *Commun. Pure Appl. Math.* **65** (2012), 381–429.

14 W.-P. Düll and G. Schneider. Validity of Whitham's equations for the modulation of wave-trains in the NLS equation. *J. Nonlin. Sci.* **19** (2009), 453–466.

15 W.-P. Düll, K. Sanei Kashani and G. Schneider. The validity of Whitham's approximation for a Klein–Gordon–Boussinesq model. *SIAM J. Math. Analysis* **48** (2016), 4311–4334.

16 T. Gallay, G. Schneider and H. Uecker. Stable transport of information near essentially unstable localized structures. *Discrete Contin. Dynam. Syst.* B **4** (2004), 349–390.

17 T. Iguchi. A long wave approximation for capillary–gravity waves and an effect of the bottom. *Commun. PDEs* **32** (2007), 37–85.

18 T. Iguchi. A shallow water approximation for water waves. *J. Math. Kyoto Univ.* **49** (2009), 13–55.

19 T. Kano and T. Nishida. A mathematical justification for Korteweg–de Vries equation and Boussinesq equation of water surface waves. *Osaka J. Math.* **23** (1986), 389–413.

20 L. V. Ovsyannikov. Cauchy problem in a scale of Banach spaces and its application to the shallow water theory justification. In *Applications of methods of functional analysis to problems in mechanics*. Lecture Notes in Mathematics, vol. 503, pp. 426–437 (Springer, 1976).

21 G. Schneider. Limits for the Korteweg–de Vries-approximation. *Z. Angew. Math. Mech.* **76** (1996), 341–344.

22 G. Schneider. The role of the nonlinear Schrödinger equation in nonlinear optics. In *Photonic crystals: mathematical analysis and numerical approximation* (ed. W. Dörfler, A. Lechleiter, M. Plum, G. Schneider and C. Wieners). Oberwolfach Seminars, vol. 42 (Basel: Birkhäuser, 2011).

23 G. Schneider and C. E. Wayne. The long-wave limit for the water wave problem. I. The case of zero surface tension. *Commun. Pure Appl. Math.* **53** (2000), 1475–1535.

24 G. Schneider and C. E. Wayne. The rigorous approximation of long-wavelength capillary–gravity waves. *Arch. Ration. Mech. Analysis* **162** (2002), 247–285.

25 G. B. Whitham. *Linear and nonlinear waves*. Pure and Applied Mathematics (New York: Wiley, 1999).