Double Degeneracy in Multiphase Modulation and the Emergence of the Boussinesq Equation

By Daniel J. Ratliff

In recent years a connection between conservation law singularity, or more generally zero characteristics arising within the linear Whitham equations, and the emergence of reduced nonlinear PDEs from systems generated by a Lagrangian density has been made in conservative systems. Remarkably the conservation laws form part of the reduced nonlinear system. Within this paper, the case of double degeneracy is investigated in multiphase wavetrains, characterised by a double zero characteristic of the linearised Whitham system, with the resulting modulation of relative equilibrium (which are a generalisation of the modulation of wavetrains) leading to a vector two-way Boussinesq equation. The derived PDE adheres to the previous results (such as Ratliff and Bridges, 2016) in the sense that all but one of its coefficients are related to the conservation laws along the relative equilibrium solution, which are then projected to form a corresponding scalar system. The theory is applied to two examples to highlight how both the criticality can be assessed and the two-way Boussinesq equation's coefficients are obtained. The first is the coupled NLS system and is the first time the two-way Boussinesq equation has been shown to arise in such a context, and the second is a stratified shallow water model which validates the theory against existing results.

keywords: Whitham modulation, Lagrangian dynamics, nonlinear waves, partial differential equations, asymptotic analysis.

1. Introduction

Whitham modulation remains a widely used tool in the study of nonlinear waves ([2, 3] and references therein for an overview of the Whitham

University of Surrey, Guildford, United Kingdom, GU2 7XH; email: d.ratliff@surrey.ac.uk

DOI: 10.1111/Bou.ss-Multi-Revision2-SIAPM
STUDIES IN APPLIED MATHEMATICS 00:1–30

methodology), which has been applied to various systems of physical interest such as water waves [4, 5], plasmas [6], optics [7, 8] and Bose-Einstein condensates [9, 10]. This technique provides equations for the slow evolution of the wavenumbers and frequencies for waves in such systems and thus insight may be obtained by studying these comparatively simpler systems. Although typically robust, there are points in the wavenumber-frequency space where the Whitham equations have zero characteristics. Such instances in the case of a single phase wavetrain have been shown to generate dispersion within the system and lead to well-known nonlinear PDEs governing the wavenumber emerging from the Euler-Lagrange equations [11, 12]. One particular instance of degeneracy in the single phase case, where the characterisitic of the Whitham system coalesce, leads to the two-way Boussinesq equation:

$$u_{tt} + \left(\frac{1}{2}u^2 \pm u_{xx}\right)_{xx} = 0\,,$$

where the function $u(x,t)$ perturbs the local wavenumber [12]. The two-way Boussinesq equation appears in physical contexts, primarily as a model in fluid dynamics [13, 14], and is named as such due to the second order time derivative allowing for left and right travelling waves. The aim of this current work is to generalise the reduction of the Euler-Lagrange equations in the case of two phase wavetrains (and in general, two phase relative equilibria) to the two-way Boussinesq equation, as well as formulating the criterion for such a reduction to hold using degeneracies of the linear Whitham equations.

In general, two phase wavetrains form relative equilibria when underlying symmetries are present, characterised through multiple parameters [16]. In the single phase case, relative equilibria are simply equilibria that move along some group orbit and are associated with a one parameter symmetry group, such as an affine or $SO(2)$ symmetry [17]. These may be written as

$$u(\theta,k,\omega) = G_\theta u_0(k,\omega)\,, \quad \theta = kx + \omega t + \theta_0$$

for equilibrium u_0, parameters k, ω and θ_0, which are constant, and group action G_θ. In the two phase case, such relative equilibria take a similar form,

$$\mathbf{u} = \hat{\mathbf{u}}(\theta_1,\theta_2;k_1,k_2,\omega_1,\omega_2) \equiv \hat{\mathbf{u}}(\boldsymbol{\theta};\mathbf{k},\boldsymbol{\omega}), \quad \theta_i = k_i x + \omega_i t + \theta_i^0, \qquad (1.1)$$

where we have introduced

$$\boldsymbol{\theta} = \begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix}k_1,\\ k_2\end{pmatrix}, \quad \boldsymbol{\omega} = \begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix},$$

and the θ_i^0 are arbitrary constants representing shifts in the phases. These are also associated with symmetries, in this case two independent one parameter symmetry groups. General multiphase wavetrains lead to small divisors [18, 19], however in the context of symmetry, relative equilibria and phase averaging considered in this paper these problems do not occur.

The principle role of wavetrains in this paper is as a mechanism in the generation of nonlinear PDEs through modulation arguments. Phase dynamics in conservative contexts has been well documented in the case of single phase wavetrains, highlighting the fact that such a methodology leads to coefficients that can be determined using only knowledge of the basic state [11, 12]. This lends real weight to the method, as the derivation itself need only be done once for any system that can be cast in the general form of the Lagrangian considered.

Unsurprisingly, these results carry over to the phase dynamics of multiple phases. The principle idea is to take a solution depending on two wave variables like (1.1) and use the invariance of phase translations to construct a new solution to the problem,

$$\mathbf{u} = \hat{\mathbf{u}}(\boldsymbol{\theta} + \varepsilon^p \boldsymbol{\phi}, \ldots), \ \varepsilon \ll 1. \tag{1.2}$$

Here $\boldsymbol{\phi}$ is a function of slowly scaled variables, and the other variables of the wavetrain are perturbed in similar ways. The substitution of this perturbed solution into the original system and subsequent Taylor expansion close to $\varepsilon = 0$ yields a series of equations (one for each power of ε) that, once solved, eventually lead to $\boldsymbol{\phi}$ or one of the other functions satisfying some vector PDE. The emergent vector systems in the framework considered here have analogous connections between the tensors appearing in the final PDE and the conservation laws for the system [1]. It has remained, up until now, an open question as to how double degeneracy, that is a loss of coefficients in both space and time derivative terms, takes form within the multiphase framework and how the theory is modified in such situations. This is the principle aim of this paper, and the outcome is that the most suitable model is the two-way Boussinesq equation as one may expect from the previous single phase studies.

We restrict ourselves to the class of PDEs generated from a Lagrangian density. In particular it is assumed to be in multisymplectic form, so that the Lagrangian considered is given by

$$\mathscr{L}(Z, Z_x, Z_t) = \iint \frac{1}{2}\langle Z, \mathbf{M}Z_t\rangle + \frac{1}{2}\langle Z, \mathbf{J}Z_x\rangle - S(Z)\ \mathrm{d}x\,\mathrm{d}t\,. \tag{1.3}$$

The Lagrangian density is integrated over some box $[x_1, x_2] \times [t_1, t_2]$, $Z(x, t)$ denotes the state vector for the system and the standard inner product is denoted by $\langle \cdot, \cdot \rangle$. The benefits of this formulation are twofold - firstly, it relates the conservation laws to the geometric formulation of the system. Secondly, it allows for a large amount of simplification within the modulation analysis. The main construct studied here will be the Euler-Lagrange equation for the above Lagrangian,

$$\mathbf{M}Z_t + \mathbf{J}Z_x = \nabla S(Z) \tag{1.4}$$

for skew-symmetric matrices $\mathbf{M}$ and $\mathbf{J}$, ∇ denotes the gradient with respect to Z and some Hamiltonian function S generated through the Legendre transforms. The modulational ansatz will eventually be substituted into this system, and it is precisely within this structure that the emergent nonlinear PDEs arise most clearly and with the desired form for the coefficients. A key assumption of this paper is that the equation (1.4) possesses a two phase solution of the form

$$Z = \widehat{Z}(\boldsymbol{\theta}, \mathbf{k}, \boldsymbol{\omega})\,. \tag{1.5}$$

It will be this solution that is modulated, which will give rise to the two-way Boussinesq obtained in this paper.

The assumption is then made that the system possesses two conservation laws, each with its own generator. For two phases, these conservation laws take the form

$$A(x,t)_t + B(x,t)_x = 0, \quad A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \quad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}.$$

Evaluating these along the wavetrain solution $\widehat{Z}$ obtains their $(\mathbf{k}, \boldsymbol{\omega})$ space counterparts:

$$\mathbf{A}(\mathbf{k}, \boldsymbol{\omega}) = \begin{pmatrix} \mathscr{A}_1 \\ \mathscr{A}_2 \end{pmatrix}, \quad \mathbf{B}(\mathbf{k}, \boldsymbol{\omega}) = \begin{pmatrix} \mathscr{B}_1 \\ \mathscr{B}_2 \end{pmatrix}. \tag{1.6}$$

The first key result of this paper is that through phase dynamical arguments, one is able to obtain the linear Whitham equations

$$\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\Omega}_T + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}\boldsymbol{\Omega}_X + \mathrm{D}_{\mathbf{k}}\mathbf{A}\mathbf{q}_T + \mathrm{D}_{\mathbf{k}}\mathbf{B}\mathbf{q}_X = 0, \quad \mathbf{q}_T - \boldsymbol{\Omega}_X = 0, \tag{1.7}$$

where D denotes the directional derivative of the respective subscript, and the functions $\mathbf{q}$, $\boldsymbol{\Omega}$ are defined as

$$\mathbf{q} = \boldsymbol{\phi}_X, \quad \boldsymbol{\Omega} = \boldsymbol{\phi}_T,$$

for the $\boldsymbol{\phi}$ introduced in (1.2). The linear Whitham equations are typically nondegenerate, but scenarios in which one or more of the terms become singular are of interest as a rescaling of the slow variables leads to the emergence of nonlinearity and dispersion from the phase dynamics. The previous multiphase modulation studies in [1] focused on the case where the tensor of the last term in (1.7) is singular,

$$\det\big[\mathrm{D}_{\mathbf{k}}\mathbf{B}\big] = 0, \tag{1.8}$$

meaning that the conservation law $\mathbf{B}$ is considered critical with respect to $\mathbf{k}$ along some surface $\mathbf{k}_0(\boldsymbol{\omega})$. Throughout the paper, this zero eigenvalue is assumed to be simple. Along such surfaces, the mapping $\mathbf{k} \mapsto \mathbf{B}(\mathbf{k}, \boldsymbol{\omega})$ is no longer bijective. Such a condition is primarily mathematical, but has some physical connotations such as representing a stability threshold when considered in stratified flows [20] and the interaction of two wave groups [16], and these connection are further discussed within the paper. In cases where (1.8) holds a KdV equation is attainable as a reduction of the Euler-Lagrange equations [1] and is of the form

$$\boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta} U_T + \boldsymbol{\zeta}^T \mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta}, \boldsymbol{\zeta}) U U_X + \boldsymbol{\zeta}^T \mathbf{K} U_{XXX} = 0\,, \tag{1.9}$$

for a scalar unknown function $U(x, t)$, and the vector $\mathbf{K}$ arises from a linear algebra analysis which is detailed within the paper. A remarkable feature of the above PDE is that the tensors appearing from the phase dynamics are again related to directional derivatives of the vectors (1.6), and the eigenvector associated with the zero eigenvalue (henceforth denoted as the zero eigenvector) of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$, denoted as $\boldsymbol{\zeta}$, has been used to project in the direction of the kernel to a scalar PDE.

Of interest now though is additional degeneracies of (1.7), namely of the middle two terms of this equation. The theory of the paper indicates that this second degeneracy occurs when the quantity

$$\boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta} = 0, \tag{1.10}$$

indicating the time derivative term in the above KdV vanishes and as a consequence the emergence of a generalised eigenvector $\boldsymbol{\gamma}$, given by

$$\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\gamma} = (\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta}\,.$$

Interestingly, this corresponds to a double zero eigenvalue condition of (1.7) when cast as a 4×4 matrix PDE. The relevance of such a link is that one may in fact now predict the relevant modulation equations by characterising the changes of behaviour in (1.7), rather than just relying on conservation law criticality. Again, these thresholds often signal a change in system stability and so the reduction about these regions offers insight into system behaviour close to, and either side, of stability boundaries.

The phase dynamical theory, in light of both the first and second singularities, must be slightly modified. This is because both the linear Whitham equations or the KdV equation can no longer be attained in such cases. To account for the degeneracy (1.10) we must introduce a second set of modulational functions that, although are dependent on the same slowly scaled variables, have a different prefactor scaling. Principally we use the ansatz

$$Z = \widehat{Z}(\boldsymbol{\theta}+\varepsilon\boldsymbol{\phi}+\varepsilon^2\boldsymbol{\psi}, \mathbf{k}+\varepsilon^2\mathbf{q}+\varepsilon^3\boldsymbol{r}, \boldsymbol{\omega}+\varepsilon^3\boldsymbol{\Omega}+\varepsilon^4\boldsymbol{\tau})+\varepsilon^3 W(\boldsymbol{\theta}, X, T, \varepsilon) \quad (1.11)$$

where W is a remainder term used to regulate the analysis and the modulation functions depend on the slowly scaled variables

$$X = \varepsilon x, \quad T = \varepsilon^2 t.$$

The benefit of using the ansatz (1.11) is apparent for three reasons. Firstly, many of the terms in the analysis leading to (1.12) cancel due to properties of the solution $\widehat{Z}$. Secondly, it is this ansatz along with the multisymplectic form of the Euler-Lagrange equations (1.4) that ultimately leads to the conservation laws emerging as coefficients. Finally, by considering an abstract form on the solution and governing equations means that the result of this paper may be applied to any system generated by a Lagrangian density with a two parameter symmetry.

The process to obtain (1.12) below is much the same as the preceding works: substitute (1.11) into (1.4), compute the Taylor expansion around $\varepsilon = 0$ and solve at each order in ε. Full details of this procedure and the terms involved will be discussed within the paper. The small parameter ε typically characterises the distance of the perturbation from the chosen point in $(\mathbf{k}, \boldsymbol{\omega})$ space, but in other derivations of the Boussinesq equation it represents a physical property of the system (such as wave steepness in [13, 15]).

The effect of the degeneracies (1.8) and (1.10) discussed is two-fold. The first of these relates the elements of the vector $\mathbf{q}$ through constants of proportionality determined by $\boldsymbol{\zeta}$,

$$\boldsymbol{\zeta} U = \mathbf{q}$$

for some slowly varying, initially arbitrary scalar function $U(X, T)$ which will play the role of the unknown function in the final PDE. The second relates the first set of modulation functions $(\boldsymbol{\phi}, \mathbf{q}, \boldsymbol{\Omega})$, or more accurately U, to the second set $(\boldsymbol{\psi}, \boldsymbol{r}, \boldsymbol{\tau})$ through

$$\boldsymbol{r}_X = -\boldsymbol{\gamma} U_T.$$

We demonstrate the ultimate result of the above in the modulation theory is the vector Boussinesq equation:

$$\Big(\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\zeta} - (\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\gamma}\Big)U_{TT} + \Big(\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta}, \boldsymbol{\zeta})UU_X + \mathbf{K}U_{XXX}\Big)_X + \mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\alpha}_{XXX} = 0\,, \quad (1.12)$$

for some arbitrary vector function $\boldsymbol{\alpha}(X,T)$. In this form the system is not closed, as the vector $\boldsymbol{\alpha}$ is unknown and causes the vector Boussinesq equation to be inhomogeneous. Without it, the vector Boussinesq equation generically admits trivial dynamics since U is scalar valued and would have two scalar PDEs with different coefficients governing the evolution of U. As $\boldsymbol{\zeta} \in \ker(\mathrm{D}_{\mathbf{k}}\mathbf{B})$ and $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ is symmetric, by multiplying by $\boldsymbol{\zeta}$ on the left one is able to eliminate the $\boldsymbol{\alpha}_{XX}$ term and obtain the scalar version of the PDE,

$$\begin{aligned}\boldsymbol{\zeta}^T\Big(\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\zeta} - (\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\gamma}\Big)U_{TT}& \\ + \Big(\boldsymbol{\zeta}^T\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta})UU_X + \boldsymbol{\zeta}^T\mathbf{K}U_{XXX}\Big)_X &= 0.\end{aligned} \tag{1.13}$$

The advantage revealed through the phase dynamics approach is that all but one of the coefficients have a direct relation to the conservation laws and allows them to be computed a priori using information from the relative equilibrium solution (henceforth referred to as the basic state). The coefficient of dispersion can be obtained either through calculation of the linear dispersion relation or through a Jordan chain argument. We use the Jordan chain approach within the asymptotics as it arises naturally within the theory. In some contexts the use of the Jordan chain approach may be easier, where the linear dispersion relation may be difficult to compute, although since the dispersive term is linear the coefficient may be achieved from the dispersion relation with equal validity.

To illustrate how such an approach can be applied, we present two examples of where the two-way Boussinesq equation emerges. The first discusses the theory in the context of a set of coupled Nonlinear Schrödinger equations, and in fact presents the first such reduction using any approach from this system to the two-way Boussinesq equation that the author is aware of. The second example used to illustrate the result of this paper is a two-layered stratified shallow water system, and although such reductions exist in the literature, the primary aim here is to show that the coefficients may be determined simply using rudimentary linear algebra.

The structure of the paper follows. We introduce the necessary properties of the basic state for the theory, including a discussion of the linear operator, the conservation laws and the emergent Jordan chain theory. This is followed by a summary of the asymptotic analysis arising from the phase dynamics that leads to (1.7). A brief discussion of this system's degeneracy are presented before the reconstruction of the modulation argument is undertaken leading to (1.12). To apply the theory, we use the examples of a coupled NLS system and stratified shallow water hydrodynamics to show how the two-way Boussinesq equation can be derived in these contexts using the ideas built up in this paper. Finally, concluding remarks and areas for future study are discussed.

2. Governing equations, linearisation and conserved quantities

For the purposes of this analysis, (1.4) is considered our governing equation. The matrices are assumed skew-symmetric so that $\mathbf{M}^T = -\mathbf{M}$, $\mathbf{J}^T = -\mathbf{J}$, and $\mathbf{J}$ is assumed invertible for simplicity. In the cases where $\mathbf{J}$ is non-invertible, we would have to make the assumption instead that none of the resulting Jordan chain elements appearing later lie in its kernel, a scenario we do not consider here. Now assume the existence of a two-phase solution to (1.4) of the form

$$Z = \widehat{Z}(\boldsymbol{\theta};\mathbf{k},\boldsymbol{\omega}),$$

and when substituted into (1.4) shows that $\widehat{Z}$ satisfies the PDE

$$\big(\omega_1\mathbf{M} + k_1\mathbf{J}\big)\widehat{Z}_{\theta_1} + \big(\omega_2\mathbf{M} + k_2\mathbf{J}\big)\widehat{Z}_{\theta_2} = \nabla S(\widehat{Z}). \tag{2.1}$$

Such a solution is an example of a multiparameter relative equilibrium (referred to herein as the basic state) when continuous symmetries are present [16], which we assume for the purposes of this paper is the case.

Linearisation of (1.4) about this solution gives the linear operator

$$\mathbf{L}\mathbf{v} = \mathrm{D}^2 S(\widehat{Z})\mathbf{v} - \big(\omega_1\mathbf{M} + k_1\mathbf{J}\big)\mathbf{v}_{\theta_1} - \big(\omega_2\mathbf{M} + k_2\mathbf{J}\big)\mathbf{v}_{\theta_2}\,, \tag{2.2}$$

which is formally self adjoint with respect to the inner product

$$\langle\!\langle \cdot,\cdot \rangle\!\rangle = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi}\!\!\int_0^{2\pi} \langle \cdot,\cdot\rangle \,\mathrm{d}\theta_1 \mathrm{d}\theta_2\ , \tag{2.3}$$

for $2\pi\times 2\pi$ periodic solutions, where $\langle\cdot,\cdot\rangle$ is the standard inner product. For systems with purely affine symmetry (as is the case for the second example of this paper) the averaging is dropped.

The modulation analysis takes advantage of the equations for the derivatives of $\widehat{Z}$, with respect to its phases, wavenumbers, and frequencies. The following expressions are of particular note:

$$\mathbf{L}\widehat{Z}_{\theta_i} = 0, \tag{2.4a}$$

$$\mathbf{L}\widehat{Z}_{k_i} = \mathbf{J}\widehat{Z}_{\theta_i}, \tag{2.4b}$$

$$\mathbf{L}\widehat{Z}_{\omega_i} = \mathbf{M}\widehat{Z}_{\theta_i}. \tag{2.4c}$$

The first equation highlights the zero eigenvalue of $\mathbf{L}$ is not simple, and so the assumption on its kernel is made to be

$$\ker(\mathbf{L}) = \mathrm{span}\{\widehat{Z}_{\theta_1},\widehat{Z}_{\theta_2}\} \tag{2.5}$$

and that it is no larger. This, and the self-adjointness of $\mathbf{L}$, imply that any system of the form

$$\mathbf{L}W = F \quad \text{is solvable when} \quad \langle\!\langle \widehat{Z}_{\theta_1}, F \rangle\!\rangle = \langle\!\langle \widehat{Z}_{\theta_2}, F \rangle\!\rangle = 0\,. \tag{2.6}$$

This forms the primary mechanism for which the conditions (1.8), (1.10) emerge as well as how the vector coefficients appearing in (1.12) will eventually be generated. The remaining equations (2.4b) and (2.4c) hint at the importance of Jordan chains within the analysis. In this paper, only the one involving $\mathbf{J}$ will pivotal to the resulting PDE, but in other contexts (such as in further time degeneracy or in the dual setting) the other chain may play a more substantial role in the analysis.

2.1. Jordan chain theory

The analysis which leads to (1.12) relies on a short Jordan chain argument in relation to the solvability of certain terms. This section deals with the necessary theory that justifies this.

Under the assumption (2.5), the zero eigenvalue of $\mathbf{L}$ has geometric multiplicity two. This fact determines that there must be two Jordan blocks:

$$\left.\begin{array}{rcl} \mathbf{L}\widehat{Z}_{\theta_1} &=& 0 \\ \mathbf{L}\widehat{Z}_{k_1} &=& \mathbf{J}\widehat{Z}_{\theta_1} \end{array}\right\} \quad \text{and} \quad \left\{\begin{array}{rcl} \mathbf{L}\widehat{Z}_{\theta_2} &=& 0 \\ \mathbf{L}\widehat{Z}_{k_2} &=& \mathbf{J}\widehat{Z}_{\theta_2} \end{array}\right. .$$

The modulation analysis leading to the linear multiphase Whitham system discussed in section 3 assumes these chains are no longer than this. In the case that they are, the linear Whitham equations are degenerate as the matrix $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ becomes singular. A consequence of this is that the system

$$\mathbf{L}\xi_5 = \zeta_1 \mathbf{J}\widehat{Z}_{k_1} + \zeta_2 \mathbf{J}\widehat{Z}_{k_2} \tag{2.7}$$

becomes solvable, where ζ_1, ζ_2 are the elements of the zero eigenvector of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$, $\boldsymbol{\zeta}$.

A longer Jordan chain now exists through combining elements of the above two blocks. The existence of a solution to (2.7) defines ξ_5. As all symplectic Jordan chains contain an even number of elements due to the problem possessing an even characteristic polynomial [22, chapter 3]. This is seen when one considers the characteristic polynomial for this problem, $\Delta(\lambda)$:

$$\begin{aligned} \Delta(\lambda) =& \det\big[\mathbf{L} - \lambda\mathbf{J}\big] = \det\big[(\mathbf{L} - \lambda\mathbf{J})^T\big] \\ =& \det\big[\mathbf{L}^T - \lambda\mathbf{J}^T\big] = \det\big[\mathbf{L} + \lambda\mathbf{J}\big] = \Delta(-\lambda) \end{aligned}$$

Therefore, the characteristic polynomial is even in λ and so when the zero eigenvalue occurs it does so with even multiplicity. As a consequence, the existence of ξ_5 guarantees the next element ξ_6 also exists and satisfies

$$\mathbf{L}\xi_6 = \mathbf{J}\xi_5 \,. \tag{2.8}$$

A $2 \oplus 4$ structure can in fact be formulated through a change of basis, resulting in two Jordan chains. These take the form

$$\begin{cases} \mathbf{L}(\zeta_1 \widehat{Z}_{\theta_2} - \zeta_2 \widehat{Z}_{\theta_1}) &= 0 \\ \mathbf{L}(\zeta_1 \widehat{Z}_{k_2} - \zeta_2 \widehat{Z}_{k_1}) &= \mathbf{J}(\zeta_1 \widehat{Z}_{\theta_2} - \zeta_2 \widehat{Z}_{\theta_1}) \end{cases}$$

$$\text{and} \begin{cases} \mathbf{L}(\zeta_1 \widehat{Z}_{\theta_1} + \zeta_2 \widehat{Z}_{\theta_2}) &= 0 \\ \mathbf{L}(\zeta_1 \widehat{Z}_{k_1} + \zeta_2 \widehat{Z}_{k_2}) &= \mathbf{J}(\zeta_1 \widehat{Z}_{\theta_1} + \zeta_2 \widehat{Z}_{\theta_2}) \\ \mathbf{L}\xi_5 &= \mathbf{J}(\zeta_1 \widehat{Z}_{k_1} + \zeta_2 \widehat{Z}_{k_2}) \\ \mathbf{L}\xi_6 &= \mathbf{J}\xi_5 \,. \end{cases}$$

The top chain contains only two elements, owing to the fact that zero is strictly a simple eigenvalue of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$. It is assumed that the right-hand chain terminates at four. Define now

$$\mathbf{K} = \begin{pmatrix} \mathscr{K}_1 \\ \mathscr{K}_2 \end{pmatrix} := - \begin{pmatrix} \langle\!\langle \widehat{Z}_{\theta_1}, \mathbf{J}\xi_6 \rangle\!\rangle \\ \langle\!\langle \widehat{Z}_{\theta_2}, \mathbf{J}\xi_6 \rangle\!\rangle \end{pmatrix} . \tag{2.9}$$

The non-solvability of $\mathbf{L}\xi_7 = \mathbf{J}\xi_6$ then assures that $\|\mathbf{K}\|^2 = \mathscr{K}_1^2 + \mathscr{K}_2^2 > 0$, and thus the presence of dispersion in the projected system.

The role of the Jordan chain in the modulation theory is to form the dispersive part of the reduction. In the case of two independent chains of length two, it will become clear that these chains form the first order dispersive terms in X. When the chain becomes longer in the way outlined above, the asymptotics will highlight that it is precisely the Jordan chain of length four that generates the coefficient of dispersion through $\mathbf{K}$ for the final PDE.

2.2. Conservation laws

With the Lagrangian in canonical form (1.3) the conservation laws now possess a geometric form, since the presence of the skew-symmetric matrices in their formulation associates them to the symplectic structure [23]. The wave action vector evaluated along the 2 phase wavetrain from this formulation can be found as

$$\mathbf{A}(\mathbf{k}, \boldsymbol{\omega}) = \begin{pmatrix} \mathscr{A}_1 \\ \mathscr{A}_2 \end{pmatrix} := \frac{1}{2} \begin{pmatrix} \langle\!\langle \mathbf{M}\widehat{Z}_{\theta_1}, \widehat{Z} \rangle\!\rangle \\ \langle\!\langle \mathbf{M}\widehat{Z}_{\theta_2}, \widehat{Z} \rangle\!\rangle \end{pmatrix} ,$$

as well as the associated flux vector

$$\mathbf{B}(\mathbf{k}, \boldsymbol{\omega}) = \begin{pmatrix} \mathscr{B}_1 \\ \mathscr{B}_2 \end{pmatrix} := \frac{1}{2} \begin{pmatrix} \langle\!\langle \mathbf{J}\widehat{Z}_{\theta_1}, \widehat{Z} \rangle\!\rangle \\ \langle\!\langle \mathbf{J}\widehat{Z}_{\theta_2}, \widehat{Z} \rangle\!\rangle \end{pmatrix} .$$

in the periodic case. The affine case is almost identical but without the factors of $\frac{1}{2}$. The periodic case can alternatively be obtained through the k and ω derivatives of the Lagrangian (1.3) averaged over the two-phase solution:

$$\mathscr{L}(\mathbf{k},\boldsymbol{\omega}) = \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi}\left[\frac{1}{2}\sum_{j=1}^{2}\left[\langle \widehat{Z},\omega_j\mathbf{M}\widehat{Z}_{\theta_j}+k_j\mathbf{J}\widehat{Z}_{\theta_j}\rangle\right]-S(\widehat{Z})\right]\mathrm{d}\theta_1\mathrm{d}\theta_2\,.$$

By these definitions, we then have the following directional derivatives:

$$\mathrm{D}_{\mathbf{k}}\mathbf{A} = \begin{pmatrix}\partial_{k_1}\mathscr{A}_1 & \partial_{k_2}\mathscr{A}_1\\ \partial_{k_1}\mathscr{A}_2 & \partial_{k_2}\mathscr{A}_2\end{pmatrix} = \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}^T,$$

$$\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A} = \begin{pmatrix}\partial_{\omega_1}\mathscr{A}_1 & \partial_{\omega_2}\mathscr{A}_1\\ \partial_{\omega_1}\mathscr{A}_2 & \partial_{\omega_2}\mathscr{A}_2\end{pmatrix},\quad \mathrm{D}_{\mathbf{k}}\mathbf{B} = \begin{pmatrix}\partial_{k_1}\mathscr{B}_1 & \partial_{k_2}\mathscr{B}_1\\ \partial_{k_1}\mathscr{B}_2 & \partial_{k_2}\mathscr{B}_2\end{pmatrix},$$

$$\mathrm{D}^2_{\mathbf{k}}\mathbf{B} = \left(\begin{array}{cc|cc}\partial_{k_1k_1}\mathscr{B}_1 & \partial_{k_2k_1}\mathscr{B}_1 & \partial_{k_1k_2}\mathscr{B}_1 & \partial_{k_2k_2}\mathscr{B}_1\\ \partial_{k_1k_1}\mathscr{B}_2 & \partial_{k_2k_1}\mathscr{B}_2 & \partial_{k_1k_2}\mathscr{B}_2 & \partial_{k_2k_2}\mathscr{B}_2\end{array}\right).$$

The entries of these tensors are related to solutions via

$$\partial_{k_j}\mathscr{A}_i = \langle\!\langle \mathbf{M}\widehat{Z}_{\theta_i},\widehat{Z}_{k_j}\rangle\!\rangle, \tag{2.10a}$$

$$\partial_{\omega_j}\mathscr{A}_i = \langle\!\langle \mathbf{M}\widehat{Z}_{\theta_i},\widehat{Z}_{\omega_j}\rangle\!\rangle, \tag{2.10b}$$

$$\partial_{k_j}\mathscr{B}_i = \langle\!\langle \mathbf{J}\widehat{Z}_{\theta_i},\widehat{Z}_{k_j}\rangle\!\rangle, \tag{2.10c}$$

$$\partial_{k_jk_m}\mathscr{B}_i = \langle\!\langle \mathbf{J}\widehat{Z}_{\theta_ik_m},\widehat{Z}_{k_j}\rangle\!\rangle + \langle\!\langle \mathbf{J}\widehat{Z}_{\theta_i},\widehat{Z}_{k_jk_m}\rangle\!\rangle. \tag{2.10d}$$

Notice that

$$\partial_{k_i}\mathscr{B}_j = \langle\!\langle \mathbf{J}\widehat{Z}_{\theta_j},\widehat{Z}_{k_i}\rangle\!\rangle = \langle\!\langle \mathbf{L}\widehat{Z}_{k_j},\widehat{Z}_{k_i}\rangle\!\rangle = \langle\!\langle \widehat{Z}_{k_j},\mathbf{L}\widehat{Z}_{k_i}\rangle\!\rangle = \langle\!\langle \widehat{Z}_{k_j},\mathbf{J}\widehat{Z}_{\theta_i}\rangle\!\rangle = \partial_{k_j}\mathscr{B}_i \tag{2.11}$$

as well as

$$\partial_{k_j}\mathscr{A}_i = \langle\!\langle \mathbf{M}\widehat{Z}_{\theta_i},\widehat{Z}_{k_i}\rangle\!\rangle = \langle\!\langle \widehat{Z}_{\omega_i},\mathbf{J}\widehat{Z}_{k_j}\rangle\!\rangle = \partial_{\omega_i}\mathscr{B}_j\,.$$

We say that a conservation law is critical in the multiphase case if a zero determinant condition, such as (1.8), is attained with respect to one of the parameter vectors $\mathbf{k}$ or $\boldsymbol{\omega}$. The details of the modulation analysis when this is the case are presented in section 4.

The criticality condition (1.8) is primarily considered in a mathematical way in this paper, with the fact that an abstract Lagrangian is considered suggests no physical connotations are necessary. This is not to say that this condition has no physical relevance, as it can be shown to emerge in several contexts. In fluid mechanics, the condition (1.8) corresponds to one of the characteristic speeds of the uniform flow solution vanishes [14] and forms the boundary between stability and instability [20]. In other systems, such as the coupled NLS system considered later, this criticality condition again forms a stability boundary [16]. Therefore the criterion (1.8) may be viewed as a marginal stability curve for the system considered. This connection between mathematical and physical criticality is further discussed in [1].

3. Linear Whitham equations

We begin our discussion with the simplest case of multiphase modulation, which leads to the linear multiphase Whitham equations. To obtain them, we construct the ansatz as

$$Z = \widehat{Z}(\boldsymbol{\theta} + \boldsymbol{\phi}, \mathbf{k} + \varepsilon\mathbf{q}, \boldsymbol{\omega} + \varepsilon\boldsymbol{\Omega}) + \varepsilon^2 W(\boldsymbol{\theta}, X, T). \tag{3.1}$$

In the above, we define $X = \varepsilon x,\ T = \varepsilon t$ and the perturbations are related through the following

$$\boldsymbol{\phi}_X = \mathbf{q}, \quad \boldsymbol{\phi}_T = \boldsymbol{\Omega} \quad \Rightarrow \quad \mathbf{q}_T = \boldsymbol{\Omega}_X. \tag{3.2}$$

The fact that these functions are related this way can either be viewed as a definition for $\mathbf{q}$, $\boldsymbol{\Omega}$ or as a relation between them. We substitute the ansatz into (1.4) and undertake a Taylor expansion around the $\varepsilon = 0$ state. Leading order obtains (2.1), whereas at first order we obtain

$$\sum_{i=1}^{2} \left[q_i \mathbf{L}\widehat{Z}_{k_i} + \Omega_i \mathbf{L}\widehat{Z}_{\omega_i} \right] = \sum_{i=1}^{2} \left[(\phi_i)_X \mathbf{J}\widehat{Z}_{\theta_i} + (\phi_i)_T \mathbf{M}\widehat{Z}_{\theta_i} \right],$$

which is satisfied through the phase consistency conditions (3.2) along with (2.4b), (2.4c). The next order once simplified gives that

$$\mathbf{L}W = \sum_{i=1}^{2} \left[(q_i)_X \mathbf{J}\widehat{Z}_{k_i} + (\Omega_i)_X \mathbf{J}\widehat{Z}_{\omega_i} + (q_i)_T \mathbf{M}\widehat{Z}_{k_i} + (\Omega_i)_T \mathbf{M}\widehat{Z}_{\omega_i} \right].$$

Imposing the solvability condition (2.6) and using (2.10a - 2.10c) reveals that the modulational functions must satisfy

$$\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\Omega}_T + \mathrm{D}_{\mathbf{k}}\mathbf{A}\mathbf{q}_T + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}\boldsymbol{\Omega}_X + \mathrm{D}_{\mathbf{k}}\mathbf{B}\mathbf{q}_X = \mathbf{0}. \tag{3.3}$$

This abstract approach generalises results such as [19] to general Lagrangians with symmetries, albeit only in the linear case. It should be noted that, although these appear to be the conservation laws linearised about fixed wavenumbers and frequencies, it is unlikely that this reduction will satisfy the original system's conservation laws. It will only do so approximately, since the higher order terms in ε that would normally contribute to the true conservation law are in essence discarded. This leads to the notion of approximate conservation laws [21].

Generically the above system is nondegenerate (that is, all the terms in the above are nonvanishing) in $(\mathbf{k},\boldsymbol{\omega})$-space, but in some contexts there are curves in this space for which one or more of the Jacobians is singular. These singularities then give rise to nonlinear reductions of the Euler-Lagrange equations (1.4) along these curves, leading to dispersion. The criterion for their emergence can be formulated using the degeneracies of (3.3). This will form the remainder of the discussion in this section.

3.1. Zero eigenvalues of the Whitham equations

By defining $\mathbf{Q} = (\mathbf{q},\boldsymbol{\Omega})^T$, one may rewrite (3.3) as the matrix system

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathrm{D}_{\boldsymbol{\omega}}\mathbf{A} \end{pmatrix} \mathbf{Q}_T + \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathrm{D}_{\mathbf{k}}\mathbf{B} & \mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B} \end{pmatrix} \mathbf{Q}_X = 0. \tag{3.4}$$

We can invert the first matrix readily providing $\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}$ is nonsingular to write the linear Whitham equations as

$$\mathbf{Q}_T + \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ (\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A})^{-1}\mathrm{D}_{\mathbf{k}}\mathbf{B} & (\mathrm{D}_{\omega}\mathbf{A})^{-1}(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}) \end{pmatrix} \mathbf{Q}_X = 0. \tag{3.5}$$

The zero eigenvalues of the matrix are then of interest, since these determine the characteristics of this system. Computation of the characteristic polynomial for this matrix gives that the eigenvalues satisfy

$$a_4\lambda^4 + a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0\,, \tag{3.6}$$

with

$$\begin{aligned} a_4 &= \det\big[\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\big]\,, \\ a_3 &= -\,\mathrm{Trace}\big[(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})^\sharp \mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\big]\,, \\ a_2 &= \det\big[\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}\big] + \mathrm{Trace}\big[\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}^\sharp \mathrm{D}_{\mathbf{k}}\mathbf{B}\big]\,, \\ a_1 &= -\,\mathrm{Trace}\big[(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})^\sharp \mathrm{D}_{\mathbf{k}}\mathbf{B}\big]\,, \\ a_0 &= \det\big[\mathrm{D}_{\mathbf{k}}\mathbf{B}\big]\,, \end{aligned}$$

where $\sharp$ denotes the cofactor matrix, defined as

$$\mathbf{P}^\sharp = \det\big[\mathbf{P}\big]\mathbf{P}^{-1}.$$

From this, it is clear that this matrix has a simple zero eigenvalue (and thus the linear Whitham equations have a zero characteristic) when

$$\det\big[\mathrm{D}_{\mathbf{k}}\mathbf{B}\big] = 0, \tag{3.7}$$

since we assume that $\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}$ is invertible. This will be referred to as primary criticality as it will be the first condition met in the reduction, but also because it is the main criteria responsible for both the generation of dispersion and nonlinearity from the phase dynamics. It has previously been demonstrated that when this criticality is met the modulation approach leads to the KdV equation [1]. This formulation of the linear Whitham equations provides an alternate definition of the eigenvector associated with the zero eigenvalue (henceforth referred to as the zero eigenvector) $\boldsymbol{\zeta}$ in the following way:

$$\begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ (\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A})^{-1}\mathrm{D}_{\mathbf{k}}\mathbf{B} & (\mathrm{D}_{\omega}\mathbf{A})^{-1}(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}) \end{pmatrix} \begin{pmatrix} \boldsymbol{\zeta} \\ \mathbf{0} \end{pmatrix} = \mathbf{0}\,.$$

The vector $\boldsymbol{\zeta}$ can be scaled freely at this stage as the above problem is linear, and such a scaling does not affect the linear coefficients emerging within the theory of this paper (since the scaling can be cancelled out in the final PDE). A choice of scaling will affect the final coefficient of nonlinearity, which is typical of nonlinear PDEs. Therefore no scaling is imposed on $\boldsymbol{\zeta}$ (such as normalisation) in this paper, nor will it be on any of the other generalised eigenvector appearing within the theory.

The main result of this paper occurs when the zero eigenvalue arising from (3.6) has algebraic multiplicity two. By inspection, this occurs when

$$\mathrm{Trace}\Big[(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})^{\sharp}\mathrm{D}_{\mathbf{k}}\mathbf{B}\Big] = 0 \tag{3.8}$$

This forms the secondary criticality condition, as it will be the second condition to arise in the reduction but necessarily requires the first, (3.7), to hold in order to be relevant. To prevent the zero having higher algebraic multiplicity we also additionally impose that

$$\det\Big[\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}\Big] + \mathrm{Trace}\Big[\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}^{\sharp}\mathrm{D}_{\mathbf{k}}\mathbf{B}\Big] \neq 0.$$

This condition prevents the new time derivative term arising from the phase dynamics from having a zero coefficient, and thus assures that the derived two-way Boussinesq equation is the dominant balance.

How does this double zero condition arise within the modulation theory? Since $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ is a 2×2 matrix with (assumed) simple zero eigenvalue, the presence of a zero with algebraic multiplicity 2 in (3.5) suggests a generalised eigenvector problem of the form emerges:

$$\begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ (\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A})^{-1}\mathrm{D}_{\mathbf{k}}\mathbf{B} & (\mathrm{D}_{\omega}\mathbf{A})^{-1}(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}) \end{pmatrix} \begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\delta} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\zeta} \\ \mathbf{0} \end{pmatrix}\,.$$

Solving this, we find that $\boldsymbol{\delta} = -\boldsymbol{\zeta}$ and $\boldsymbol{\gamma}$ must satisfy

$$(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta} = \mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\gamma}. \tag{3.9}$$

Multiplying on the left by $\boldsymbol{\zeta}$ leads to (3.8) after some manipulation. Although not immediately obvious, the implication is that an additional set of modulational functions are needed in the ansatz. This is to facilitate the generation of the suitable term on the right hand side of (3.9) to solve the term on the left. This incorporation of an additional set of functions into the derivation is also supported by the ideas presented in the derivation of the two-way Boussinesq equation equation for stratified flows [14].

4. Modulation leading to the Boussinesq equation

In the presence of the singularities (3.7) and (3.8), we must use an ansatz of the form (1.11) and the slow variables are scaled as $X = \varepsilon x,\ T = \varepsilon^2 t$. In this modulational analysis, we also consider the function W as a simple asymptotic expansion,

$$W = W_0 + \varepsilon W_1 + \varepsilon^2 W_2 + \dots .$$

Expressions that automatically cancel are ignored within this analysis to focus on the important terms to streamline the discussion. Below is a summary of the calculations leading to (1.12) and by extension (1.13).

4.1. Orders 1, ε and ε^2

The leading order recovers the equation for the basic state (2.1), whereas the first order in ε reads

$$\phi_1 \mathbf{L}\widehat{Z}_{\theta_1} + \phi_2 \mathbf{L}\widehat{Z}_{\theta_2} = 0.$$

This is automatically satisfied since the θ derivatives lie in the nullspace of the linear operator. At second order in ε we have

$$q_1 \mathbf{L}\widehat{Z}_{k_1} - (\phi_1)_X \mathbf{J}\widehat{Z}_{\theta_1} + q_2 \mathbf{L}\widehat{Z}_{k_2} - (\phi_2)_X \mathbf{J}\widehat{Z}_{\theta_2} = 0.$$

which is true, by (2.4b) and the phase consistency condition (3.2).

4.2. Third order

The non-cancelling terms at this order, including those that recover the relation $\boldsymbol{\phi}_T = \boldsymbol{\Omega}$ in (3.2), are given by

$$\mathbf{L}W_0 = (q_1)_X \mathbf{J}\widehat{Z}_{k_1} + (q_2)_X \mathbf{J}\widehat{Z}_{k_2}.$$

These are solvable when

$$\begin{pmatrix} \langle\!\langle \widehat{Z}_{\theta_1}, \mathbf{J}\widehat{Z}_{k_1} \rangle\!\rangle & \langle\!\langle \widehat{Z}_{\theta_1}, \mathbf{J}\widehat{Z}_{k_2} \rangle\!\rangle \\ \langle\!\langle \widehat{Z}_{\theta_2}, \mathbf{J}\widehat{Z}_{k_1} \rangle\!\rangle & \langle\!\langle \widehat{Z}_{\theta_2}, \mathbf{J}\widehat{Z}_{k_2} \rangle\!\rangle \end{pmatrix} \mathbf{q}_X = -\mathrm{D}_{\mathbf{k}}\mathbf{B}\mathbf{q}_X = \mathbf{0}\,, \tag{4.1}$$

which in turn is nontrivially solvable when $\det[\mathrm{D}_{\mathbf{k}}\mathbf{B}] = 0$. This allows us to define a zero eigenvector $\boldsymbol{\zeta} = (\zeta_1, \zeta_2)^T$, which means there exists some function U such that

$$\mathbf{q} = U(X,T)\zeta. \tag{4.2}$$

In general U also depends on ε, however only the leading order term is necessary for the analysis in this paper and so we take $U(X,T) = U(X,T;\varepsilon)|_{\varepsilon=0}$. Excluding the further expansion of U in ε has no effect on the asymptotics up to the order considered. The solution at this order is therefore

$$W_0 = \alpha_1 \widehat{Z}_{\theta_1} + \alpha_2 \widehat{Z}_{\theta_2} + U_X \xi_5,$$

and ξ_5 satisfies (2.7). The functions α_1, α_2 are arbitrary functions of the slow space and time variables, and will go on to form the inhomogeneity in the final vector equation.

4.3. Fourth order

The terms that persevere at fourth order are

$$\mathbf{L}(W_1 - \widetilde{W}_1) = \zeta_1 U_T(\mathbf{M}\widehat{Z}_{k_1} + \mathbf{J}\widehat{Z}_{\omega_1}) + \zeta_2 U_T(\mathbf{M}\widehat{Z}_{k_2} + \mathbf{J}\widehat{Z}_{\omega_2}) + U_{XX}\mathbf{J}\xi_5 \\ + (r_1)_X \mathbf{J}\widehat{Z}_{k_1} + (r_2)_X \mathbf{J}\widehat{Z}_{k_2}.$$

The function $\widetilde{W}_1$ contains terms for which the solution may already be computed. Explicitly, it is given by

$$\begin{aligned} \widetilde{W}_1 \;\; = \;\; & (\alpha_1)_X \widehat{Z}_{k_1} + (\alpha_2)_X \widehat{Z}_{k_2} + \zeta_1 U \alpha_1 \widehat{Z}_{\theta_1\theta_1} + (\zeta_2 U \alpha_1 + \zeta_1 U \alpha_2)\widehat{Z}_{\theta_1\theta_2} \\ & + \zeta_2 U \alpha_2 \widehat{Z}_{\theta_2\theta_2} + \phi_1 U_X (\xi_5)_{\theta_1} + \phi_2 U_X (\xi_5)_{\theta_2} \, . \end{aligned} \tag{4.3}$$

Collecting these terms in this expression is to highlight that these are not important at this order, and instead that the remaining terms are of interest. The second derivative term is solvable since the zero eigenvalue of $\mathbf{L}$ is even, as discussed in §2.1. The rest of the terms are solvable when

$$(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta} U_T + \mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{r}_X = 0. \tag{4.4}$$

This is precisely the problem that appears in (3.9), and is solvable providing that

$$\boldsymbol{r}_X = -\boldsymbol{\gamma} U_T. \tag{4.5}$$

This allows us to define ϱ with

$$\mathbf{L}\varrho = \zeta_1(\mathbf{M}\widehat{Z}_{k_1} + \mathbf{J}\widehat{Z}_{\omega_1}) + \zeta_2(\mathbf{M}\widehat{Z}_{k_2} + \mathbf{J}\widehat{Z}_{\omega_2}) - \gamma_1 \mathbf{J}\widehat{Z}_{k_1} - \gamma_2 \widehat{Z}_{k_2}.$$

Now everything is related to a single modulational function U. Overall, this allows us to obtain a solution for W_1 as

$$W_1 = \widetilde{W}_1 + \beta_1(X,T)\widehat{Z}_{\theta_1} + \beta_2(X,T)\widehat{Z}_{\theta_2} + U_{XX}\xi_4 + U_T\varrho. \tag{4.6}$$

4.4. Fifth order

The remaining terms, upon using that $\boldsymbol{r}_T = \boldsymbol{\tau}_X$ at this order are

$$\begin{aligned}
\mathbf{L}\widetilde{W}_2 =&(r_1)_T(\mathbf{M}\widehat{Z}_{k_1} + \mathbf{J}\widehat{Z}_{\omega_1}) + (r_2)_T(\mathbf{M}\widehat{Z}_{k_2} + \mathbf{J}\widehat{Z}_{\omega_2}) \\
&+ (\Omega_1)_T\mathbf{M}\widehat{Z}_{\omega_1} + (\Omega_2)_T\mathbf{M}\widehat{Z}_{\omega_2} \\
&+ \zeta_1 UU_X\big(D^3S(\widehat{Z})(\widehat{Z}_{k_1},\xi_5) - \mathbf{J}(\xi_5)_{\theta_1} - \zeta_2\mathbf{J}\widehat{Z}_{k_1k_2} - \zeta_1\mathbf{J}\widehat{Z}_{k_1k_1}\big) \\
&+ \zeta_2 UU_X\big(D^3S(\widehat{Z})(\widehat{Z}_{k_2},\xi_5) - \mathbf{J}(\xi_5)_{\theta_2} - \zeta_2\mathbf{J}\widehat{Z}_{k_2k_2} - \zeta_1\mathbf{J}\widehat{Z}_{k_1k_2}\big) \\
&+ U_{XT}(\mathbf{J}\varrho + \mathbf{M}\xi_5) + U_{XXX}\mathbf{J}\xi_4 + (\alpha_1)_{XX}\mathbf{J}\widehat{Z}_{k_1} + (\alpha_2)_{XX}\mathbf{J}\widehat{Z}_{k_2}.
\end{aligned} \tag{4.7}$$

Terms that are solvable have been absorbed into the linear operator in the form of $\widetilde{W}_2$, and are not important since the analysis terminates at this order. All that remains is to determine the solvability condition for the equation at this order, which generates the two-way Boussinesq equation.

The first set of terms, multiplied by $(r_i)_T$, have appeared at the previous order and simply result in the term

$$(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{r}_T.$$

The next set of terms, involving $(\Omega_i)_T$, generate the tensor

$$\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\Omega}_T\,,$$

by using (2.10b). The U_{XT} term vanishes again from the fact that the zero eigenvalue of $\mathbf{L}$ is even. The tensor acting on U_{XXX} is determined using the definition (2.9) to give $\mathbf{K}U_{XXX}$. Solvability of the $\boldsymbol{\alpha}_{XX}$ terms is almost identical to the calculation appearing at third order in (4.1) and gives the term $\big[\mathrm{D}_{\mathbf{k}}\mathbf{B}\big]\boldsymbol{\alpha}_{XX}$. The final tensor to determine is the one multiplying the quadratic term UU_X. Notice that

$$\begin{aligned}
\langle\!\langle \widehat{Z}_{\theta_i},&\mathrm{D}^3S(\widehat{Z})(\widehat{Z}_{k_j},\xi_5) - \mathbf{J}(\xi_5)_{\theta_j} - \zeta_1\mathbf{J}\widehat{Z}_{k_1k_j} - \zeta_2\mathbf{J}\widehat{Z}_{k_jk_2}\rangle\!\rangle \\
&= \langle\!\langle \mathrm{D}^3S(\widehat{Z})(\widehat{Z}_{k_j},\widehat{Z}_{\theta_i}) - \mathbf{J}\widehat{Z}_{\theta_i\theta_j},\xi_5\rangle\!\rangle - \langle\!\langle \widehat{Z}_{\theta_i},\zeta_1\mathbf{J}\widehat{Z}_{k_1k_j} + \zeta_2\mathbf{J}\widehat{Z}_{k_jk_2}\rangle\!\rangle\,, \\
&= -\,\langle\!\langle \widehat{Z}_{\theta_ik_j},\mathbf{L}\xi_5\rangle\!\rangle - \langle\!\langle \widehat{Z}_{\theta_i},\zeta_1\mathbf{J}\widehat{Z}_{k_1k_j} + \zeta_2\mathbf{J}\widehat{Z}_{k_jk_2}\rangle\!\rangle\,, \\
&= -\,\langle\!\langle \widehat{Z}_{\theta_ik_j},\zeta_1\mathbf{J}\widehat{Z}_{k_1} + \zeta_2\mathbf{J}\widehat{Z}_{k_2}\rangle\!\rangle - \langle\!\langle \widehat{Z}_{\theta_i},\zeta_1\mathbf{J}\widehat{Z}_{k_1k_j} + \zeta_2\mathbf{J}\widehat{Z}_{k_jk_2}\rangle\!\rangle\,, \\
&= \zeta_1\partial_{k_1k_j}\mathscr{B}_i + \zeta_2\partial_{k_jk_2}\mathscr{B}_i.
\end{aligned}$$

where we have used that $\mathbf{L}\widehat{Z}_{\theta_ik_j} = \mathbf{J}\widehat{Z}_{\theta_i\theta_j} - \mathrm{D}^3S(\widehat{Z})(\widehat{Z}_{k_j},\widehat{Z}_{\theta_i})$, seen by differentiating (2.4a) with respect to k_j. Therefore, the tensor acting on the nonlinearity takes the form

$$\begin{pmatrix} \zeta_1(\zeta_1\partial_{k_1k_1}\mathscr{B}_1+\zeta_2\partial_{k_1k_2}\mathscr{B}_1)+\zeta_2(\zeta_1\partial_{k_1k_2}\mathscr{B}_1+\zeta_2\partial_{k_2k_2}\mathscr{B}_1) \\ \zeta_1(\zeta_1\partial_{k_1k_1}\mathscr{B}_2+\zeta_2\partial_{k_1k_2}\mathscr{B}_2)+\zeta_2(\zeta_1\partial_{k_1k_2}\mathscr{B}_2+\zeta_2\partial_{k_2k_2}\mathscr{B}_1) \end{pmatrix} = \mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta})\,.$$

Combining all of the above results gives the matrix system

$$\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\Omega}_T+(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\mathbf{k}}\mathbf{A}^T)\boldsymbol{r}_T+\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta})UU_X+\mathbf{K}U_{XXX}+\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\alpha}_{XX}=0. \tag{4.8}$$

Differentiation with respect to X and using (4.5), as well as $\boldsymbol{\zeta}U_{TT}=\mathbf{q}_{TT}=\boldsymbol{\Omega}_{XT}$, gives

$$\Big(\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\zeta}-(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\gamma}\Big)U_{TT} + \Big(\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta})UU_X+\mathbf{K}U_{XXX}\Big)_X+\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\alpha}_{XXX}=0. \tag{4.9}$$

This is the two-way Boussinesq equation given in (1.12), and can be formed into a scalar equation upon projection in the direction of the kernel of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ by multiplying on the left by $\boldsymbol{\zeta}$, giving

$$AU_{TT}+\big(BUU_X+KU_{XXX}\big)_X=0\,, \tag{4.10}$$

with

$$A=\boldsymbol{\zeta}^T\Big(\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\zeta}-(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\gamma}\Big),\quad B=\boldsymbol{\zeta}^T\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta}),\quad K=\boldsymbol{\zeta}^T\mathbf{K}.$$

The other projection, involving the other eigenvector of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$, leading to an equation from which $\boldsymbol{\alpha}$ may be determined. This equation isn't necessary for our analysis, since it is only required if one were to continue the analysis to higher orders of ε which is not undertaken here. Also required to proceed to higher orders of ε would be the inclusion of the previously neglected ε terms in the expansion of U, as these would appear nontrivially at higher orders. Finally, the conservation laws of the original system are only approximately satisfied by this reduction. The inclusion of further ε orders would be required to alleviate this, much like the linear Whitham equations.

5. Example 1: coupled NLS model

The primary example of the paper is to demonstrate how one may reduce the coupled nonlinear Schrödinger equations (CNLS) to the two-way Boussinesq equation. This presents the first such reduction using any theoretical approach, and so provides a novel result by applying the theory of the paper. The CNLS is a natural candidate to illustrate the theory of this paper, possessing the necessary number of symmetries and a Lagrangian density. By reducing the CNLS to the two-way Boussinesq equation, dynamics along (or close to) curves where (3.7) and (3.8) hold may be investigated using a scalar PDE instead. The benefit of doing so is that it reduces the study of a coupled system to that of a single equation with a comparatively large literature regarding its properties and solutions, and so some insight might be given through the use of the phase dynamical equation.

The CNLS takes the form

$$\begin{aligned} i(\Psi_1)_t + \alpha_1(\Psi_1)_{xx} + (\beta_{11}|\Psi_1|^2 + \beta_{12}|\Psi_2|^2)\Psi_1 = 0\,, \\ i(\Psi_2)_t + \alpha_2(\Psi_2)_{xx} + (\beta_{21}|\Psi_1|^2 + \beta_{22}|\Psi_2|^2)\Psi_2 = 0\,, \end{aligned} \tag{5.1}$$

where the unknowns $\Psi_i(x,t)$ are complex valued functions and $\alpha_i, \beta_{ij} \in \mathbb{R}$ constants. In order for this system to possess a generating Lagrangian density, we must impose $\beta_{12} = \beta_{21}$ and so in subsequent working we replace the latter with the former. Such an equation arises in many physical contexts, such as within the study of rogue waves [24, 25] and as a model for a pair of weakly interacting Bose gases [26].

The relative equilibrium of interest is generated by the toral symmetry of each function, giving the symmetry group as $S^1 \times S^1 := \mathbb{T}$. As such, we seek the plane wave solution $\Psi_i = \Psi_i^{(0)} e^{i\theta_i}$ as this solution respects the symmetry, and upon substitution obtain that the amplitudes $\Psi_i^{(0)}$ satisfy

$$\begin{aligned} |\Psi_1^{(0)}|^2 &= \frac{1}{\beta}\big(\beta_{22}(\alpha_1 k_1^2 + \omega_1) - \beta_{12}(\alpha_2 k_2^2 + \omega_2)\big), \\ |\Psi_2^{(0)}|^2 &= \frac{1}{\beta}\big(\beta_{11}(\alpha_2 k_2^2 + \omega_2) - \beta_{12}(\alpha_1 k_1^2 + \omega_1)\big), \end{aligned}$$

where $\beta = \beta_{11}\beta_{22} - \beta_{12}^2$. Moreover, one can obtain the relevant conservation laws as

$$A = \frac{1}{2}\begin{pmatrix} |\Psi_1|^2 \\ |\Psi_2|^2 \end{pmatrix}, \quad B = \Im \begin{pmatrix} (\Psi_1)_x \Psi_1^* \\ (\Psi_2)_x \Psi_2^* \end{pmatrix},$$

where $\Im$ denotes the imaginary part is taken and $*$ denotes the complex conjugate. We can evaluate these along the relative equilibrium solution to obtain the tensors relevant for the theory:

$$\mathbf{A} = \frac{1}{2}\begin{pmatrix} |\Psi_1^{(0)}|^2 \\ |\Psi_2^{(0)}|^2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} k_1|\Psi_1^{(0)}|^2 \\ k_2|\Psi_2^{(0)}|^2 \end{pmatrix}. \tag{5.2}$$

5.1. Criticality of the plane waves

The primary criticality, the one relating solely to $\mathbf{B}$, is discussed first. Evaluating the directional derivative we obtain the matrix

$$\mathrm{D}_{\mathbf{k}}\mathbf{B} = \frac{1}{\beta}\begin{pmatrix} \alpha_1|\Psi_1^{(0)}|^2(1+\beta_{22}E_1^2) & -\frac{2\alpha_1\alpha_2k_1k_2\beta_{12}}{\beta} \\ -\frac{2\alpha_1\alpha_2k_1k_2\beta_{12}}{\beta} & \alpha_2|\Psi_2^{(0)}|^2(1+\beta_{22}E_2^2) \end{pmatrix},$$

where we have introduced the dimensionless quantities

$$E_1^2 = \frac{2\alpha_1k_1^2}{\beta|\Psi_1^{(0)}|^2}, \quad E_2^2 = \frac{2\alpha_2k_2^2}{\beta|\Psi_2^{(0)}|^2}.$$

There is a natural analogy between the above quantities and Froude numbers in stratified fluid flows, which will be introduced in the second example, as they are both dimensionless quantities that characterise subcritical and supercritical behaviour. It is worth noting that the above quantities do not yet appear in the typical literature, but are introduced here for simplicity. Of interest is the zero eigenvalue condition of the above matrix which gives the condition for primary criticality, which requires that

$$(1+\beta_{22}E_1^2)(1+\beta_{11}E_2^2) = \beta_{12}^2E_1^2E_2^2. \tag{5.3}$$

This forms a surface of values in $(\mathbf{k}, \boldsymbol{\omega})$ space, whose nature is determined by the β_{ij}. Assuming that one chooses values of these parameters lying on this surface, we may find the zero eigenvector of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ as

$$\boldsymbol{\zeta} = \begin{pmatrix} \frac{2\alpha_1\alpha_2k_1k_2\beta_{12}}{\beta} \\ \alpha_1|\Psi_1^{(0)}|^2(1+\beta_{22}E_1^2)\,. \end{pmatrix}.$$

The second criticality relates to the tensor

$$\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B} = \frac{1}{\beta}\begin{pmatrix} 2\alpha_1\beta_{22}k_1 & -\beta_{12}(\alpha_1k_1+\alpha_2k_2) \\ -\beta_{12}(\alpha_1k_1+\alpha_2k_2) & 2\alpha_2\beta_{11}k_2 \end{pmatrix}.$$

By considering the relevant projection, one can show that the second criticality condition requires

$$k_2|\Psi_1^{(0)}|^2(\beta_{22}+\beta E_1^2) + k_1|\Psi_2^{(0)}|^2(\beta_{11}+\beta E_2^2) = 0.$$

There are a number of ways this can be achieved due to the degrees of freedom afforded by the parameters β_{ij}. To illustrate, we use the parameter values from [26] and set

$$\alpha_i = 1,\ \beta_{11} = \beta_{22} = -1 \quad \text{and} \quad \beta_{12} = \beta_{21} = -\alpha \tag{5.4}$$

for some $\alpha \in (0,1)$. By choosing $k = k_1 = -k_2$ and $\omega = \omega_1 = \omega_2$ one automatically satisfies the second condition and thus the first is met when

$$E_1^2 = E_2^2 = \frac{1}{1 \pm \alpha}\,.$$

As a consequence, one requires $\omega < 0$ as $k = \pm\sqrt{-\frac{(1-\alpha)\omega}{2(1\pm\alpha)+(1-\alpha)}}$. There is scope for additional ways for both conditions to be met simultaneously, but this merely illustrates a simple way that these are satisfied in this setting. Under the assumption that the above criticality conditions are met, we can find the generalised eigenvector $\boldsymbol{\gamma}$ as

$$\boldsymbol{\gamma} = \frac{\beta_{12}}{\beta}\begin{pmatrix} \frac{2\alpha_2\beta_{22}E_1^2k_2}{\alpha_1|A_0|^2(1+\beta_{22}E_1^2)} - (\alpha_1k_1 + \alpha_2k_2) \\ 0 \end{pmatrix}\,,$$

up to arbitrary shifts in $\boldsymbol{\zeta}$.

5.2. *Emergence of the two-way Boussinesq equation*

Assuming the above criticalities are met, all that remains is to compute the relevant coefficients. The additional matrices relevant to these computations are

$$\begin{aligned}
\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A} =& \frac{1}{\beta}\begin{pmatrix} \beta_{22} & -\beta_{12} \\ -\beta_{21} & \beta_{11} \end{pmatrix}, \\
\mathrm{D}_{\mathbf{k}}^2\mathbf{B} =& \frac{2}{\beta}\left(\begin{array}{cc|cc} 3\alpha_1^2\beta_{22}k_1 & -\alpha_1\alpha_2\beta_{12}k_2 & -\alpha_1\alpha_2\beta_{12}k_2 & -\alpha_1\alpha_2\beta_{12}k_1 \\ -\alpha_1\alpha_2\beta_{12}k_2 & -\alpha_1\alpha_2\beta_{12}k_1 & -\alpha_1\alpha_2\beta_{12}k_1 & 3\alpha_2^2\beta_{11}k_2 \end{array}\right).
\end{aligned}$$

Determining the time term first, its first component is given by

$$\boldsymbol{\zeta}^T\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\zeta} = |\Psi_1^{(0)}|^2|\Psi_2^{(0)}|^2\frac{\kappa}{2}\Bigg(\frac{\alpha_1\beta_{11}(1+\beta_{22}E_1^2)}{|\Psi_2^{(0)}|^2} + \frac{\alpha_2\beta_{22}(1+\beta_{11}E_2^2)}{|\Psi_1^{(0)}|^2} - \frac{4\beta_{12}^2\alpha_1\alpha\cdots}{\beta|\Psi_1^{(0)}|^2|\cdots}\cdots$$

with $\kappa = \alpha_1\beta^{-1}|\Psi_1^{(0)}|^2(1+\beta_{22}E_1^2)$, and its second by

$$\begin{aligned}
&\boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\gamma} \\
&\qquad = -|\Psi_1^{(0)}|^2|\Psi_2^{(0)}|^2\kappa\Bigg(\frac{2(2\beta_{11}\beta_{22}-\beta_{12}^2)\alpha_1\alpha_2k_1k_2}{\beta|\Psi_1^{(0)}|^2|\Psi_2^{(0)}|^2} - \frac{\alpha_1E_1^2\beta_{12}^2}{2|\Psi_2^{(0)}|^2} - \frac{\alpha_2E_2^{2}\cdots}{2|\Psi_1^{(\cdots}}\cdots
\end{aligned}$$

Therefore, it can be seen that the coefficient of the time derivative term is given by

$$\boldsymbol{\zeta}^T \mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\zeta} - \boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\gamma} = \\ |\Psi_1^{(0)}|^2|\Psi_2^{(0)}|^2\kappa\left[\frac{\alpha_1(\beta_{11}+\beta E_1^2)}{2|\Psi_2^{(0)}|^2} + \frac{\alpha_2(\beta_{22}+\beta E_2^2)}{2|\Psi_1^{(0)}|^2} + \frac{4\alpha_1\alpha_2 k_1 k_2}{|\Psi_1^{(0)}|^2|\Psi_2^{(0)}|^2}\right]. \tag{5.5}$$

The quadratic term will have the coefficient

$$\boldsymbol{\zeta}^T \mathrm{D}_{\mathbf{k}}^2\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta}) = 6\alpha_1^2\alpha_2^2\kappa|\Psi_1^{(0)}|^2\Big(|\Psi_1^{(0)}|^2(1+\beta_{22}E_1^2)(\beta_{11}+\beta E_1^2) \\ - \beta_{12}|\Psi_2^{(0)}|^2(1+\beta_{11}E_2^2)\Big) \tag{5.6}$$

The dispersive term is generated by the relevant Jordan chain analysis, and as such we may use the result in [27] to state the projection

$$\boldsymbol{\zeta}^T\mathbf{K} = \frac{1}{2}\kappa\alpha_1\alpha_2\Big(\alpha_2|A_0|^2(\beta_{11}+\beta E_1^2) + \alpha_1|B_0|^2(\beta_{22}+\beta E_2^2)\Big). \tag{5.7}$$

Thus, once (5.5), (5.6) and (5.7) are substituted into (1.13) we obtain the relevant two-way Boussinesq equation at this criticality. For example, using the Salman and Berloff parameters given in (5.4), one can show that the relevant two-way Boussinesq equation at the discussed criticality is given by

$$\big(2 \mp 3\alpha\big)U_{TT} \pm \alpha\left(\frac{3\alpha k|\Psi|^2}{1\pm\alpha}U^2 + U_{XX}\right)_{XX} = 0\,,$$

where

$$|\Psi|^2 = |\Psi_1^{(0)}|^2 = |\Psi_2^{(0)}|^2 = -\frac{2\omega(1\pm\alpha)}{(1+\alpha)\big(2(1\pm\alpha)+(1-\alpha)\big)} \neq 0\,.$$

6. Example 2: two-layer shallow water flow

We now apply the theory to the case of shallow water hydrodynamics with two layers of differing density bounded above by a free surface. There are examples in the literature where the two-way Boussinesq equation has been derived in this setting (for example [14, 28, 29, 30]), and so we demonstrate here that the theory recovers these results using elementary calculations.

The governing equations for this system are

$$(\rho_1\eta)_t + (\rho_1\eta u_1)_x = 0, \tag{6.1a}$$

$$(\rho_2\chi)_t + (\rho_2\chi u_2)_x = 0, \tag{6.1b}$$

$$(\rho_1 u_1)_t + \left(\frac{\rho_1}{2}u_1^2 + g\rho_1\eta + g\rho_2\chi\right)_x = a_{11}\eta_{xxx} + a_{12}\chi_{xxx}, \tag{6.1c}$$

$$(\rho_2 u_2)_t + \left(\frac{\rho_2}{2}u_2^2 + g\rho_2\eta + g\rho_2\chi\right)_x = a_{21}\eta_{xxx} + a_{22}\chi_{xxx}, \tag{6.1d}$$

with

$$\begin{aligned}
a_{11} =& -\frac{1}{3}\rho_1 g\eta_0^2 - \rho_2 g\eta_0\chi_0 - \frac{1}{2}g\chi_0^2,\\
a_{12} =a_{21} =& -\frac{1}{6}\rho_2 g\eta_0^2 - \frac{1}{4}\rho_2 g\eta_0\chi_0 - \frac{\rho_2^2}{2\rho_1}g\eta_0\chi_0 - \frac{5}{12}\rho_2 g\chi_0^2,\\
a_{22} =& -\frac{\rho_2^2}{2\rho_1}g\eta_0\chi_0 - \frac{1}{3}\rho_2 g\chi_0^2\,.
\end{aligned}$$

In these equations, ρ_1, η and u_1 are the density, layer thickness, and the horizontal and vertical velocities in the lower layer, and ρ_2, χ, and u_2 are the corresponding quantities in the upper layer. In the dispersion coefficients, η_0 and χ_0 are quiescent thicknesses in the two layers. The dispersionless version of these equations is derived in [31], and the dispersive terms in x are derived in [29] (see also [14]).

The first two equations (6.1a) are conservation laws and the symmetry associated with them is a constant shift of the velocity potentials which are defined by $u_1 = \partial_x\varphi_1$ and $u_2 = \partial_x\varphi_2$. Time-dependent multiphase wavetrains associated with this symmetry take the form

$$\varphi_1 = \theta_1 := k_1 x + \omega_1 t + \theta_1^0 \quad\text{and}\quad \varphi_2 = \theta_2 := k_2 x + \omega_2 t + \theta_2^0\,,$$

where the k_i are constant velocities in each layer. Substitution into the governing equations gives the thicknesses of each flow as

$$\begin{aligned}
\eta_0 =& \frac{1}{g(\rho_1-\rho_2)}\left(\tfrac{1}{2}(\rho_2 k_2^2 - \rho_1 k_1^2) + R_1 - R_2 - \rho_1\omega_1 + \rho_2\omega_2\right),\\
\chi_0 =& \frac{\rho_1}{g(\rho_1-\rho_2)}\left(R_2 - R_1 - \omega_2 + \omega_1 + \tfrac{1}{2}(k_1^2 - k_2^2)\right),
\end{aligned}$$

where R_1, R_2 are constants of integration (which can be thought of as Bernoulli constants in each layer). These two equations are used to express η_0 and χ_0 in terms of k_i and ω_i in the conservation laws.

6.1. Conservation laws and criticality

The first two equations of this system (6.1a, 6.1b) form the conservation laws for the system. Therefore one has

$$A = \begin{pmatrix} \rho_1\eta \\ \rho_2\chi \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} \rho_1\eta u_1 \\ \rho_2\chi u_2 \end{pmatrix}. \tag{6.2}$$

Evaluating these along the basic state, and differentiating, gives the following first necessary condition for the emergence of the two-way Boussinesq equation,

$$\det\left[\mathrm{D_k}\mathbf{B}\right] = \det\begin{pmatrix} \rho_1\eta_0 - \frac{\rho_1 k_1^2}{g(1-r)} & \frac{\rho_2 k_1 k_2}{g(1-r)} \\ \frac{\rho_2 k_1 k_2}{g(1-r)} & \rho_2\chi_0 - \frac{\rho_2 k_2^2}{g(1-r)} \end{pmatrix} = 0\,,$$

and we have introduced the quantity $r = \frac{\rho_2}{\rho_1} < 1$ as the ratio of the two densities for convenience. Upon evaluation, the determinant condition can be reduced to

$$(1 - F_1^2)(1 - F_2^2) = r, \tag{6.3}$$

where we have introduced the Froude numbers

$$F_1^2 = \frac{k_1^2}{g\eta_0}, \quad F_2^2 = \frac{k_2^2}{g\chi_0}\,. \tag{6.4}$$

This condition for emergence agrees with the classical primary criticality condition found within other works [32, 20, 33]. Assuming that the trace of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ is nonzero, so that zero eigenvalue is simple, one has the eigenvector associated with the zero eigenvalue

$$\boldsymbol{\zeta} = \begin{pmatrix} -\rho_2 k_1 k_2 \\ g\rho_1\eta_0(1 - r - F_1^2) \end{pmatrix}.$$

There is now the second criticality to consider, involving the matrix

$$\mathrm{D}_{\mathbf{k}}\mathbf{A} = \frac{1}{g(1-r)}\begin{pmatrix} -\rho_1 k_1 & \rho_2 k_2 \\ \rho_2 k_1 & -\rho_2 k_2 \end{pmatrix} = \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}^T\,.$$

The condition (3.9) can be solved when

$$\begin{aligned}\boldsymbol{\zeta}^T&(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta} \\ &= -2g^2\rho_1^2\rho_2\chi_0\eta_0^2(1 - r - F_1^2)\left(\frac{k_1}{g\eta_0}(1 - F_2^2) + \frac{k_2}{g\chi_0}(1 - F_1^2)\right) = 0\,.\end{aligned} \tag{6.5}$$

This is exactly the coefficient of the time derivative term of the KdV appearing within the shallow water example in [1], and is also proportional to the second characteristic speed appearing in [20]. One consequence of this is that k_1 and k_2 must be of opposite sign, as $(1-F_1^2)(1-F_1^2) = r > 0$

from the first criticality condition. These criterion form a continuum of solutions in (k_1, k_2, r) space for any chosen η_0, χ_0 as pictured in figure 1. For example,

$$g\eta_0 = 4, \quad g\chi_0 = 10, \quad k_1 = -1, \quad k_2 = 2, \quad r = \frac{9}{20},$$

are values for which both conditions are met and thus at this point the two-way Boussinesq equation is applicable.

Figure 1: An example of a curve in (k_1, k_2, r) on which both criticality conditions (6.3), (6.5) are satisfied, for $g\eta_0 = 4$, $g\chi_0 = 10$.

When this condition is met, one is able to find $\boldsymbol{\gamma}$ as

$$\boldsymbol{\gamma} = \frac{\rho_2}{g\eta_0(1 - r - F_1^2)} \begin{pmatrix} 2k_1^2 k_2 + g(k_1 + k_2)\eta_0(1 - r - F_1^2) \\ 0\,. \end{pmatrix}, \tag{6.6}$$

up to shifts in $\boldsymbol{\zeta}$.

6.2. Emergence of the two-way Boussinesq equation at criticality

The relevant coefficient matrices for the vector Boussinesq equation (1.12) are

$$\begin{aligned} \mathrm{D}_{\boldsymbol{\omega}}\mathbf{A} =& \frac{1}{g(1-r)} \begin{pmatrix} -\rho_1 & \rho_2 \\ \rho_2 & -\rho_2 \end{pmatrix}, \\ \mathrm{D}^2_{\mathbf{k}}\mathbf{B} =& \frac{1}{g(1-r)} \left(\begin{array}{cc|cc} -3\rho_1 k_1 & \rho_2 k_2 & \rho_2 k_2 & \rho_2 k_1 \\ \rho_2 k_2 & \rho_2 k_1 & \rho_2 k_1 & -3\rho_2 k_2 \end{array} \right). \end{aligned}$$

Now project to obtain the coefficients of the Boussinesq equation using the theory presented in §4. The first term appearing in the time coefficient projects to

$$\boldsymbol{\zeta}^T \mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\zeta} = -\frac{g^2\rho_1^2\rho_2\eta_0^2\chi_0(1 - r - F_1^2)}{1 - r} \left(\frac{1 - r - F_2^2}{g\eta_0} + \frac{2rk_1k_2}{g^2\eta_0\chi_0} + \frac{1 - r - F_1^2}{g\chi_0} \right),$$

and the second gives

$$\boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\gamma} = -\frac{g^2\rho_1^2\rho_2\eta_0^2\chi_0(1-r-F_1^2)}{1-r}\left(\frac{2(2-r)k_1k_2}{g^2\eta_0\chi_0}-\frac{rF_1^2}{g\chi_0}-\frac{rF_2^2}{g\eta_0}\right),$$

so that the overall coefficient of the time derivative term is

$$\begin{aligned}\boldsymbol{\zeta}^T\mathrm{D}_{\boldsymbol{\omega}}\mathbf{A}\boldsymbol{\zeta} - \boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\gamma} \\ = -g^2\rho_1^2\rho_2\eta_0^2\chi_0(1-r-F_1^2)\left(\frac{4k_1k_2}{g^2\eta_0\chi_0}-\frac{1-F_1^2}{g\chi_0}-\frac{1-F_2^2}{g\eta_0}\right).\end{aligned} \tag{6.7}$$

The factor within the bracket is proportional to the coefficient of the quadratic term of the characteristic polynomial discussed in the appendix of [20], and so this coefficient is supported by the literature.

For the second derivative,

$$\begin{aligned}\boldsymbol{\zeta}^T\mathrm{D}_{\mathbf{k}}^2\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta}) \\ = 3g^2\rho_1^3\rho_2k_2\eta_0^2(1-r-F_1^2)\big(\chi_0 r(1-F_2^2)F_1^2-\eta_0(1-F_1^2)^2F_2^2\big).\end{aligned}$$

Since the dispersion vector $\mathbf{K}$ is a coefficient of a linear term, it can be calculated using the dispersion relation (associated with the linearisation about the basic state) or using the Jordan chain. The details are omitted and we just state the result

$$\mathbf{K} = -\frac{1}{g(1-r)}\begin{pmatrix}\rho_1k_1T_1\\ \rho_2k_2T_2\end{pmatrix},$$

with

$$\begin{aligned}T_1 &= \frac{\zeta_2(a_{11}r+(1+r)a_{12}+a_{22})k_2-\zeta_1(a_{11}+2a_{12}+a_{22})k_1}{g\rho_1(1-r)},\\ T_2 &= \frac{\zeta_1(a_{11}r+(1+r)a_{12}+a_{22})k_1-\zeta_2(a_{11}r^2+2a_{12}r+a_{22})k_2}{g\rho_2(1-r)}.\end{aligned}$$

This generates the coefficient of dispersion as

$$\zeta_1\mathscr{K}_1+\zeta_2\mathscr{K}_2 = g\rho_1^2\eta_0^2\chi_0(1-r-F_1^2)\big(a_{11}r(1-F_2^2)-2ra_{12}+(1-F_1^2)a_{22}\big).$$

Therefore by gathering the coefficients discussed above, one finds that the emergent two-way Boussinesq equation, once simplified, is given by

$$\begin{aligned}&\chi_0\left(\frac{1-F_1^2}{g\chi_0}+\frac{1-F_2^2}{g\eta_0}-\frac{4k_1k_2}{g^2\eta_0\chi_0}\right)U_{TT}\\ &+\Big[3\rho_1k_2\big(\chi_0r(1-F_2^2)F_1^2-\eta_0(1-F_1^2)^2F_2^2\big)UU_X\\ &\quad+\frac{\chi_0}{\rho_2 g}\big(a_{11}r(1-F_2^2)-2ra_{12}+(1-F_1^2)a_{22}\big)U_{XXX}\Big]_X = 0.\end{aligned} \tag{6.8}$$

Without further investigation, one is able to anticipate that the time dispersion coefficients are to be of opposite sign. This can be inferred from the fact that $k_1 \neq k_2$ (as $k_1 k_2 < 0$) and there is no surface tension present and so the solution $U = 0$ is expected to be unstable due to the system residing in the Kelvin-Helmholtz instability regime [34].

7. Concluding remarks

This paper has illustrated how, if given a multiphase wavetrain in a Lagrangian framework and provided suitable conditions hold, a two-way Boussinesq equation emerges as a reduction about the multiphase relative equilibrium solution. This result was then illustrated on two examples of physical interest to show how the conditions for the reduction can be assessed and the coefficients calculated.

There are many ways in which the result here may be extended. Primarily, the formulation of the modulational analysis in this paper lends itself to the N-phase generalisation. It is expected that as long as the zero eigenvalue of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ is simple one that the same result to hold for arbitrarily many phases. One may also consider the addition of a transverse spatial variable, in which case the expectation is that the vector analogy of the $2+1$ two-way Boussinesq equation obtained in [12] as the result of the 2+1 dimensional analysis.

The generalised zero eigenvectors of the matrix appearing in the linear Whitham system that have been discussed in this paper have an interesting connection with time degeneracies, which in turn leads to higher order time derivatives. One can continue to the case of a third generalised eigenvector under suitable conditions and induce a further time degeneracy of the system. Abridging the theory in a suitable way seems to indicate a KdV-like equation emerges as the suitable nonlinear phase equation, but with a a third order time term, given by

$$a_0 U_{TTT} + \left(\frac{1}{2} a_1 U^2 + a_2 U_{XX}\right)_{XXX} = 0\,,$$

for coefficients a_i. The fact that the multiphase setting has at least 4 parameters and the criticality to obtain such a system requires three conditions to be met implies that this may be the first time such a system has emerged from any approach. Although such a system appears to have an unstable trivial state in all cases, it is the case that bounded solutions can still be obtained in a similar way to both the KdV and two-way Boussinesq equations. For example this system possesses the solitary wave solution

$$U = \frac{3a_0c^3}{a_1}\text{sech}^2\left(\frac{1}{2}\sqrt{\frac{c^3a_0}{a_2}}(X-cT)+\phi_0\right)$$

for some wavespeed c and phase shift ϕ_0, whenever $a_0a_2 > 0$.

The main focus of the paper has been around conditions that relate to the multiplicity of the zero eigenvalue of the Whitham system, however there are other degeneracies which can occur that cannot be linked directly to the linear Whitham equations. For example, higher order dispersion is expected when the coefficient of the highest spatial derivative vanishes. One also expects higher order nonlinearities when $\boldsymbol{\zeta}^T \mathrm{D}_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta}) = 0$. The methodology to derive phase equations in these cases is expected to be similar as those presented in this paper, albeit with some adjustments of the scales involved so that the relevant terms balance.

Another open question the theory of this paper poses is the physical significance of the second criticality condition. In previous works (such as [20]) the connection between the first criticality and the stability of the physical systems in question has been made, but for the other criterion (3.8) the link is not so clear nor does it appear to be explored. Analysis of the link between the second criticality and system stability may yield better insight into the nature of the instabilities beyond this threshold, as well as how well the two-way Boussinesq equation predicts the qualitative behaviour of the wavenumber in such regions.

Acknowledgements

The author would like to thank Prof. Tom Bridges for his insightful discussions regarding this work and the referees for their helpful suggestions. The author is supported by a fully funded PhD. studentship under the EPSRC grant EP/L505092/1.

References

1. D.J. Ratliff & T.J. Bridges, *Multiphase wavetrains, singular wave interactions, and the emergence of the KdV equation*, Proc. R. Soc. Lond. A **472**, 20160456 (2016).
2. G.B. Whitham, *Linear and nonlinear waves*, John Wiley & Sons (2011).
3. G.A. El & M.A. Hoefer, *Dispersive shock waves and modulation theory*, Physica D **333**, 11–65 (2016).
4. M. J. Lighthill, *Some special cases treated by the Whitham theory*, Proc.Roy. Soc. Lond. A **299**, 1456 (1967).

5. G. A. El, R. H. J. Grimshaw, & N.F. Smyth, *Unsteady undular bores in fully nonlinear shallow-water theory*, Phys. Fluids **18.2**, 027104 (2006).
6. A. V. Tyurina, & G. A. El, *Hydrodynamics of modulated finite-amplitude waves in dispersive media*, J. Exp. Theor. Phys. **88.3**, 615–625 (1999).
7. G. A. El, A. Gammal, E. G. Khamis, R. A. Kraenkel & A. M. Kamchatnov *Theory of optical dispersive shock waves in photorefractive media*, Phys. Rev. A **76**, 053183 (2007).
8. M. Crosta, S. Trillo, A. Fratalocchi *The Whitham approach to dispersive shocks in systems with cubic-quintic nonlinearities*, New J. Phys. **14**, 093019 (2012).
9. M. A. Hoefer, M. J. Ablowitz, I. Coddington, E. A. Cornell, P. Engels, & V. Schweikhard, *Dispersive and classical shock waves in Bose-Einstein condensates and gas dynamics*, Phys. Rev. A **74**, 023623 (2006).
10. M. A. Hoefer, P. Engels, & J. J. Chang, *Matterwave interference in BoseEinstein condensates: A dispersive hydrodynamic perspective*, Physica D **238.15**, 1311-1320 (2009).
11. T.J. Bridges, *A universal form for the emergence of the Korteweg–de Vries equation*, Proc. Roy. Soc. Lond. A **469**, 20120707 (2013).
12. D.J. Ratliff & T.J. Bridges, *Whitham modulation equations, coalescing characteristics, and dispersive Boussinesq dynamics*, Physica D, **333** (2016).
13. J.V. Boussinesq, *Théorie des ondes et des remous qui se propagent le long dun canal rectangulaire horizontal, en communiquant au liquide contene dans ce canal des vitesses sensiblement pareilles de la surface au fond*, J. Math. Pure Appl. **17** 55–108 (1872).
14. T.J. Bridges and D.J. Ratliff, *Double criticality and the two-way Boussinesq equation in stratified shallow water hydrodynamics*, Phys. Fluids **28** 062103 (2016).
15. R.S. Johnson, *A two-dimensional Boussinesq equation for water waves and some of its solutions*, J. Fluid Mech. **323** 65–78 (1996).
16. T.J. Bridges & F.E. Laine-Pearson, *Multisymplectic relative equilibria, multiphase wavetrains, and coupled NLS equations*, Stud. Appl. Math. **107**, 137-155 (2001).
17. T.J. Bridges, *Symmetry, Modulation and Nonlinear Waves* Lectures on the Theory of Water Waves **426**, 261–283 (2016).
18. M.J. Ablowitz, *Approximate methods for obtaining multi-phase modes in nonlinear dispersive wave problems*, Stud. Appl. Math. **51**, 17–55 (1972).
19. M.J. Ablowitz & D.J. Benney, *The evolution of multi-phase modes for nonlinear dispersive waves*, Stud. Appl. Math. **49**, 225–238 (1970).
20. G.A. Lawrence, *On the hydraulics of Boussinesq and non-Boussinesq two-layer flows*, J. Fluid Mech. **215**, 457–480 (1990).
21. A. Ali and H. Kalisch, *Mechanical balance laws for Boussinesq models of surface water waves*, J. Nonlinear Sci. **22**, 371–398 (2012)
22. R. Abraham and J. E. Marsden, *Foundations of mechanics*, Benjamin/Cummings Publishing Company (1978).
23. T.J. Bridges, *A geometric formulation of the conservation of wave action and its implications for signature and the classification of instabilities*, Proc. R. Soc. Lond. A **453** 1365–1395 (1997).
24. M.J. Ablowitz, & T.P. Horikis, *Interacting nonlinear wave envelopes and rogue wave formation in deep water*, Phys. Fluids, **27.1** (2015).
25. M. Onorato, A. R. Osborne, and M. Serio, *Modulational instability in crossing sea states: A possible mechanism for the formation of freak waves* Phys. Rev. Lett., **96.1** (2006).

26. H. Salman & N.G. Berloff, *Condensation of classical nonlinear waves in a two-component system*, Physica D **238** 1482–1489 (2009).
27. D.J. Ratliff, *Reduction of coupled NLS equations to KdV via modulation of a two-phase wavetrain*, Preprint (2017).
28. W. Craig, P. Guyenne & H. Kalisch, *Hamiltonian long wave expansions for free surfaces and interfaces*, Commun. Pure Appl. Math., **58**, 15871641 (2005).
29. N.M. Donaldson, *Criticality theory and conformal mapping techniques for single and two layer water-wave systems*, Ph.D. thesis, University of Surrey, (2006).
30. H. Y. Nguyen & F. Dias, *A Boussinesq system for two-way propagation of interfacial waves*, Physica D, **237**, 23652389 (2008).
31. P.G. Baines, *A general method for determining upstream effects in stratified flow of finite depth over long two-dimensional obstacles*, J. Fluid Mech. **188**, 1–22 (1988).
32. G.S. Benton, *The occurrence of critical flow and hydraulic jumps in a multi-layered fluid system*, J. Meteor. **11**, 139–150 (1954).
33. T.J. Bridges and N.M. Donaldson, *Criticality manifolds and their role in the generation of solitary waves for two-layer flow with a free surface*, Euro. J. Mech. B/Fluids **28**, 117–126 (2007).
34. T.B. Benjamin and T.J. Bridges, *Reappraisal of the Kelvin-Helmholtz problem. Part 1. Hamiltonian structure*, J. Fluid Mech. **333**, 301–325 (1997).