Flux Singularities in Multiphase Wavetrains and the Kadomtsev-Petviashvili Equation with Applications to Stratified Hydrodynamics

By Daniel J. Ratliff

This paper illustrates how the singularity of the wave action flux causes the Kadomtsev-Petviashvili (KP) equation to arise naturally from the modulation of a two-phased wavetrain, causing the dispersion to emerge from the classical Whitham modulation theory. Interestingly, the coefficients of the resulting KP are shown to be related to the associated conservation of wave action for the original wavetrain, and so may be obtained prior to the modulation. This provides a universal form for the KP as a dispersive reduction from any Lagrangian with the appropriate wave action flux singularity. The theory is applied to the full water wave problem with two layers of stratification, illustrating how the KP equation arises from the modulation of a uniform flow state and how its coefficients may be extracted from the system.

keywords: modulation, Lagrangian dynamics, nonlinear waves, Water waves and fluid dynamics, asymptotic analysis.

1. Introduction

In the study of nonlinear waves in two spatial dimensions, one of the fundamental equations which arises is the Kadomtsev-Petviashvili (KP) equation,

$$\big(U_T + UU_X + U_{XXX}\big)_X \pm U_{YY} = 0\,, \tag{1}$$

for scalar function $U(X,Y,T)$. The equation itself models waves for which dispersion balances nonlinear steepening effects and weak transverse effects, and is a generalisation of the one spatial dimension Korteweg-de Vries (KdV) equation. First arising within the context of

Loughborough University, Loughborough, United Kingdom, LE11 3TU; email: d.j.ratliff@lboro.ac.uk

DOI: 10.1111/SIA.PM-Euler-RFinal

STUDIES IN APPLIED MATHEMATICS 00:1–34

plasma physics [1], it has since been recognised to be a universal equation in the sense that it arises across a variety of different applications such as water waves [2, 3], nonlinear optics [4, 5] and Bose-Einstein condensates [6, 7].

The principal aim of this paper is to illustrate how the KP equation (1) arises naturally from a phase dynamical perspective utilising solutions parameterised by multiple phase variables, such as wavetrains. Although the reduction will be presented for the case of general multiphase periodic wavetrains, the emphasis will be to illustrate how the theory suggests the KP equation emerges from the full water wave problem with stratification via the modulation of a uniform flow state. This has been previously accomplished for zero velocity states in [8, 9]), the outcome of this paper provides not only a new mechanism for its emergence but also demonstrates how it arises from nontrivial uniform background states.

The method of phase dynamics has a rich history, and has been developed over several years within multiple fields. The first use of the technique stems from the original studies of modulation undertaken by Whitham [10] and their extension by Hayes into higher dimension [11], deriving nonlinear equations governing the wavenumbers and frequency of the wave to determine the evolution of the wave. The strategy has been utilised and developed by several authors since, with a review of various contributions appearing in [12] and the references therein. Additionally, an analogous system being derived for reaction diffusion systems by Cross and Newell [13]. The ideas of phase dynamics were further developed within the works of Kuramoto [14, 15], which demonstrated how the degeneracy of the phase diffusion equation, the linearisation of the Cross-Newell system, led to the emergence of nonlinearity and higher order dissipation. These concepts were later used by Doelman et al. [16] to derive Burgers' equation in similar settings. The ideas of Whitham, Kuramoto and Doelman et al. were combined within the Lagrangian setting by Bridges [17, 18], demonstrating that the degeneracy of the linearised Whitham equations could be formulated in terms of conservation law criticality. Ultimately, through using ideas similar to Kuramoto and Doelman et al., the KdV equation arose from the modulation and thus the emergence of dispersion from the breakdown of the dispersionless Whitham equations via modulation. The approach of Bridges forms the principal methodology to obtain the KP within this paper.

To best see how the phase dynamics considered in this paper ties into classical Whitham modulation theory, we start with the generic family of PDEs generated from a Lagrangian,

$$\mathscr{L}(V) = \iiint \mathcal{L}(V, V_x, V_y, V_t, \dots)\; dx\, dy\, dt\,, \tag{2}$$

for state vector $V(x,y,t)\in\mathbb{R}^n$ and Lagrangian density $\mathcal{L}$. The modulation of wavetrains with multiple phases is of a similar nature to those of a single phase. We assume that there exists a doubly periodic wavetrain solution to the Euler-Lagrange equations associated with (2) with two phases, which take the form

$$
\begin{gathered}
V=\hat{V}(\boldsymbol{\theta};\mathbf{k},\mathbf{m},\boldsymbol{\omega})\,,\quad \boldsymbol{\theta}=\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}=\begin{pmatrix}k_1x+m_1y+\omega_1t\\ k_2x+m_2y+\omega_2t\end{pmatrix}\,,\\
\mathbf{k}=\begin{pmatrix}k_1\\ k_2\end{pmatrix}\,,\ \mathbf{m}=\begin{pmatrix}m_1\\ m_2\end{pmatrix}\,,\ \boldsymbol{\omega}=\begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix}\,,\\
\hat{V}(\theta_1+2\pi,\theta_2)=\hat{V}(\boldsymbol{\theta})=\hat{V}(\theta_1,\theta_2+2\pi)\,.
\end{gathered}\tag{3}
$$

The dependence of $\hat{V}$ on the wavenumbers $\mathbf{k}$, $\mathbf{m}$ and frequencies $\boldsymbol{\omega}$ is parameter-like in nature, as expressed by the semicolon, however derivatives with respect to these may be taken and are meaningful [16]. One may then follow the strategy of Whitham and Hayes by imposing that the phases are slowly varying functions of space and time, so that instead $\boldsymbol{\theta}=\varepsilon^{-1}\boldsymbol{\Theta}(X,Y,T)$ with $X=\varepsilon x$, $Y=\varepsilon y$, $T=\varepsilon t$ and small parameter $0<\varepsilon\ll 1$. As such, the wavenumbers and frequencies become defined as the slow derivatives of this slow phase, giving that now $\mathbf{k}=\boldsymbol{\Theta}_X$, $\mathbf{m}=\boldsymbol{\Theta}_Y$ and $\boldsymbol{\omega}=\boldsymbol{\Theta}_T$, so that

$$
V=\widehat{V}(\varepsilon^{-1}\boldsymbol{\Theta};\,\boldsymbol{\Theta}_X,\,\boldsymbol{\Theta}_Y,\,\boldsymbol{\Theta}_T)\,.\tag{4}
$$

By substituting the above slowly varying wave into the Lagrangian (2) and taking variations with respect to $\boldsymbol{\Theta}$, one obtains a system of equations

$$
\begin{gathered}
\mathbf{A}(\mathbf{k},\mathbf{m},\boldsymbol{\omega})_T+\mathbf{B}(\mathbf{k},\mathbf{m},\boldsymbol{\omega})_X+\mathbf{C}(\mathbf{k},\mathbf{m},\boldsymbol{\omega})_Y=\mathbf{0}\,,\\
\text{with}\quad \mathbf{k}_Y=\boldsymbol{\omega}_X,\ \mathbf{m}_X=\mathbf{k}_Y,\ \mathbf{m}_T=\boldsymbol{\omega}_Y\,,
\end{gathered}\tag{5}
$$

which form the multiple phase analogue of the Whitham modulation equations. It transpires that the vector valued function $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ have the surprising property that their elements are related to the components of the conservation of wave action associated with each phase of the wavetrain. Although the Whitham equations are typically derived for wavetrains, they may also be obtained for any solution which are steady with respect to some group action. Such solutions are known as relative equilibria and can be considered as the generalisation of periodic waves (see [18] and references therein). This dramatically increases the range of solutions whose slow evolution can be modelled by the system (5), as well as the number of problems for which the Whitham equations may be used to gain insight.

In general, the Whitham modulation equations (5) are robust, however an interesting case arises when one of the functions $\mathbf{A}$, $\mathbf{B}$ or $\mathbf{C}$ fail to be injective. Throughout this paper it will be assumed that this occurs in $\mathbf{B}$ in the variable $\mathbf{k}$, leading to a loss of injectivity when its Jacobian is singular,

$$\det\big[\mathrm{D}_{\mathbf{k}}\mathbf{B}\big] = 0\,, \tag{6}$$

along some curve $\mathbf{k} = \mathbf{k}_0(\boldsymbol{\omega}, \mathbf{m})$. Throughout this paper, this condition will be referred to as criticality. There is an interesting connection between this condition and the stability of the original solution $\widehat{V}$ [19, 20, 21], which gives a physical interpretation of criticality. Assuming that the zero eigenvalue arising from the condition (6) is simple, we may define the eigenvector $\boldsymbol{\zeta}$,

$$\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\zeta} = \mathbf{0}\,.$$

The emergence of this eigenvector is the key difference between the multiphase modulation and that of single phase solutions, where no such object arises, and highlights the increased role criticality plays in the multiphase case. The consequence of criticality in the modulation approach is to lead to the appearance of dispersive terms from the phase dynamics, which will be illustrated within this paper. As such, (6) may also be viewed as a diagnostic for when one expects dispersion to enter the phase dynamics of the wavetrain.

At points where the criticality condition (6) holds, the Whitham approach must be altered via a rescaling in order to rebalance the degenerate X component of the equation. The strategy to do is to introduce new scalings to account for the singularity in $\mathrm{D}_{\mathbf{k}}\mathbf{B}$. To derive the KP equation within the setting of this paper, this is achieved by constructing a modulation ansatz of the form

$$V = \widehat{V}(\boldsymbol{\theta} + \varepsilon\boldsymbol{\Phi}(X, Y, T; \varepsilon); \mathbf{k} + \varepsilon^2\boldsymbol{\Phi}_X,\, \mathbf{m} + \varepsilon^3\boldsymbol{\Phi}_Y,\, \boldsymbol{\omega} + \varepsilon^4\boldsymbol{\Phi}_T\big) + \varepsilon^3 W(\boldsymbol{\theta}, X, Y, T; \varepsilon)\,, \tag{7}$$

for the slow variables

$$X = \varepsilon x\,, \quad Y = \varepsilon^2 y\,, \quad T = \varepsilon^3 t\,,$$

and the modulation function $\boldsymbol{\Phi}$. There is a clear parallel between this ansatz and the one used for the Whitham modulation theory (4), but instead perturbs about fixed wave variables $\mathbf{k}$, $\mathbf{m}$, $\boldsymbol{\omega}$. The various criteria arising within the analysis will impose conditions on these in order for certain reductions to emerge, for example (6). The scalings are chosen in anticipation of the KP equation arising from the modulation analysis, whereas the form of the ansatz is inspired by the previous work on single phase wavetrains [17, 22], as well as from the initial studies into the modulation of multiphase wavetrains [23, 24, 25]. In particular, the presence of an additional spatial dimension requires one to split up the modulation function $\mathbf{\Phi}$ as

$$\mathbf{\Phi} = \boldsymbol{\phi}(X, Y, T; \varepsilon) + \varepsilon \boldsymbol{\psi}(X, Y, T; \varepsilon)\,,$$

inspired by the latter multiphase works. This represents another major divergence from the modulation of single phases, in particular from previous studies in which the KP equation was obtained [28], where only one phase perturbation was required. Full details of the role of this decomposition in the derivation of the KP will be presented within the manuscript. The overall strategy of the modulation approach presented here is to substitute this ansatz into the Euler-Lagrange equations associated with the Lagrangian (2), Taylor expand around $\varepsilon = 0$ and then to solve the resulting system at each power of ε.

There are several benefits to using the above modulation approach over the standard asymptotic series expansion technique. Firstly, by perturbing the independent variables of the wave, derivatives of the original wave solution are naturally introduced into the analysis, leading to several unimportant terms which do not contribute cancelling automatically. Secondly, this approach most readily leads to the connection between the system's conservation laws and the various solvability conditions which arise, as well as the final coefficients appearing in the KP equation. Finally, the consideration of an abstract Lagrangian and wavetrain allow the results derived here to be applicable to any Lagrangian possessing the relevant solution and meeting the criterion necessary for the KP to emerge, without any further asymptotics.

By following the phase dynamical procedure, the result is the vector equation

$$(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta} U_T + \mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta})UU_X + \mathscr{K}U_{XXX} + \mathrm{D}_{\mathbf{m}}\mathbf{C}\boldsymbol{\phi}_{YY} + (\mathrm{D}_{\mathbf{k}}\mathbf{C} + \mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\psi}_{XY} + \mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\alpha}_{XX} = \mathbf{0} \quad (8)$$

for the scalar function $U(X, Y, T)$, arising from the theory due to the criticality condition (6) through the relation

$$\boldsymbol{\phi}_X = \boldsymbol{\zeta} U\,,$$

and an additional solvability condition imposes that

$$\boldsymbol{\psi}_{XX} = -\boldsymbol{\eta} U_Y \quad \text{with} \quad \mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\eta} = (\mathrm{D}_{\mathbf{k}}\mathbf{C} + \mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\zeta}\,.$$

The way in which this condition arises from the analysis will be discussed within the manuscript. The system (8) possesses the surprising property that the majority of its coefficients are formed from derivatives of the conservation law components evaluated along the original wavetrain. The only one that isn't, $\mathscr{K}$, appears from a Jordan chain analysis arising from the theory. The equation (8) is almost the vector form of the KP equation, except for the presence of an inhomogeneity, owing to the arbitrary function $\boldsymbol{\alpha}(X, Y, T)$. This is yet another element of the analysis unique to the modulation of multiple phases, which is required to prevent U from being trivial. This vector equation can in fact be projected to form a scalar PDE to remove the inhomogeneity, achieved by multiplying on the left by $\boldsymbol{\zeta}$, as $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ is symmetric. Doing so, and then differentiating with respect to X, leads to the scalar KP equation

$$\left[\boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta}U_T+\boldsymbol{\zeta}^T\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta})UU_X+\boldsymbol{\zeta}^T\mathscr{K}U_{XXX}\right]_X \\ +\boldsymbol{\zeta}^T\left[\mathrm{D}_{\mathbf{m}}\mathbf{C}\boldsymbol{\zeta}-(\mathrm{D}_{\mathbf{k}}\mathbf{C}+\mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\eta}\right]U_{YY}=0\,. \quad (9)$$

Thus, the key result of this paper will be to show how through use of the ansatz (7) about points where (6) holds the KP equation arises as the dispersive correction to the Whitham equations.

The principal example to illustrate the application of the theory will be to show that the full water wave problem with two stratified layers may be reduced to the KP equation. As a nonlinear wave equation, the KP equation already has a central role in the modelling of water waves [8, 26, 27]. The theory presented within this paper offers several additions to the existing literature on the KP equation. Firstly, it provides a new mechanism for the equation to arise from water wave problems via phase modulation. Secondly, the result of this paper shows how the KP equation is obtained for *finite* uniform velocity flows, extending the existing literature which obtains it about zero velocity flows (for example, see [8, 9]). The final major contribution this paper presents is that arising from the modulation approach is a condition, which acts as a diagnostic to determine scenarios when the KP is expected to appear from the water wave problem, which is formulated using properties of the flow itself. This will be illustrated within the manuscript, and the example will demonstrate how the relevant criticality may be assessed for the full stratified water wave problem and how the relevant coefficients of the KP equation may be constructed via the conservation laws.

An outline of the paper is as follows. In §2 the relevant abstract setup required for the modulation approach is developed. This includes a discussion of the multisymplectic structure, the geometric form of the conservation law and the linearisation about the basic state which generates the Jordan chain structures. Following this, the details of the reduction procedure leading to the KP equation are given in §3, demonstrating how the multisymplectic structure leads to the coefficients of the KP equation having a form tied to the conservation laws for the system. Using these ideas, §4 provides the details for how the full two-layered water wave problem can be reduced to the KP equation about a uniform flow background state. Concluding remarks, as well as extensions to the results presented here, are discussed in §5.

2. Abstract Set-up

We begin by considering the abstract Lagrangian (2), although it is more desirable to recast it into multisymplectic form. The process to do so is to take a sequence of Legendre transforms in time (which transforms the original Lagrangian into its Hamiltonian representation) and in both of its spatial components. This then leads to the *multisymplectic formulation* of the Lagrangian,

$$\mathcal{L} = \iint \frac{1}{2}\langle Z, \boldsymbol{\alpha}(Z)_t + \boldsymbol{\beta}(Z)_x + \boldsymbol{\gamma}(Z)_y\rangle - S(Z)\; dx\, dy\, dt\,, \tag{10}$$

for new state vector Z, which is the previous state vector V along with the relevant conjugate state variables, and $\langle\cdot,\cdot\rangle$ denotes the standard inner product on $\mathbb{R}^2$. The vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are two-forms, which are typically nonlinear. In the previous analyses involving modulation within multisymplectic frameworks, these were assumed linear so that $\boldsymbol{\alpha} = \mathbf{M}Z$, $\boldsymbol{\beta} = \mathbf{J}Z$ and $\boldsymbol{\gamma} = \mathbf{K}Z$ for constant skew-symmetric matrices $\mathbf{M}$, $\mathbf{J}$ and $\mathbf{K}$ [17, 23, 28]. Many cases are covered by this assumption, such as shallow water hydrodynamics and various nonlinear Schrödinger models, however this is not generic. There are Lagrangians of physical interest, such as those associated with the water wave problem, whose two-forms are nonlinear. Thus, the paper will proceed with the nonlinear case so that the results generated here will apply to such systems. In cases where these two-forms are indeed nonlinear, similar structures to those in the linear case emerge by defining

$$\mathbf{M}(Z) = \frac{1}{2}\big(\mathrm{D}\boldsymbol{\alpha} - \mathrm{D}\boldsymbol{\alpha}^T\big), \quad \mathbf{J}(Z) = \frac{1}{2}\big(\mathrm{D}\boldsymbol{\beta} - \mathrm{D}\boldsymbol{\beta}^T\big), \quad \mathbf{K}(Z) = \frac{1}{2}\big(\mathrm{D}\boldsymbol{\gamma} - \mathrm{D}\boldsymbol{\gamma}^T\big),$$

This guarantees that $\mathbf{M}$, $\mathbf{J}$ and $\mathbf{K}$ are skew-symmetric. Variation of the multisymplectic Lagrangian (10) leads to the Euler Lagrange equation

$$\mathbf{M}(Z)Z_t + \mathbf{J}(Z)Z_x + \mathbf{K}(Z)Z_y = \nabla S(Z) \tag{11}$$

A key assumption which allows the theory of this paper to proceed is that the above equation has a doubly periodic wavetrain solution $\widehat{Z}$ with periods 2π and wave variables $\theta_i = k_i x + \omega_i t + m_i y + \theta_i^0$ of the form

$$Z = \widehat{Z}(\boldsymbol{\theta}; \mathbf{k}, \mathbf{m}, \boldsymbol{\omega})\,, \tag{12}$$

with

$$\boldsymbol{\theta} = \begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix}k_1\\ k_2\end{pmatrix}, \quad \mathbf{m} = \begin{pmatrix}m_1\\ m_2\end{pmatrix}, \quad \boldsymbol{\omega} = \begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix}.$$

The formulation of the abstract results in this paper will be outlined for the two phased wavetrain case, however one should note that these results also apply to relative equilibria providing suitable modifications are made (for example, to the inner product used), which is discussed within §2.1.

The doubly periodic wavetrain (12) satisfies the PDE

$$\sum_{i=1}^{2}\big(\omega_i\mathbf{M}(\widehat{Z})+k_i\mathbf{J}(\widehat{Z})+m_i\mathbf{K}(\widehat{Z})\big)\widehat{Z}_{\theta_i}=\nabla S(\widehat{Z}). \tag{13}$$

For convenience in the later modulation reduction, we may define the linear operator about this wavetrain solution as

$$\mathbf{L}=\mathrm{D}^2S(\widehat{Z})-\sum_{i=1}^{2}\Bigg[\big(\omega_i\mathbf{M}(\widehat{Z})+k_i\mathbf{J}(\widehat{Z})+m_i\mathbf{K}(\widehat{Z})\big)\partial_{\theta_i} + \big(\omega_i\mathrm{D}\mathbf{M}(\widehat{Z})+k_i\mathrm{D}\mathbf{J}(\widehat{Z})+m_i\mathrm{D}\mathbf{K}(\widehat{Z})\big)\widehat{Z}_{\theta_i}\Bigg].$$

Immediately we see that the following hold:

$$\mathbf{L}\widehat{Z}_{\theta_i}=0,\quad \mathbf{L}\widehat{Z}_{k_i}=\mathbf{J}(\widehat{Z})\widehat{Z}_{\theta_i},\quad \mathbf{L}\widehat{Z}_{\omega_i}=\mathbf{M}(\widehat{Z})\widehat{Z}_{\theta_i},\quad \mathbf{L}\widehat{Z}_{m_i}=\mathbf{K}(\widehat{Z})\widehat{Z}_{\theta_i}\,. \tag{14}$$

The first of these suggests that the kernel of $\mathbf{L}$ contains at least two elements, which within the setting of this paper we assume that this is no larger. This means that inhomogeneous problems have the following solvability condition:

$$\mathbf{L}Q=R\quad\text{is solvable if and only if}\quad \langle\!\langle \widehat{Z}_{\theta_1},R\rangle\!\rangle=\langle\!\langle \widehat{Z}_{\theta_2},R\rangle\!\rangle=0\,, \tag{15}$$

and $\langle\!\langle\cdot,\cdot\rangle\!\rangle$ is a suitable inner product for the problem. For doubly periodic waves, this takes the form

$$\langle\!\langle V,W\rangle\!\rangle=\frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi}\langle V,W\rangle\;d\theta_1\,d\theta_2\,.$$

The remaining expressions in (14) highlight the presence of Jordan chains arising from the linearisation. The focus of this paper will be the one involving the matrix $\mathbf{J}$, as this will be the mechanism for which dispersive terms enter the phase dynamics. Details of the necessary Jordan chain theory for the modulation theory which leads to the KP equation are given in §2.3.

2.1. From Periodic Wavetrains to Relative Equilibrium

Classically, modulation analyses are undertaken for periodic wavetrains and utilise the invariance of phase translations (associated with the underlying toral symmetries) to proceed. However, it is the case that one may treat a more generic family of solutions with the modulation approach, requiring only that these are associated with a continuous symmetry group which leaves the Lagrangian invariant. Such solutions are known as *relative equilibria*, and can be thought of as a generalisation of periodic solutions. The modulation of these proceeds in an almost identical way to the wavetrain case, with a few modifications and interpretation of the solution parameters $\mathbf{k}$, $\mathbf{m}$ and $\boldsymbol{\omega}$.

For the purposes of this paper, we need only be concerned with relative equilibrium arising from continuous group actions with two (independent) parameters. These are most naturally taken to be the classical wave variables as defined in (3), so that the solutions traverse the group orbit at a constant rate in each of the spatial and time variables. Most readily, such actions are associated with some combination of affine (invariance under translation), toral (invariance under rotation) or phase translational symmetries. The case of doubly periodic wavetrains falls within the remit of combining two actions that represent phase translations, but other Lagrangians may possess different and more exotic combinations. For example, stratified shallow water systems have relative equilibrium solutions associated with a combination of affine symmetries [23] and is also the case considered in the stratified Euler equations of this paper, presented in §4. Other interesting cases include systems like the Dysthe equation [29, 30], the Benney-Roskes equation [31] or current-wave couplings [32, 33] which have mixed group actions which result from both an affine and toral symmetry being present.

Denote such actions by $G_{\theta_1,\theta_2} \in G$, for some Lie group G. We assume invariance of (11) under the actions of G, so that for each $G_{\theta_1,\theta_2} \in G$

$$G_{\theta_1,\theta_2}\mathbf{J} = \mathbf{J}G_{\theta_1,\theta_2}\,, \quad G_{\theta_1,\theta_2}\mathbf{K} = \mathbf{K}G_{\theta_1,\theta_2}\,,$$
$$G_{\theta_1,\theta_2}\mathbf{M} = \mathbf{M}G_{\theta_1,\theta_2}\,, \quad S(G_{\theta_1,\theta_2}Z) = S(Z)\,,$$

so that if Z solves (11) then so does $G_{\theta_1,\theta_2}Z$. From these notions, we may define relative equilibria solutions as

$$Z = G_{\theta_1,\theta_2}\hat{z}(\mathbf{k},\mathbf{m},\boldsymbol{\omega}) \equiv \widehat{Z}(\boldsymbol{\theta};\mathbf{k},\mathbf{m},\boldsymbol{\omega})\,,$$

for a suitable equilibrium $\hat{z}$ of (11). Thus, like periodic wavetrains, one is able to think of these solutions as equilibrium with respect to the group orbit in the same way that travelling waves are equilibrium within the moving frame. From here, the majority of the abstract set-up presented before follows through identically, with only the inner product requiring further consideration for the case of relative equilibria. For example, in the case of a doubly affine symmetry one does not need to average over each θ_i and instead the standard inner product suffices, but averaging is required for each variable that is associated with a toral symmetry. More complicated inner products may also be required for more complex systems. This is the case for the primary example this paper uses to apply the modulation theory, where an inner product which utilises depth averaging is required, which is detailed within §4.

In the setting of relative equilibria, the parameters $\mathbf{k}$, $\mathbf{m}$, $\boldsymbol{\omega}$ can respresent more general quantities aside from wavenumbers and frequencies. For example, Whitham [10, chapter 16] presents a case of multiphase modulation where the wavenumber of the second phase (denoted as the *pseudophase*) represents the mean value of the flow velocity. This is similar to the case with the example presented in this paper, where the vector $\mathbf{k}$ assumes the role of the velocity in each layer of the stratified fluid and $\boldsymbol{\omega}$ represents a quantity akin to the Bernoulli head of each flow, similar to the psuedophase used in the case of Whitham [32]. Other settings are likely to relate these parameters to traditional modulation variables, such as the amplitude of the wave (see [12] and references therein), but this is likely to depend on the problem being considered and its symmetries.

2.2. Conservation of Wave Action

Conservation laws have a substantial role in the emergence of the KP equation from modulational arguments. Not only can the criteria for the KP equation to emerge be formulated using the singularity of the wave action flux, but the coefficients of the final PDE are formed from the derivatives of the conservation laws. This occurs because the multisymplectic formulation of the Lagrangian makes an explicit connection between the structure of the Euler Lagrange equation (11) and the conservation laws for the system through the skew-symmetric matrices $\mathbf{M}$, $\mathbf{J}$ and $\mathbf{K}$.

We may extract two conservation laws associated with the conservation of wave action from the Lagrangian, one associated with each phase θ_i. The components of these take the form

$$A_i(x,y,t) = \frac{1}{2}\langle\!\langle Z, \boldsymbol{\alpha}_{\theta_i}\rangle\!\rangle\,, \quad B_i(x,y,t) = \frac{1}{2}\langle\!\langle Z, \boldsymbol{\beta}_{\theta_i}\rangle\!\rangle\,, \quad C_i(x,y,t) = \frac{1}{2}\langle\!\langle Z, \boldsymbol{\gamma}_{\theta_i}\rangle\!\rangle\,, \tag{16}$$

obtained either using multisymplectic Noether theory [34, 35] or by differentiation of the Lagrangian evaluated along the wavetrain and averaged over the phases. Herein, denote

$$A(x,y,t) = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \quad B(x,y,t) = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad C(x,y,t) = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix},$$

so we may write the conservation law

$$A_t + B_x + C_y = 0\,,$$

for any solution satisfying (11).

Central to the theory of this paper are the conservation law components evaluated for the wavetrain solution $\widehat{Z}$,

$$
\begin{array}{rcl}
\mathscr{A}_i(\mathbf{k},\mathbf{m},\boldsymbol{\omega}) & = & \frac{1}{2}\langle\!\langle \widehat{Z}, \boldsymbol{\alpha}(\widehat{Z})_{\theta_i}\rangle\!\rangle\,, \\
\mathscr{B}_i(\mathbf{k},\mathbf{m},\boldsymbol{\omega}) & = & \frac{1}{2}\langle\!\langle \widehat{Z}, \boldsymbol{\beta}(\widehat{Z})_{\theta_i}\rangle\!\rangle\,, \\
\mathscr{C}_i(\mathbf{k},\mathbf{m},\boldsymbol{\omega}) & = & \frac{1}{2}\langle\!\langle \widehat{Z}, \boldsymbol{\gamma}(\widehat{Z})_{\theta_i}\rangle\!\rangle\,,
\end{array} \tag{17}
$$

which comprise the following vectors:

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

These components then can be differentiated with respect to their parameters:

$$
\begin{aligned}
\partial_{k_j}\mathscr{A}_i =& \langle\!\langle \widehat{Z}_{k_j}, \mathbf{M}(\widehat{Z})\widehat{Z}_{\theta_i}\rangle\!\rangle = \langle\!\langle \widehat{Z}_{\omega_i}, \mathbf{J}(\widehat{Z})\widehat{Z}_{\theta_j}\rangle\!\rangle = \partial_{\omega_i}\mathscr{B}_j\,,\\
\partial_{\omega_j}\mathscr{A}_i =& \langle\!\langle \widehat{Z}_{\omega_j}, \mathbf{M}(\widehat{Z})\widehat{Z}_{\theta_i}\rangle\!\rangle\,,\\
\partial_{k_j}\mathscr{C}_i =& \langle\!\langle \widehat{Z}_{k_j}, \mathbf{K}(\widehat{Z})\widehat{Z}_{\theta_i}\rangle\!\rangle = \langle\!\langle \widehat{Z}_{m_i}, \mathbf{J}(\widehat{Z})\widehat{Z}_{\theta_j}\rangle\!\rangle = \partial_{m_i}\mathscr{B}_j\,,\\
\partial_{m_j}\mathscr{C}_i =& \langle\!\langle \widehat{Z}_{m_j}, \mathbf{K}(\widehat{Z})\widehat{Z}_{\theta_i}\rangle\!\rangle\,,\\
\partial_{m_j}\mathscr{A}_i =& \langle\!\langle \widehat{Z}_{m_j}, \mathbf{M}(\widehat{Z})\widehat{Z}_{\theta_i}\rangle\!\rangle = \langle\!\langle \widehat{Z}_{\omega_i}, \mathbf{K}(\widehat{Z})\widehat{Z}_{\theta_j}\rangle\!\rangle = \partial_{\omega_i}\mathscr{C}_j\,,\\
\partial_{k_j}\mathscr{B}_i =& \langle\!\langle \widehat{Z}_{k_j}, \mathbf{J}(\widehat{Z})\widehat{Z}_{\theta_i}\rangle\!\rangle\,,\\
\partial_{k_i}\partial_{k_j}\mathscr{B}_l =& \langle\!\langle \widehat{Z}_{k_ik_j}, \mathbf{J}(\widehat{Z})\widehat{Z}_{\theta_l}\rangle\!\rangle + \langle\!\langle \widehat{Z}_{k_i}, \mathbf{J}(\widehat{Z})\widehat{Z}_{\theta_l k_j} + \mathrm{D}\mathbf{J}(\widehat{Z})(\widehat{Z}_{\theta_l}, \widehat{Z}_{k_j})\rangle\!\rangle.
\end{aligned}
$$

and overall gives the tensors to be considered as

$$
\begin{array}{rclrcl}
\mathrm{D}_{\mathbf{k}}\mathbf{A} & = & \begin{pmatrix}\partial_{k_1}\mathscr{A}_1 & \partial_{k_2}\mathscr{A}_1\\ \partial_{k_1}\mathscr{A}_2 & \partial_{k_2}\mathscr{A}_2\end{pmatrix} = \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B}^T\,, & \mathrm{D}_{\mathbf{k}}\mathbf{B} & = & \begin{pmatrix}\partial_{k_1}\mathscr{B}_1 & \partial_{k_2}\mathscr{B}_1\\ \partial_{k_1}\mathscr{B}_2 & \partial_{k_2}\mathscr{B}_2\end{pmatrix}\,,\\
\mathrm{D}_{\mathbf{k}}\mathbf{C} & = & \begin{pmatrix}\partial_{k_1}\mathscr{C}_1 & \partial_{k_2}\mathscr{C}_1\\ \partial_{k_1}\mathscr{C}_2 & \partial_{k_2}\mathscr{C}_2\end{pmatrix} = \mathrm{D}_{\mathbf{m}}\mathbf{B}^T\,, & \mathrm{D}_{\boldsymbol{\omega}}\mathbf{A} & = & \begin{pmatrix}\partial_{\omega_1}\mathscr{A}_1 & \partial_{\omega_2}\mathscr{A}_1\\ \partial_{\omega_1}\mathscr{A}_2 & \partial_{\omega_2}\mathscr{A}_2\end{pmatrix}\,,\\
\mathrm{D}_{\mathbf{m}}\mathbf{A} & = & \begin{pmatrix}\partial_{m_1}\mathscr{A}_1 & \partial_{m_2}\mathscr{A}_1\\ \partial_{m_1}\mathscr{A}_2 & \partial_{m_2}\mathscr{A}_2\end{pmatrix} = \mathrm{D}_{\boldsymbol{\omega}}\mathbf{C}^T\,, & \mathrm{D}_{\mathbf{m}}\mathbf{C} & = & \begin{pmatrix}\partial_{m_1}\mathscr{C}_1 & \partial_{m_2}\mathscr{C}_1\\ \partial_{m_1}\mathscr{C}_2 & \partial_{m_2}\mathscr{C}_2\end{pmatrix}\,,
\end{array}
$$

as well as the tensor

$$
\mathrm{D}^2_{\mathbf{k}}\mathbf{B} = \left[\begin{pmatrix}\partial_{k_1k_1}\mathscr{B}_1 & \partial_{k_1k_2}\mathscr{B}_1\\ \partial_{k_1k_1}\mathscr{B}_2 & \partial_{k_1k_2}\mathscr{B}_2\end{pmatrix}\middle|\begin{pmatrix}\partial_{k_1k_2}\mathscr{B}_1 & \partial_{k_2k_2}\mathscr{B}_1\\ \partial_{k_1k_2}\mathscr{B}_2 & \partial_{k_2k_2}\mathscr{B}_2\end{pmatrix}\right]
$$

We are now in a position to introduce the notion of criticality, which is fundamental to the emergence of nonlinearity and dispersion from the phase dynamics. We say that a conservation law component is critical when it fails to be injective with respect to one of its arguments. The particular case this paper concerns itself with is when the wave action flux $\mathbf{B}$ is no longer injective in $\mathbf{k}$. This means that there exists some curve $\mathbf{k} = \mathbf{k}_0(\mathbf{m},\boldsymbol{\omega})$ along which one has

$$
\det\left[\mathrm{D}_{\mathbf{k}}\mathbf{B}\right] = 0\,. \tag{18}
$$

The above condition can be shown to be the criterion for the emergence of a zero characteristic from the linear Whitham equations associated with the Lagrangian considered [24], but also has recently been shown to relate to stability boundaries emerging from physical problems [19, 20, 21]. Thus, there are both abstract and physical interpretations for this condition. Throughout this paper it will be assumed that the zero eigenvalue arising from (18) is simple, so that we may introduce a unique eigenvector $\boldsymbol{\zeta}$ associated with the zero eigenvalue, with

$$\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\zeta} = \mathbf{0}\,. \tag{19}$$

This eigenvector has a central role in the modulation theory, both within the modulation procedure as well as in the generation of the coefficients of the resulting KP equation. This differs from the modulation of single phase wavetrains, where no such term emerges from the corresponding scalar criticality, and presents one of the novel components within the modulation of multiphase wavetrains.

Due to the presence of an additional spatial dimension, there is a second condition which arises from the reduction procedure that must also be satisfied. This is related to the conservation law component $\mathbf{C}$, and takes the form

$$\boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{C} + \mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\zeta} = 0\,, \tag{20}$$

which may be understood as the condition for the linear system

$$\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\eta} = (\mathrm{D}_{\mathbf{k}}\mathbf{C} + \mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\zeta}\,,$$

to be solvable. The fact that one must impose an additional condition in order for the KP to emerge seems surprising, given that only one is required for the KdV to arise, thus the KP equation is expected to be just as prevalent. It transpires, however, that one can show that if the governing equations are transverse reversible (that is, invariant under the mapping $y \mapsto -y$) it is sufficient to choose $\mathbf{m} = \mathbf{0}$ in order to satisfy (20). The argument illustrating this adapts the reasoning from §7 in [28] to the multiphase case. Start by introducing a reversor symmetry operator $\mathbf{R}$, defined as

$$\mathbf{R}Z(x, y, t) = Z(x, -y, t)$$

and is isometric and an involution. Supposing the system is transverse reversible in the y direction means that whenever $Z(x, y, t)$ is a solution of 11), then so is $Z(x, -y, t)$. It follows that this reversor must have the properties

$$\mathbf{MR} = \mathbf{RM}, \quad \mathbf{JR} = \mathbf{RJ}, \quad \mathbf{RK} = -\mathbf{KR}, \quad \nabla S(\mathbf{R}Z) = \mathbf{R}\nabla S(Z).$$

This means that whenever Z is a solution, $\mathbf{R}Z(x, -y, t)$ is also a solution as expected. An immediate implication of this is that $\mathbf{R}\widehat{Z}(\boldsymbol{\theta}; \mathbf{k}, \mathbf{m}, \boldsymbol{\omega}) =$

$\widehat{Z}(\boldsymbol{\theta};\,\mathbf{k},-\mathbf{m},\boldsymbol{\omega})$. Now we observe how $\mathscr{C}_i$ is affected under the reversor symmetry:

$$\begin{aligned}\mathscr{C}_i(\mathbf{k},-\mathbf{m},\boldsymbol{\omega}) =&\frac{1}{2}\langle\!\langle \mathbf{K}\widehat{Z}_{\theta_i}(\boldsymbol{\theta};\mathbf{k},-\mathbf{m},\boldsymbol{\omega}),\widehat{Z}(\boldsymbol{\theta};\mathbf{k},-\mathbf{m},\boldsymbol{\omega})\rangle\!\rangle\\ =&\frac{1}{2}\langle\!\langle \mathbf{K}\mathbf{R}\widehat{Z}_{\theta_i}(\boldsymbol{\theta};\mathbf{k},\mathbf{m},\boldsymbol{\omega}),\widehat{Z}(\boldsymbol{\theta};\mathbf{k},-\mathbf{m},\boldsymbol{\omega})\rangle\!\rangle\\ =&-\frac{1}{2}\langle\!\langle \mathbf{R}\mathbf{K}\widehat{Z}_{\theta_i}(\boldsymbol{\theta};\mathbf{k},\mathbf{m},\boldsymbol{\omega}),\widehat{Z}(\boldsymbol{\theta};\mathbf{k},-\mathbf{m},\boldsymbol{\omega})\rangle\!\rangle\\ =&-\frac{1}{2}\langle\!\langle \mathbf{K}\widehat{Z}_{\theta_i}(\boldsymbol{\theta};\mathbf{k},\mathbf{m},\boldsymbol{\omega}),\mathbf{R}\widehat{Z}(\boldsymbol{\theta};\mathbf{k},-\mathbf{m},\boldsymbol{\omega})\rangle\!\rangle\\ =&-\frac{1}{2}\langle\!\langle \mathbf{K}\widehat{Z}_{\theta_i}(\boldsymbol{\theta};\mathbf{k},\mathbf{m},\boldsymbol{\omega}),\widehat{Z}(\boldsymbol{\theta};\mathbf{k},\mathbf{m},\boldsymbol{\omega})\rangle\!\rangle = -\mathscr{C}_i(\mathbf{k},\mathbf{m},\boldsymbol{\omega})\end{aligned}$$

Thus, whenever a reversor symmetry is present, $\mathbf{C}$ is an odd function of $\mathbf{m}$. A similar calculation verifies that $\mathscr{A}_i,\mathscr{B}_i$ are even in $\mathbf{m}$. A consequence of this is that the parameter derivatives of $\mathscr{C}_i$ other than those with respect to m_i are automatically zero when $\mathbf{m}=\mathbf{0}$. Consequently, it follows that

$$\mathrm{D}_{\mathbf{k}}\mathbf{C}\big|_{\mathbf{m}=\mathbf{0}} = \mathbf{0}\,. \tag{21}$$

Therefore, the condition (20) will automatically hold at $\mathbf{m}=\mathbf{0}$ whenever the system is transverse reversible.

2.3. Review of the Jordan Chain Theory for Multiple Phases

It is clear from (14) that multiple Jordan chains emerge from the linearisation about the wavetrain solution $\widehat{Z}$, each associated with one of the skew-symmetric matrices. The one of importance within this paper is the one involving $\mathbf{J}$, which has the form

$$\mathbf{L}\xi_1 = 0,\quad \mathbf{L}\xi_i = \mathbf{J}\xi_{i-1},\ i>1.$$

Two such chains exist, each starting with a θ_i derivative and are followed by the respective k_i derivative. We denote these in the following way:

$$\xi_1 = \widehat{Z}_{\theta_1}\,,\quad \xi_2 = \widehat{Z}_{k_1}\,,\quad \xi_3 = \widehat{Z}_{\theta_2}\,,\quad \xi_4 = \widehat{Z}_{k_2}\,,$$

so that the first two indices form the first chain, and the latter two the second. It will be made clear within the modulation analysis that these two chains must coalesce in order for dispersion to arise and the KP equation emerge. Consider the equation

$$\mathbf{L}\xi_5 = \sum_{i=1}^{2}\zeta_i\mathbf{J}\widehat{Z}_{k_i}\,. \tag{22}$$

Assessing the solvability of the above according to (15) generates the system

$$\begin{pmatrix} \langle\!\langle \widehat{Z}_{\theta_1}, \mathbf{J}\widehat{Z}_{k_1} \rangle\!\rangle & \langle\!\langle \widehat{Z}_{\theta_1}, \mathbf{J}\widehat{Z}_{k_2} \rangle\!\rangle \\ \langle\!\langle \widehat{Z}_{\theta_2}, \mathbf{J}\widehat{Z}_{k_1} \rangle\!\rangle & \langle\!\langle \widehat{Z}_{\theta_2}, \mathbf{J}\widehat{Z}_{k_2} \rangle\!\rangle \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \equiv - \begin{pmatrix} \partial_{k_1}\mathscr{B}_1 & \partial_{k_2}\mathscr{B}_1 \\ \partial_{k_1}\mathscr{B}_2 & \partial_{k_2}\mathscr{B}_2 \end{pmatrix} \boldsymbol{\zeta} \equiv -\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\zeta} = \mathbf{0}\,. \tag{23}$$

Therefore, if $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ has a zero eigenvalue with associated eigenvector $\boldsymbol{\zeta}$, which is precisely when (18) holds, then one is able to solve (22). This allows the definition of ξ_5 with

$$\mathbf{L}\xi_5 = \sum_{i=1}^{2} \zeta_i \mathbf{J}\widehat{Z}_{k_i}.$$

The other eigenvalue for the system (23) is given by the trace of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$, which results in the eigenvalue problem

$$\mathrm{D}_{\mathbf{k}}\mathbf{B} \begin{pmatrix} \zeta_2 \\ -\zeta_1 \end{pmatrix} = (\partial_{k_1}\mathscr{B}_1 + \partial_{k_2}\mathscr{B}_2) \begin{pmatrix} \zeta_2 \\ -\zeta_1 \end{pmatrix}.$$

It follows that the equation

$$\mathbf{L}F = \zeta_2 \mathbf{J}\widehat{Z}_{k_1} - \zeta_1 \mathbf{J}\widehat{Z}_{k_2}\,,$$

cannot be solved for, due to the assumption that the zero eigenvalue of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ is simple.

One may show that the zero eigenvalue of $\mathbf{L}$ is of even algebraic multiplicity [36, chapter 3], and as a consequence the existence of ξ_5 guarantees the existence of ξ_6 with

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

Within this paper we will assume that this chain is no longer. This imposes that

$$\mathbf{L}\xi_7 = \mathbf{J}\xi_6\,,$$

is not solvable. This facilitates the definition of the nontrivial vector

$$\mathscr{K} = \begin{pmatrix} \mathscr{K}_1 \\ \mathscr{K}_2 \end{pmatrix} := - \begin{pmatrix} \langle\!\langle \widehat{Z}_{\theta_i}, \mathbf{J}\xi_6 \rangle\!\rangle \\ \langle\!\langle \widehat{Z}_{\theta_2}, \mathbf{J}\xi_6 \rangle\!\rangle \end{pmatrix} \neq \mathbf{0}\,. \tag{24}$$

This vector arises directly from the modulation theory and ultimately forms the dispersive term in X appearing in (33).

3. Modulation Leading to the KP equation

The strategy to obtain the KP equation in this paper is to utilise the method of phase modulation. In particular we use the multiphase method developed by Bridges and Ratliff [23, 24, 25], which itself is an extension of the approach used by Kuramoto [14, 15], Doelman et al. [16] and Bridges [17, 28] for single phase wavetrains. The ansatz taken to obtain the KP equation is

$$\begin{aligned} Z(x,y,t) = \widehat{Z}\big(&\boldsymbol{\theta} + \varepsilon\boldsymbol{\phi}(X,Y,T;\varepsilon) + \varepsilon^2\boldsymbol{\psi}(X,Y,T;\varepsilon);\\ &\mathbf{k} + \varepsilon^2\mathbf{q}(X,Y,T;\varepsilon) + \varepsilon^3\mathbf{p}(X,Y,T;\varepsilon),\\ &\mathbf{m} + \varepsilon^3\mathbf{r}(X,Y,T;\varepsilon) + \varepsilon^4\mathbf{s}(X,Y,T;\varepsilon),\\ &\boldsymbol{\omega} + \varepsilon^4\boldsymbol{\Omega}(X,Y,T;\varepsilon) + \varepsilon^5\boldsymbol{\Gamma}(X,Y,T;\varepsilon)\big)\\ &\quad + \varepsilon^3 W(\theta,X,Y,T;\varepsilon)\,, \end{aligned} \tag{25}$$

for $X = \varepsilon x,\ Y = \varepsilon^2 y,\ T = \varepsilon^3 t$ with the following phase relations:

$$\begin{gathered} \boldsymbol{\phi}_X = \mathbf{q}\,, \quad \boldsymbol{\phi}_Y = \mathbf{r}\,, \quad \boldsymbol{\phi}_T = \boldsymbol{\Omega}\,, \quad \Rightarrow \quad \mathbf{q}_T = \boldsymbol{\Omega}_X\,, \quad \mathbf{q}_Y = \mathbf{r}_X\,, \quad \mathbf{r}_T = \boldsymbol{\Omega}_Y\,,\\ \boldsymbol{\psi}_X = \mathbf{p}\,, \quad \boldsymbol{\psi}_Y = \mathbf{s}\,, \quad \boldsymbol{\psi}_T = \boldsymbol{\Gamma}\,, \quad \Rightarrow \quad \mathbf{p}_T = \boldsymbol{\Gamma}_X\,, \quad \mathbf{p}_Y = \mathbf{s}_X\,, \quad \mathbf{s}_T = \boldsymbol{\Gamma}_Y\,. \end{gathered} \tag{26}$$

The first set of these, related to $\boldsymbol{\phi}$ are the typical modulational functions that arise from this approach. The second of these, the ones involving $\boldsymbol{\psi}$, are required in order to resolve the secondary criticality (20). Ultimately, these will become related to the $\boldsymbol{\phi}$ family of modulation functions via the condition (20), which will be made apparent in the analysis.

This ansatz and the resulting approach may seem unnecessarily complex, but it allows for several simplifications within the analysis. Firstly, by including the unknown functions as perturbations to the wave variables, one naturally introduces derivatives of $\widehat{Z}$ into the analysis, which leads to the cancellation of several unimportant terms which do not contribute to the analysis. Secondly, this ansatz leads to both the key solvability criteria and the coefficients of the resulting nonlinear PDE to be related to derivatives of the conservation law components, allowing these computations to be done in advance. Finally, although there are initially two unknown vector valued functions $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ within the analysis, the solvability requirements (18) and (20) will reduce this to a problem in some scalar valued unknown function over the course of the reduction.

The idea is to substitute the ansatz (25) into the Euler Lagrange equation (11) and Taylor expand around the $\varepsilon = 0$ state to generate a system of equations to solve at each order of ε. For convenience, we also expand W in a simple asymptotic series,

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

Below is the summary of the analysis, ignoring the leading order problem which simply recovers (13).

First order

This simply reads

$$\sum_{i=1}^{2}\phi_i\mathbf{L}\widehat{Z}_{\theta_i}=0,$$

which is true, as $\widehat{Z}_{\theta_i}\in\ker(\mathbf{L})$.

Second Order

The system at this order is

$$\sum_{i=1}^{2}\left(q_i\mathbf{L}\widehat{Z}_{k_i}-(\phi_i)_X\mathbf{J}(\widehat{Z})\widehat{Z}_{\theta_i}\right)=0$$

which is true by an earlier stated linear operator result in (14), as well as recalling from (26) that $q_i=(\phi_i)_X$.

Third Order

At this order, we obtain

$$\mathbf{L}W_0=\sum_{i=1}^{2}(q_i)_X\mathbf{J}(\widehat{Z})\widehat{Z}_{k_i}.$$

This system is solvable precisely when

$$\begin{pmatrix}\langle\!\langle\widehat{Z}_{\theta_1},\mathbf{J}(\widehat{Z})\widehat{Z}_{k_1}\rangle\!\rangle & \langle\!\langle\widehat{Z}_{\theta_1},\mathbf{J}(\widehat{Z})\widehat{Z}_{k_2}\rangle\!\rangle\\ \langle\!\langle\widehat{Z}_{\theta_2},\mathbf{J}(\widehat{Z})\widehat{Z}_{k_1}\rangle\!\rangle & \langle\!\langle\widehat{Z}_{\theta_2},\mathbf{J}(\widehat{Z})\widehat{Z}_{k_2}\rangle\!\rangle\end{pmatrix}\mathbf{q}_X\equiv-\mathrm{D}_{\mathbf{k}}\mathbf{B}\mathbf{q}_X=0\,. \tag{27}$$

So that this may be satisfied, we require that

$$\det\big[\mathrm{D}_{\mathbf{k}}\mathbf{B}\big]=0\,,\quad \mathbf{q}=\boldsymbol{\zeta}U\quad\text{with}\quad \mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\zeta}=\mathbf{0}\,,$$

which is seen by appealing to (18) and (19). When the above holds, we may solve the system at this order, giving that

$$W_0=U_X\xi_5,\quad\text{with}\quad \mathbf{L}\xi_5=\sum_{i=1}^{2}\zeta_i\mathbf{J}(\widehat{Z})\widehat{Z}_{k_i}.$$

Fourth Order

After simplifying, one finds the equation at this order is given by

$$\mathbf{L}\Big(W_1 - \sum_{i=1}^{2}\phi_i U_X \partial_{\theta_i}\xi_5\Big) = \sum_{i=1}^{2}\big((\phi_i)_T\mathbf{M}(\widehat{Z})\widehat{Z}_{\theta_i} - \Omega_i\mathbf{L}\widehat{Z}_{\omega_i}\big) + U_{XX}\mathrm{D}\mathbf{J}\xi_3 + U_Y\sum_{i=1}^{2}\zeta_i(\mathbf{K}(\widehat{Z})\widehat{Z}_{k_i} + \mathbf{J}(\widehat{Z})\widehat{Z}_{m_i}) + \sum_{i=1}^{2}(p_i)_X\mathbf{J}(\widehat{Z})\widehat{Z}_{k_i}\,. \tag{28}$$

The U_{XX} component of this equation lies in the range of $\mathbf{L}$ as the zero eigenvalue of $\mathbf{L}$ is of even algebraic multiplicity. The ϕ_T term cancels with the Ω term, and so in order for the remaining terms to also lie in the range of $\mathbf{L}$ it is required that

$$-\mathrm{D}_{\mathbf{k}}\mathbf{B}\mathbf{p}_X + \begin{pmatrix}\langle\!\langle \widehat{Z}_{\theta_1}, \mathbf{K}(\widehat{Z})\widehat{Z}_{k_1} + \mathbf{J}(\widehat{Z})\widehat{Z}_{m_1}\rangle\!\rangle & \langle\!\langle \widehat{Z}_{\theta_1}, \mathbf{K}(\widehat{Z})\widehat{Z}_{k_2} + \mathbf{J}(\widehat{Z})\widehat{Z}_{m_2}\rangle\!\rangle \\ \langle\!\langle \widehat{Z}_{\theta_2}, \mathbf{K}(\widehat{Z})\widehat{Z}_{k_1} + \mathbf{J}(\widehat{Z})\widehat{Z}_{m_1}\rangle\!\rangle & \langle\!\langle \widehat{Z}_{\theta_2}, \mathbf{K}(\widehat{Z})\widehat{Z}_{k_2} + \mathbf{J}(\widehat{Z})\widehat{Z}_{m_2}\rangle\!\rangle\end{pmatrix}\boldsymbol{\zeta} \equiv -\mathrm{D}_{\mathbf{k}}\mathbf{B}\mathbf{p}_X - (\mathrm{D}_{\mathbf{k}}\mathbf{C} + \mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\zeta}U_Y = \mathbf{0}\,.$$

This matrix system is only solvable when

$$\boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{C} + \mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\zeta} = 0\,. \tag{29}$$

Thus the second condition for the KP to arise emerges from the modulation approach naturally, which can be met automatically by choosing $\mathbf{m} = \mathbf{0}$ as illustrated in §2.2. This additionally imposes that

$$\mathbf{p}_X = -\boldsymbol{\eta}U_Y \qquad \text{with} \qquad \mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\eta} = (\mathrm{D}_{\mathbf{k}}\mathbf{C} + \mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\zeta}\,. \tag{30}$$

Thus, all the modulation functions are now related to the scalar unknown U. As discussed in §2.2, when a transverse reversor symmetry is present whenever $\mathbf{m} = \mathbf{0}$, one can choose $\boldsymbol{\eta} = \mathbf{0}$ and so $\mathbf{p}$ terms vanish. In cases where (29) is satisfied, the equation (28) is solvable and we have

$$W_1 = \sum_{i=1}^{2}\phi_i U_X\partial_{\theta_i}\xi_5 + U_{XX}\xi_6\,, \quad \text{with} \quad \mathbf{L}\xi_6 = \mathbf{J}(\widehat{Z})\xi_5\,.$$

Fifth Order

The final order reads

$$
\begin{aligned}
\mathbf{L}\widetilde{W}_2 \;=\; & U_T \textstyle\sum_{i=1}^2 \zeta_i\big(\mathbf{M}(\widehat{Z})\widehat{Z}_{k_i} + \mathbf{J}(\widehat{Z})\widehat{Z}_{\omega_i}\big) + U_{XXX}\mathbf{J}(\widehat{Z})\xi_6 \\
& + UU_X \textstyle\sum_{i=1}^2 \zeta_i\Big[\big(\mathbf{J}(\widehat{Z})(\xi_5)_{\theta_i} + \mathrm{D}\mathbf{J}(\widehat{Z})(\xi_5,\widehat{Z}_{\theta_i})\\
& + \textstyle\sum_{j=1}^2 \big(k_j \mathrm{D}^2\mathbf{J}(\widehat{Z})(\widehat{Z}_{\theta_j},\xi_3,\widehat{Z}_{k_i}) + \omega_j \mathrm{D}^2\mathbf{M}(\widehat{Z})(\widehat{Z}_{\theta_j},\xi_3,\widehat{Z}_{k_i})\big)\\
& - \mathrm{D}^3 S(\widehat{Z})(\widehat{Z}_{k_i},\xi_5) + \textstyle\sum_{j=1}^2 \zeta_j\big(\mathbf{J}(\widehat{Z})\widehat{Z}_{k_ik_j} + \mathrm{D}\mathbf{J}(\widehat{Z})(\widehat{Z}_{k_i},\widehat{Z}_{k_j})\big)\Big]\\
& + \textstyle\sum_{i=1}^2 (r_i)_Y \mathbf{K}(\widehat{Z})\widehat{Z}_{m_i} + \sum_{i=1}^2(\alpha_i)_{XX}\mathbf{J}\widehat{Z}_{k_i}\,.
\end{aligned}
\tag{31}
$$

The expression $\widetilde{W}_2$ is the sum of the function W_2 from the expansion of W with a collection of the preimages of the terms which lie in the range of $\mathbf{L}$ at this order. The explicit expression for $\widetilde{W}_2$ is no longer important, since the analysis terminates at this order, but would be necessary if further orders are investigated. All that remains is to determine the solvability of the system at this order using (15). This generates a vector equation which can then be projected using $\boldsymbol{\zeta}$ and will ultimately result in U satisfying a KP equation.

The solvability condition at this order is computed by taking the inner product of the right hand side of (31) with each kernel element $\widehat{Z}_{\theta_i}$; this results in a sequence of vectors being generated whose sum must vanish. The first vector we compute is the one which prefactors the U_T term. This is simply

$$
\begin{aligned}
\langle\!\langle \widehat{Z}_{\theta_p}, \mathbf{J}(\widehat{Z})\widehat{Z}_{\omega_i} + \mathbf{M}(\widehat{Z})\widehat{Z}_{k_i}\rangle\!\rangle &= -\langle\!\langle \widehat{Z}_{\omega_i}, \mathbf{J}(\widehat{Z})\widehat{Z}_{\theta_p}\rangle\!\rangle - \langle\!\langle \widehat{Z}_{k_i}, \mathbf{M}(\widehat{Z})\widehat{Z}_{\theta_p}\rangle\!\rangle\\
&= -\partial_{\omega_i}\mathscr{B}_p - \partial_{k_i}\mathscr{A}_p.
\end{aligned}
$$

This forms the vector

$$
-\begin{pmatrix} \sum_{i=1}^2 \zeta_i\big(\partial_{\omega_i}\mathscr{B}_1 + \partial_{k_i}\mathscr{A}_1\big)\\ \sum_{i=1}^2 \zeta_i\big(\partial_{\omega_i}\mathscr{B}_2 + \partial_{k_i}\mathscr{A}_2\big)\end{pmatrix} U_T \equiv -(\mathrm{D}_{\mathbf{k}}\mathbf{A} + \mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta} U_T\,.
$$

The vector of the quadratic term is sought next, giving

$$
\begin{aligned}
&\langle\!\langle \widehat{Z}_{\theta_p}, \big[\mathbf{J}(\widehat{Z})(\xi_5)_{\theta_i} + \mathrm{D}\mathbf{J}(\widehat{Z})(\xi_5,\widehat{Z}_{\theta_i}) - \mathrm{D}^3 S(\widehat{Z})(\widehat{Z}_{k_i},\xi_5) \\
&\quad + \sum_{j=1}^{2}\big(k_j \mathrm{D}^2\mathbf{J}(\widehat{Z})(\widehat{Z}_{\theta_j},\xi_3,\widehat{Z}_{k_i}) + \omega_j \mathrm{D}^2\mathbf{M}(\widehat{Z})(\widehat{Z}_{\theta_j},\xi_3,\widehat{Z}_{k_i})\big) \\
&\quad + \sum_{j=1}^{2}\zeta_j\big(\mathbf{J}(\widehat{Z})\widehat{Z}_{k_ik_j} + \mathrm{D}\mathbf{J}(\widehat{Z})(\widehat{Z}_{k_i},\widehat{Z}_{k_j})\big]\rangle\!\rangle \\
&= \sum_{j=1}^{2}\zeta_j\langle\!\langle \widehat{Z}_{\theta_p},(\boldsymbol{\beta}(\widehat{Z}))_{k_ik_j}\rangle\!\rangle + \langle\!\langle \xi_3, \mathbf{L}\widehat{Z}_{\theta_pk_i}\rangle\!\rangle \\
&= \sum_{j=1}^{2}\zeta_j\big(\langle\!\langle \widehat{Z}_{\theta_p},(\boldsymbol{\beta}(\widehat{Z}))_{k_ik_j}\rangle\!\rangle + \langle\!\langle \mathbf{J}(\widehat{Z})\widehat{Z}_{k_j},\widehat{Z}_{\theta_pk_i}\rangle\!\rangle\big) \\
&= \sum_{j=1}^{2}\zeta_i\partial_{k_i}(\langle\!\langle \widehat{Z}_{\theta_p},\mathbf{J}(\widehat{Z})\widehat{Z}_{k_j}\rangle\!\rangle) = -\sum_{j=1}^{2}\zeta_i\partial_{k_ik_j}\mathscr{B}_p\,.
\end{aligned}
$$

Therefore, the vector associated with the quadratic term UU_X is

$$
-\begin{pmatrix}\sum_{i,j=1}^{2}\zeta_i\zeta_j\partial_{k_ik_j}\mathscr{B}_1\\ \sum_{i,j=1}^{2}\zeta_i\zeta_j\partial_{k_ik_j}\mathscr{B}_2\end{pmatrix}UU_X \equiv -\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta})UU_X\,. \tag{32}
$$

The elements of the vector multiplying the third order dispersive term are defined as

$$
\mathscr{K}_p = -\langle\!\langle \widehat{Z}_{\theta_p},\mathbf{J}(\widehat{Z})\xi_4\rangle\!\rangle\,,
$$

coming to form the vector

$$
\begin{pmatrix}\langle\!\langle \widehat{Z}_{\theta_1},\mathbf{J}(\widehat{Z})\xi_4\rangle\!\rangle\\ \langle\!\langle \widehat{Z}_{\theta_2},\mathbf{J}(\widehat{Z})\xi_4\rangle\!\rangle\end{pmatrix}U_{XXX} \equiv -\begin{pmatrix}\mathscr{K}_1\\ \mathscr{K}_2\end{pmatrix}U_{XXX} = -\mathscr{K}U_{XXX}
$$

The $(r_i)_Y$ terms will have coefficients of the form

$$
\langle\!\langle \widehat{Z}_{\theta_p},\mathbf{K}(\widehat{Z})\widehat{Z}_{m_i}\rangle\!\rangle = -\partial_{m_i}\mathscr{C}_p
$$

giving the vector

$$
-\begin{pmatrix}\sum_{i=1}^{2}(r_i)_Y\langle\!\langle \widehat{Z}_{\theta_1},\mathbf{K}(\widehat{Z})\widehat{Z}_{m_i}\rangle\!\rangle\\ \sum_{i=1}^{2}(r_i)_Y\langle\!\langle \widehat{Z}_{\theta_2},\mathbf{K}(\widehat{Z})\widehat{Z}_{m_i}\rangle\!\rangle\end{pmatrix} \equiv -\mathrm{D}_{\mathbf{m}}\mathbf{C}\mathbf{r}_Y\,.
$$

Similarly, one finds that the $(p_i)_Y$ terms generate

$$
\langle\!\langle \widehat{Z}_{\theta_p},\mathbf{K}(\widehat{Z})\widehat{Z}_{k_i} + \mathbf{J}(\widehat{Z})\widehat{Z}_{m_i}\rangle\!\rangle = -\partial_{k_i}\mathscr{C}_p - \partial_{m_i}\mathscr{B}_p\,.
$$

leading to the vector

$$
-(\mathrm{D}_{\mathbf{k}}\mathbf{C} + \mathrm{D}_{\mathbf{m}}\mathbf{B})\mathbf{p}_Y\,.
$$

Finally, the terms involving the components of $\boldsymbol{\alpha}$ gives

$$-\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\alpha}_{XX}\,,$$

by using (27). These results come together to form the vector PDE

$$(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta}U_T+\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta})UU_X+\mathscr{K}U_{XXX}+\mathrm{D}_{\mathbf{m}}\mathbf{C}\mathbf{r}_Y \\ +(\mathrm{D}_{\mathbf{k}}\mathbf{C}+\mathrm{D}_{\mathbf{m}}\mathbf{B})\mathbf{p}_Y+\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\alpha}_{XX}=\mathbf{0}\,.$$

This may be written entirely in terms of U by differentiating the system with respect to X and noting that from (26) we have that $\mathbf{r}_{XY}=\mathbf{q}_{YY}=\boldsymbol{\zeta}U_{YY}$ as well as $\mathbf{p}_{XY}=-\boldsymbol{\eta}U_{YY}$ from (30). Overall, this gives that

$$\Big[(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta}U_T+\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta})UU_X+\mathscr{K}U_{XXX}\Big]_X \\ +\big(\mathrm{D}_{\mathbf{m}}\mathbf{C}-(\mathrm{D}_{\mathbf{k}}\mathbf{C}+\mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\eta}\big)\boldsymbol{\zeta}U_{YY}+\mathrm{D}_{\mathbf{k}}\mathbf{B}\boldsymbol{\alpha}_{XXX}=\mathbf{0}\,.$$

This system is a set of two inhomogeneous equations (since $\boldsymbol{\alpha}$ is an arbitrary function) governing the evolution of U. It is the case however that this inhomogeneity can be removed simply by multiplying by $\boldsymbol{\zeta}^T$ on the left (which is akin to projecting the above in the direction of the kernel of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$). Doing so gives rise to the scalar KP equation:

$$\big(a_0U_T+a_1UU_X+a_2U_{XXX}\big)_X+a_3U_{YY}=0\,, \tag{33}$$

with the coefficients

$$\begin{aligned} a_0 &= \boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta} \\ a_1 &= \boldsymbol{\zeta}^T\mathrm{D}^2_{\mathbf{k}}\mathbf{B}(\boldsymbol{\zeta},\boldsymbol{\zeta}) \\ a_2 &= \boldsymbol{\zeta}^T\mathscr{K}\,, \\ a_3 &= \boldsymbol{\zeta}^T\big(\mathrm{D}_{\mathbf{m}}\mathbf{C}-(\mathrm{D}_{\mathbf{k}}\mathbf{C}+\mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\eta}\big)\boldsymbol{\zeta}\,. \end{aligned}$$

4. Application to the Full Stratified Water Wave Problem

One interesting application of the modulation theory presented here is the emergence of the KP equation from the full water wave problem with two stratified fluid layers. It may appear surprising at first that the theory of this paper is applicable here, however once the inherent double affine symmetry of the system is identified then this system becomes a natural candidate for the theory to be applied to. For this example, we modulate the state of uniform flow within each layer, which is the relative equilibrium associated with the double affine symmetry. As such, the theory of the paper goes through intact by replacing the θ_i averaging inner product with an inner product with depth averaging instead.

The stratified water wave problem in the absence of surface tension effects possesses the following Lagrangian [37],

$$\widetilde{\mathscr{L}} = \iint \Bigg[\rho_1 \int_0^{\tilde{h}} \big(\tilde{\phi}_{\tilde{t}} + \frac{1}{2}(\tilde{\phi}_{\tilde{x}}^2 + \tilde{\phi}_{\tilde{y}}^2 + \tilde{\phi}_{\tilde{z}}^2) + g\tilde{z} - R_1\big)\, d\tilde{z} + \rho_2 \int_{\tilde{h}}^{\tilde{H}} \big(\tilde{\varphi}_{\tilde{t}} + \frac{1}{2}(\tilde{\varphi}_{\tilde{x}}^2 + \tilde{\varphi}_{\tilde{y}}^2 + \tilde{\varphi}_{\tilde{z}}^2) + g\tilde{z} - R_2\big)\, d\tilde{z}\Bigg] d\tilde{x}\, d\tilde{y}\, d\tilde{t}\,, \quad (34)$$

which is obtained by using Luke's Lagrangian [38] for the fluid in each layer. Thus, this is the Lagrangian for the full water wave problem with two layers of constant density, indexed by $i = 1, 2$ with layer 1 residing beneath layer 2. For stable stratification, we require that the densities ρ_i satisfy $\rho_2 < \rho_1$. The velocity potentials are denoted by $\tilde{\phi}$ in the lower layer and $\tilde{\varphi}$ in the upper layer, and the free surfaces are given by $h(x, y, t)$ and $H(x, y, t)$. The interface between the two fluids is located at $z = h$ and the uppermost free surface at $z = H$. The equations of motion for the problem are obtained by taking variations of the above Lagrangian, and result in the system

$$\rho_1\left(\tilde{\phi}_{\tilde{t}} + \frac{1}{2}(\tilde{\phi}_{\tilde{x}}^2 + \tilde{\phi}_{\tilde{y}}^2 + \tilde{\phi}_{\tilde{z}}^2) + gh - R_1\right) = \rho_2\left(\tilde{\varphi}_{\tilde{t}} + \frac{1}{2}(\tilde{\varphi}_{\tilde{x}}^2 + \tilde{\varphi}_{\tilde{y}}^2 + \tilde{\varphi}_{\tilde{z}}^2) + gh - R_2\right),$$

$$\rho_2\left(\tilde{\varphi}_{\tilde{t}} + \frac{1}{2}(\tilde{\varphi}_{\tilde{x}}^2 + \tilde{\varphi}_{\tilde{y}}^2 + \tilde{\varphi}_{\tilde{z}}^2) + gH - R_2\right) = 0\,,$$

$$\tilde{\phi}_{\tilde{x}\tilde{x}} + \tilde{\phi}_{\tilde{y}\tilde{y}} + \tilde{\phi}_{\tilde{z}\tilde{z}} = 0 \quad \text{for} \quad \tilde{z} \in (0, \tilde{h}),$$

$$\tilde{h}_t + \tilde{h}_{\tilde{x}}\tilde{\phi}_{\tilde{x}} + \tilde{h}_{\tilde{y}}\tilde{\phi}_{\tilde{y}} = \tilde{\phi}_{\tilde{z}} \quad \text{at} \quad \tilde{z} = \tilde{h},$$

$$\tilde{\phi}_{\tilde{z}} = 0 \quad \text{at} \quad \tilde{z} = 0.$$

$$\tilde{\varphi}_{\tilde{x}\tilde{x}} + \tilde{\varphi}_{\tilde{y}\tilde{y}} + \tilde{\varphi}_{\tilde{z}\tilde{z}} = 0 \quad \text{for} \quad \tilde{z} \in (\tilde{h}, \tilde{H}),$$

$$\tilde{h}_t + \tilde{h}_{\tilde{x}}\tilde{\varphi}_{\tilde{x}} + \tilde{h}_{\tilde{y}}\tilde{\varphi}_{\tilde{y}} = \tilde{\varphi}_{\tilde{z}} \quad \text{at} \quad \tilde{z} = \tilde{h}\,,$$

$$\tilde{H}_t + \tilde{H}_{\tilde{x}}\tilde{\varphi}_{\tilde{x}} + \tilde{H}_{\tilde{y}}\tilde{\varphi}_{\tilde{y}} = \tilde{\varphi}_{\tilde{z}} \quad \text{at} \quad \tilde{z} = \tilde{H},$$

The first two of these represent the Bernoulli equations at the interface and uppermost surface respectively. The next three represent the Laplace equation and the kinematic equations at the interface and bed for the velocity potential $\tilde{\phi}$. The final three give the Laplace equation and the kinematic conditions at the interface and surface which $\tilde{\varphi}$ must satisfy.

The theory of this paper is most readily applied when the above water wave problem is flattened, which is achieved by taking

$$x=\tilde{x}, \quad y=\tilde{y} \quad t=\tilde{t}, \quad z=\begin{cases}\frac{\tilde{z}}{\tilde{h}}, & \tilde{z}\in[0,\tilde{h}],\\ \frac{(z-1)\tilde{H}-(z-2)\tilde{h}}{\tilde{H}-\tilde{h}}, & \tilde{z}\in[\tilde{h},\tilde{H}]\,,\end{cases}$$

and it also becomes convenient to define

$$\begin{aligned}\phi(x,y,z,t)=\tilde{\phi}(\tilde{x},\tilde{y},\tilde{z},\tilde{t})\,, \quad \varphi(x,y,z,t)=\tilde{\varphi}(\tilde{x},\tilde{y},\tilde{z},\tilde{t})\,,\\ \Phi_1=\phi(x,y,z=1,t)\,, \quad \Psi_1=\varphi(x,y,z=1,t)\,, \quad \Psi_2=\varphi(x,y,z=2,t)\,.\end{aligned}$$

Doing so gives the flattened Lagrangian as

$$\begin{aligned}\mathscr{L}=\iint\Bigg[\rho_1\Big(-h_t\Phi_1+\int_0^1 hu_1(\phi_x-\frac{zh_x}{h}\phi_z)+hv_1(\phi_y-\frac{zh_y}{h}\phi_z)\,dz\Big)\\ +\rho_2\Big(-H_t\Psi_2+h_t\Psi_1+\int_1^2 hu_2(\varphi_x-\frac{(z-1)H_x-(z-2)h_x}{H-h}\varphi_z)\\ +hv_2(\varphi_y-\frac{(z-1)H_y-(z-2)h_y}{h}\varphi_z)\,dz\Big)-S(Z)\Bigg]\,dx\,dy\,dt \quad (35)\end{aligned}$$

where

$$\begin{aligned}u_1&=\phi_x-\frac{zh_x}{h}\phi_z, & v_1&=\phi_y-\frac{zh_y}{h}\phi_z,\\ u_2&=\varphi_x-\frac{(z-1)H_x-(z-2)h_x}{H-h}\varphi_z, & v_2&=\varphi_y-\frac{(z-1)H_y-(z-2)h_y}{H-h}\varphi_z\,,\end{aligned}$$

have been introduced via Legendre transforms, and

$$\begin{aligned}Z&=\big(h,H,\Phi_1,\Psi_1,\phi,\varphi,\Psi_2,u_1,u_2,v_1,v_2\big)^T,\\ S(Z)=&\rho_1R_1h+\rho_2R_2(H-h)-\frac{g}{2}\big((\rho_1-\rho_2)h^2+\rho_2H^2\big)\\ &+\frac{\rho_1}{2}\int_0^1 hu_1^2+hv_1^2-\frac{1}{h}\phi_z^2\,dz\\ &+\frac{\rho_2}{2}\int_1^2(H-h)u_2^2+(H-h)v_2^2-\frac{1}{(H-h)}\varphi_z^2\,dz.\end{aligned}$$

This gives (10) with

$$
\begin{aligned}
\boldsymbol{\alpha}(Z) &= (\rho_1\Phi_1 - \rho_2\Psi_1, \rho_2\Psi_2, -\rho_1 h, \rho_2 h, 0, 0, -\rho_2 H, 0, 0, 0, 0)^T\,,\\
\boldsymbol{\beta}(Z) &= \bigg(\rho_1\int_0^1 z(u_1\phi_z)_x\,dz - \rho_2\int_1^2 (z-2)(u_2\varphi_z)_x\,dz\,,\ \rho_2\int_1^2 (z-1)(u_2\varphi_z)_x\,dz,\\
&\quad -\rho_1 h_x u_1|_{z=1}\,,\ \rho_2 h_x u_2|_{z=1}\,,\ -\rho_1(h(u_1)_x - z h_x(u_1)_z\,,\\
&\quad -\rho_2\big((H-h)(u_2)_x - (z-1)H_x(u_2)_z + (z-2)h_x(u_2)_z\big)\,,\ -\rho_2 H_x u_2|_{z=2}\,,\\
&\quad \rho_1(h\phi_x - z\phi_z h_x)\,,\ \rho_2(H-h)\varphi_x - \rho_2\varphi_z((z-1)H_x - (z-2)h_x)\,,\ 0\,,\ 0\bigg)^T\,,\\
\boldsymbol{\gamma}(Z) &= \bigg(\rho_1\int_0^1 z(v_1\phi_z)_y\,dz - \rho_2\int_1^2 (z-2)(v_2\varphi_z)_x\,dz\,,\ \rho_2\int_1^2 (z-1)(v_2\varphi_z)_y\,dz,\\
&\quad -\rho_1 h_y v_1|_{z=1}\,,\ \rho_2 h_y v_2|_{z=1}\,,\ -\rho_1(h(v_1)_y - z h_y(v_1)_z\,,\\
&\quad -\rho_2\big((H-h)(v_2)_y - (z-1)H_y(v_2)_z + (z-2)h_y(v_2)_z\big)\,,\ -\rho_2 H_y v_2|_{z=2}\,,\\
&\quad 0\,,\ 0\,,\ \rho_1(h\phi_y - z\phi_z h_y)\,,\ \rho_2(H-h)\varphi_y - \rho_2\varphi_z((z-1)H_y - (z-2)h_y)\bigg)^T\,,
\end{aligned}
$$

with their gradients leading to the relevant **M**, **J**, **K** in (11). The above multisymplectic form of equations comes with the additional constraint that $\phi_z|_{z=0} = 0$, since this now does not arise from taking variations with respect to Z, but this information is still present in the flattened Lagrangian. The flattened equations of motion can be found in appendix A. The relevant inner product used for this system is the depth averaging inner product

$$
\begin{aligned}
[\![U,V]\!] = \sum_{i=1}^{4} U_iV_i + U_7V_7 &+ \int_0^1 U_5V_5 + U_8V_8 + U_{10}V_{10}\,dz\\
&+ \int_1^2 U_6V_6 + U_9V_9 + U_{11}V_{11}\,dz\,.
\end{aligned}
$$

The basic state of this system is simply

$$
\widehat{Z}(\boldsymbol{\theta};\mathbf{k},\mathbf{m},\boldsymbol{\omega}) = \big(h_0,\,H_0,\,\theta_1,\,\theta_2,\,\theta_1,\,\theta_2,\,\theta_2,\,k_1,\,k_2,\,m_1,\,m_2\big)^T\,,
$$

so that only the velocity potentials are treated as variables, and the remaining quantities (namely the velocities and thicknesses) are functions of the wave parameters $\mathbf{k}$, $\mathbf{m}$ and $\boldsymbol{\omega}$. The wavenumbers of the relative equilibrium k_i take the role of the flow velocities, and the frequencies ω_i behave like corrections to the total Bernoulli heads of each flow, as suggested by Whitham [32]. The quiescent heights of each interface h_0

and H_0 may be solved for in terms of these parameters and give that

$$h_0(\mathbf{k},\mathbf{m},\boldsymbol{\omega}) = \frac{\rho_1 R_1 - \rho_2 R_2 - \rho_1\omega_1 + \rho_2\omega_2 + \frac{1}{2}\left(\rho_2(k_2^2+m_2^2) - \rho_1(k_1^2+m_1^2)\right)}{g(\rho_1-\rho_2)},$$
$$H_0(\mathbf{k},\mathbf{m},\boldsymbol{\omega}) = \frac{1}{g}\left(R_2 - \omega_2 - \frac{k_2^2+m_2^2}{2}\right).$$

The aim now is to determine when the criticality condition (18) occurs for this uniform flow state, which leads to the KP equation (33) emerging from the stratified water wave problem.

4.1. Conservation laws and criticality

The conservation laws for this problem are the conservation of mass within each layer of the fluid, and are associated with the affine symmetry of each velocity potential. Evaluated on the uniform flow solution, the components of these conservation laws are found to be

$$\mathbf{A} = \begin{pmatrix} \rho_1 h_0 \\ \rho_2(H_0-h_0)\end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} \rho_1 k_1 h_0 \\ \rho_2 k_2(H_0-h_0)\end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \rho_1 m_1 h_0 \\ \rho_2 m_2(H_0-h_0)\end{pmatrix}. \tag{36}$$

We use these to determine the relevant criticality conditions the uniform flow solution must satisfy in order for the KP equation to emerge. The first of these is (18), and may then be computed from

$$\begin{aligned} 0 &= \det\left[\mathrm{D}_{\mathbf{k}}\mathbf{B}\right] \\ &= \det\left[\frac{1}{g(1-r)}\begin{pmatrix} g\rho_1 h_0(1-r-F_1^2) & \rho_2 k_1 k_2 \\ \rho_2 k_1 k_2 & g\rho_2(H_0-h_0)(1-r-F_2^2)\end{pmatrix}\right]. \end{aligned}$$

In the above, we have introduced the density ratio r and the Froude numbers within each stratification layer F_i as

$$r = \frac{\rho_2}{\rho_1} < 1, \quad F_1^2 = \frac{k_1^2}{gh_0}, \quad F_2^2 = \frac{k_2^2}{g(H_0-h_0)}.$$

This determinant vanishes precisely when

$$(1-F_1^2)(1-F_2^2) = r. \tag{37}$$

Such a condition appears within the existing literature as a stability boundary, across which there is a tendency for hydraulic jumps to occur (as discussed in [19, 39]). Additionally, this condition arises within the modulation of uniform flows in stratified shallow water systems as the one necessary for the KdV equation to emerge [23, 40]. In such cases where the criticality condition (37) holds, one may then define the eigenvector necessary for the theory as

$$\boldsymbol{\zeta} = \begin{pmatrix} -\rho_2 k_1 k_2 \\ g\rho_1 h_0(1-r-F_1^2)\end{pmatrix}$$

This will be used, along with the relevant tensors, to generate the coefficients of the KP equation according to (33).

The other conservation law criticality we require is

$$\boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{C}+\mathrm{D}_{\mathbf{m}}\mathbf{B})\boldsymbol{\zeta} = -2g\rho_1^2\rho_2 h_0(1-r-F_1^2)\big(k_1 m_1(1-F_2^2)+k_2 m_2(1-F_1^2)\big) = 0\,.$$

The simplest way this is met is by choosing $\mathbf{m}=\mathbf{0}$, which we will use for this example, although it is possible for other choices for $\mathbf{m}$ exist which will satisfy the above. By this choice, $\boldsymbol{\eta}=\mathbf{0}$ and the calculation of the Y derivative term in (33) is much simpler. This completes the assessment of criticality of the uniform flow and the coefficients of the relevant KP equation can now be computed assuming these are met.

We are able to determine the relevant matrices required to compute the desired coefficients,

$$\begin{aligned}
\mathrm{D}_{\mathbf{k}}\mathbf{A} =& \frac{1}{g(1-r)}\begin{pmatrix} -\rho_1 k_1 & \rho_2 k_2 \\ \rho_2 k_1 & -\rho_2 k_2 \end{pmatrix},\\
\mathrm{D}_{\mathbf{m}}\mathbf{C} =& \begin{pmatrix} \rho_1 h_0 & 0 \\ 0 & \rho_2(H_0-h_0) \end{pmatrix},\\
\mathrm{D}^2_{\mathbf{k}}\mathbf{B} =& \frac{1}{2g(1-r)}\left(\begin{pmatrix} -6\rho_1 k_1 & 2\rho_2 k_2 \\ 2\rho_2 k_2 & 2\rho_2 k_1 \end{pmatrix}\middle|\begin{pmatrix} 2\rho_1 k_2 & 2\rho_2 k_1 \\ 2\rho_2 k_1 & -6\rho_2 k_2 \end{pmatrix}\right).
\end{aligned}$$

These give the following projections:

$$\boldsymbol{\zeta}^T\mathrm{D}_{\mathbf{m}}\mathbf{C}\boldsymbol{\zeta} = \rho_1^2\rho_2 g^2 h_0^2(H_0-h_0)(1-r-F_1^2)\big(F_1^2+F_2^2-2F_1^2F_2^2\big)$$

$$\begin{aligned}
\boldsymbol{\zeta}^T(\mathrm{D}_{\mathbf{k}}\mathbf{A}+\mathrm{D}_{\boldsymbol{\omega}}\mathbf{B})\boldsymbol{\zeta} =& -2g^2\rho_1^2\rho_2(H_0-h_0)h_0^2(1-r-F_1^2)\\
&\times\left(\frac{k_1}{gh_0}(1-F_2^2)+\frac{k_2}{g(H_0-h_0)}(1-F_1^2)\right),
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\zeta}^T\mathrm{D}^2_{\mathbf{k}}\mathbf{B}\big(\boldsymbol{\zeta},\boldsymbol{\zeta}\big) =& 3g^2\rho_1^3\rho_2 k_2 h_0^2(1-r-F_1^2)\\
&\times\big((H_0-h_0)r(1-F_2^2)F_1^2-h_0(1-F_1^2)^2F_2^2)\big)
\end{aligned}$$

Additionally, we require the coefficient of the third order dispersive term, which requires the Jordan chain analysis. The details of this calculation may be found in appendix B, which result in the coefficient

$$\begin{aligned}
\boldsymbol{\zeta}^T\mathscr{K} =& g^2\rho_1^2\rho_2 h_0^2(H_0-h_0)(1-r-F_1^2)\\
&\times\left[\frac{1}{3}h_0^2F_1^2(1-F_2^2)+(H_0-h_0)^2F_2^2(1-F_1^2)\left(\frac{1}{3}-2F_2^2+F_2^4\right)\right].
\end{aligned}\tag{38}$$

Therefore, the KP equation we obtain is

$$\Big(a_0U_T+a_1UU_X+a_2U_{XXX}\Big)_X+a_3U_{YY}=0,$$

with

$$
\begin{aligned}
a_0 &= -2g^2\rho_1^2\rho_2(H_0-h_0)h_0^2(1-r-F_1^2)\left(\tfrac{k_1}{gh_0}(1-F_2^2)+\tfrac{k_2}{g(H_0-h_0)}(1-F_1^2)\right),\\
a_1 &= 3g^2\rho_1^3\rho_2k_2h_0^2(1-r-F_1^2)\big((H_0-h_0)r(1-F_2^2)F_1^2-h_0(1-F_1^2)^2F_2^2)\big),\\
a_2 &= g^2\rho_1^2\rho_2h_0^2(H_0-h_0)(1-r-F_1^2)\\
&\quad\times\left[\tfrac{1}{3}h_0^2F_1^2(1-F_2^2)+(H_0-h_0)^2F_2^2(1-F_1^2)\left(\tfrac{1}{3}-2F_2^2+F_2^4\right)\right],\\
a_3 &= \rho_1^2\rho_2g^2h_0^2(H_0-h_0)(1-r-F_1^2)\big(F_1^2+F_2^2-2F_1^2F_2^2\big),
\end{aligned}
$$

which can be simplified to

$$
\begin{aligned}
\Bigg(&\left(\frac{k_1}{gh_0}(1-F_2^2)+\frac{k_2}{g(H_0-h_0)}(1-F_1^2)\right)U_T\\
&-\frac{3g^2\rho_1k_2}{2(H_0-h_0)}\big((H_0-h_0)r(1-F_2^2)F_1^2-h_0(1-F_1^2)^2F_2^2)\big)UU_X\\
&-\frac{1}{2}\left[\frac{1}{3}h_0^2F_1^2(1-F_2^2)+(H_0-h_0)^2F_2^2(1-F_1^2)\left(\frac{1}{3}-2F_2^2+F_2^4\right)\right]U_{XXX}\Bigg)_X\\
&-\frac{1}{2}\big(F_1^2+F_2^2-2F_1^2F_2^2\big)U_{YY}=0\,.
\end{aligned}
\tag{39}
$$

5. Concluding Remarks

This paper has illustrated that, given a two-phased wavetrain or relative equilibrium solution to a general Lagrangian system, one may reduce the original Euler-Lagrange equations to the KP equation with coefficients tied to the conservation of wave action. This result has been demonstrated by reducing the full water wave problem with two layers of stratification to the KP equation using a uniform flow solution. The theory, although only appearing within the paper for two phases, may be extended to arbitrarily many with the appropriate modifications to the sums appearing with the reduction.

Although only illustrated for the full stratified water wave problem here, the theory of this paper has several other applications. Most readily, it can be applied to the coupled nonlinear Schrödinger models in 2+1, which arise across various contexts. Most interestingly though, the theory of this paper can be applied to the single layered water wave problem in the case where a mean flow couples to some surface wave profile. A weakly nonlinear version of this scenario has been covered in Donaldson and Bridges from a different perspective [33], however the modulation approach has the ability to not only relate the criticality to the degeneracy of the associated conservation laws, but to also incorporate time evolution into the analysis.

The KP equation derived for the stratified water wave problem (39) is robust for a large choice of velocities and layer thicknesses, so long as these lie on the surface (37). However, there exist cases where one of the coefficients within this equation vanishes and the KP equation is no longer a valid asymptotic reduction. Various cases of this have been considered in one spatial dimension for the shallow water variant of this problem [24, 25, 41], demonstrating how these degeneracies lead to the two-way Boussinesq equation, modified KdV equation and various other nonlinear PDEs emerging from the modulation. It is expected that the two spatial variants of these will be recovered from similar analyses. Additionally there is the potential for the coefficient of the transverse term, U_{YY}, to vanish as well. A rescaling in this case is expected to lead to a fully two-dimensional KdV equation to be the result of the modulational approach. Additionally, a set of coupled KP equations is expected when the geometric multiplicity of the zero eigenvalue of $\mathrm{D}_{\mathbf{k}}\mathbf{B}$ is greater than one, a scenario more likely in the case of three or more phases. This is since additional eigenvectors associated with the zero eigenvector enter the analysis, which facilitate additional projections of the vector KP which arises at fifth order and thus a set of coupled nonlinear PDEs.

The modulation approach presented here may also be altered to allow the derivation of the KP equation for a moving frame, providing a further connection between the existing literature and the phase dynamical approach outlined here. By doing so, nonlinear dynamics instead emerge by suitable choice of the frame's speed c to meet the solvability criteria [42]. This would in fact lead to a KP which emerges with only the condition that such a c exists and is real, rather than imposing a condition on the wavenumbers, making it arise more readily across various contexts.

Acknowledgements

The author would like to thank Tom Bridges for the motivation and helpful discussions that eventually led to this manuscript, as well as the referees for their useful comments which shaped this manuscript. The majority of this work was undertaken whilst in receipt of a fully funded EPSRC PhD studentship from grant no. EP/L505092/1.

References

1. B. B. Kadomtsev & V. I. Petviashvili, *On the stability of solitary waves in weakly dispersing media* Sov. Phys. Dokl. **15.6**, 539–541 (1970).

2. M.J. Ablowitz & H. Segur, *On the evolution of packets of water waves*, J. Fluid Mech. **92.4**, 691-715 (1979).
3. P.A. Milewski, & J.B. Keller, *Three–Dimensional Water Waves*, Stud. Appl. Math. **97.2**, 149-166 (1996).
4. F. Baronio et al., *Optical-fluid dark line and X solitary waves in Kerr media*, Optical Data Processing and Storage **3.1**, 1-7 (2017).
5. D.E. Pelinovsky, Y.A.. Stepanyants & Y.S. Kivshar, *Self-focusing of plane dark solitons in nonlinear defocusing media*, Phys. Rev. E **51.5**, 5016 (1995).
6. E. A. Kuznetsov and S. K. Turitsyn, *Instability and collapse of solitons in media with a defocusing nonlinearity* Sov. Phys. JETP **67**, 1583 (1988)
7. S. Tsuchiya, F. Dalfovo, & L. Pitaevskii, *Solitons in two-dimensional Bose-Einstein condensates*, Phys. Rev. A **77.4**, 045601 (2008).
8. V.D. Djordjevic & L. G. Redekopp, *The fission and disintegration of internal solitary waves moving over two-dimensional topography*, J. Phys. Oceanogr., **8.6**, 1016–1024 (1978).
9. C.G. Koop and, and G. Butler, *An investigation of internal solitary waves in a two–fluid system*, J. Fluid Mech., **112**, 225–251 (1981).
10. G.B. Whitham, *Linear and Nonlinear Waves*, Vol. 42. John Wiley & Sons, (2011).
11. W.D. Hayes, *Conservation of action and modal wave action* Proc. Roy. Soc. A **320** 187–208 (1970).
12. G.A. El & M. A. Hoefer, *Dispersive shock waves and modulation theory*, Physica D **333**, 11-65 (2016).
13. M.C. Cross & A.C. Newell, *Convection patterns in large aspect ratio systems*, Physica D **10.3**, 299-328 (1984).
14. Y. Kuramoto & T. Tsuzuki, *Persistent propagation of concentration waves in dissipative media far from thermal equilibrium*, Prog. Theo. Phys. **55.2**, 356-369 (1976).
15. Y. Kuramoto, *Phase dynamics of weakly unstable periodic structures*, Prog. Theo. Phys. **71.6**, 1182-1196 (1984).
16. A. Doelman, B. Sandstede, A. Scheel, & G. Schneider, *The dynamics of modulated wave trains*, American Mathematical Society (2005).
17. T.J. Bridges, *A universal form for the emergence of the Kortewegde Vries equation* Proc. Roy. Soc. A. **469**, 20120707(2013).
18. T.J. Bridges, *Symmetry, Phase Modulation and Nonlinear Waves* Cambridge University Press (2017).
19. G.S. Benton, *The occurence of critical flow and hydraulic jumps in a multi-layered fluid system*, J. Meteor. **11**, 139–150 (1954).
20. T.J. Bridges, & F.E. Laine–Pearson, *Multisymplectic relative equilibria, multiphase wavetrains, and coupled NLS equations*, Stud. Appl. Math. **107.2**, 137-155 (2001).
21. F.E. Laine–Pearson & T.J. Bridges, *Nonlinear counterpropagating waves, multisymplectic geometry, and the instability of standing waves*, SIAM J. Appl. Math. **64.6**, 2096-2120 (2004).
22. D.J. Ratliff & T.J. Bridges, *Whitham modulation equations, coalescing characteristics, and dispersive Boussinesq dynamics* Physica D **333**, 107–116 (2016).
23. D.J. Ratliff & T.J. Bridges, *Multiphase Wavetrains, Singular Wave Interactions and the Emergence of the Korteweg-de Vries Equation* Proc. Roy. Soc. A **421**, 20160456 (2016).
24. D.J. Ratliff, *Double Degeneracy in Multiphase Modulation Leading to the Boussinesq Equation*, Stud. Appl. Math **140.1**, 48–77 (2018).
25. D.J. Ratliff, *Vanishing characteristic speeds and critical dispersive points in nonlinear interfacial wave problems* Phys. Fluids **29.11**, 112104 (2017).

26. J. Hammack, N. Scheffner, & H. Segur, *Two-dimensional periodic waves in shallow water* J. Fluid Mech. **209** 567-589 (1989).
27. R.Johnson, *Water waves and Korteweg–de Vries equations*, J. Fluid Mech. **97.4**, 701-719 (1980).
28. D.J. Ratliff & T.J. Bridges, *Phase dynamics of periodic waves leading to the Kadomtsev–Petviashvili equation in 3+ 1 dimensions* Proc. R. Soc. A. **471**, 20150137 (2015).
29. K.B. Dysthe, *Note on a modification to the nonlinear Schrdinger equation for application to deep water waves*, Proc. R. Soc. Lond. A **369.1736**, 105-114 (1979).
30. K. Trulsen, Karsten, & K. B. Dysthe, *A modified nonlinear Schrdinger equation for broader bandwidth gravity waves on deep water*, Wave motion **24.3**, 281-289 (1996).
31. D. J. Benney and G. J. Roskes, *Wave instability*, Stud. Appl. Math. **48**,455–472 (1969).
32. G.B. Whitham, *Non-linear dispersion of water waves*, J. Fluid Mech. **27.2**, 399-412 (1967).
33. T.J. Bridges & N. M. Donaldson, *Secondary criticality of water waves. Part 1. Definition, bifurcation and solitary waves*, J. Fluid Mech. **565**, 381-417 (2006).
34. T.J. Bridges, P.E. Hydon, & J.K. Lawson, *Multisymplectic structures and the variational bicomplex*, Math. Proc. Cambridge **148.1**, 159–178 (2010).
35. P.E. Hydon, *Multisymplectic conservation laws for differential and differential-difference equations*, Proc. R. Soc. A **461**, 1627–1637 (2005).
36. R. Abraham and J. E. Marsden, *Foundations of mechanics*, Benjamin/Cummings Publishing Company (1978).
37. D. Ambrosi, *Hamiltonian formulation for surface waves in a layered fluid* Wave motion **31.1**, 71–76 (2000).
38. J.C. Luke, *A variational principle for a fluid with a free surface*, J. Fluid Mech. **27.2**, 395–397 (1967).
39. G.A. Lawrence, *On the hydraulics of Boussinesq and non-Boussinesq two-layer flows*, J. Fluid Mech. **215**, 457–480 (1990).
40. T.J. Bridges & D.J. Ratliff, *Double criticality and the two-way Boussinesq equation in stratified shallow water hydrodynamics*, Phys. Fluids **28** 062103 (2016).
41. D.J.Ratliff, *Conservation Laws, Modulation and the Emergence of Universal Forms*, Ph.D Thesis, University of Surrey (2017).
42. T.J. Bridges & D.J. Ratliff, *Nonlinear modulation near the Lighthill instability threshold in 2+ 1 Whitham theory*, Phil. Trans. A (2018) (in press).

Appendix A. Details of the Flattening Transformation

In order to flatten the two-layered Euler problem, one must make the following change of co-ordinates:

$$x = \tilde{x}, \quad y = \tilde{y} \quad t = \tilde{t}, \quad z = \begin{cases} \frac{\tilde{z}}{\tilde{h}}, & \tilde{z} \in [0, \tilde{h}], \\ \frac{(z-1)\tilde{H}-(z-2)\tilde{h}}{\tilde{H}-\tilde{h}}, & \tilde{z} \in [\tilde{h}, \tilde{H}] \end{cases}$$

This gives two sets of derivative transforms. The first set (for $z \in [0, 1]$) are given by

$$\begin{aligned} \frac{\partial}{\partial \tilde{t}} &= \frac{\partial}{\partial t} - \frac{zh_t}{h}\frac{\partial}{\partial z}, & \frac{\partial}{\partial \tilde{x}} &= \frac{\partial}{\partial x} - \frac{zh_x}{h}\frac{\partial}{\partial z}, \\ \frac{\partial}{\partial \tilde{y}} &= \frac{\partial}{\partial y} - \frac{zh_y}{h}\frac{\partial}{\partial z}, & \frac{\partial}{\partial \tilde{z}} &= \frac{1}{h}\frac{\partial}{\partial z}, \end{aligned}$$

The upper layer admits the following transforms:

$$\begin{aligned} \frac{\partial}{\partial \tilde{t}} &= \frac{\partial}{\partial t} - \frac{(z-1)H_t-(z-2)h_t}{H-h}\frac{\partial}{\partial z}, & \frac{\partial}{\partial \tilde{x}} &= \frac{\partial}{\partial x} - \frac{(z-1)H_x-(z-2)h_x}{H-h}\frac{\partial}{\partial z} \\ \frac{\partial}{\partial \tilde{y}} &= \frac{\partial}{\partial y} - \frac{(z-1)H_y-(z-2)h_y}{H-h}\frac{\partial}{\partial z}, & \frac{\partial}{\partial \tilde{z}} &= \frac{1}{H-h}\frac{\partial}{\partial z}, \end{aligned}$$

We also pre-emptively define the fluid velocities in each layer:

$$\begin{aligned} u_1 &= \phi_x - \frac{zh_x}{h}\phi_z, & v_1 &= \phi_y - \frac{zh_y}{h}\phi_z, \\ u_2 &= \varphi_x - \frac{(z-1)H_x-(z-2)h_x}{H-h}\varphi_z, & v_2 &= \varphi_y - \frac{(z-1)H_y-(z-2)h_y}{H-h}\varphi_z. \end{aligned}$$

After considering the relevant Lagrange multipliers for the above substitutions, this then transforms the original Lagrangian into the one appearing in (35). It then follows from taking variations that the flattened

Euler-Lagrange equations for this problem are

$$\rho_2\bigg(H_t + H_x\ u_2|_{z=2} + H_y\ v_2|_{z=2} \bigg) = \frac{\rho_2}{H-h}\ \varphi_z|_{z=2}\ ,$$

$$\rho_2\bigg(h_t + h_x\ u_2|_{z=1} + h_y\ v_2|_{z=1} \bigg) = \frac{\rho_2}{H-h}\ \varphi_z|_{z=1}\ ,$$

$$\rho_2\bigg((H-h)\partial_x u_2 - \big((z-1)H_x - (z-2)h_x\big)\partial_z u_2$$
$$+ (H-h)\partial_y v_2 - \big((z-1)H_y - (z-2)h_y\big)\partial_z v_2 + \frac{1}{H-h}\varphi_{zz}\bigg) = 0, \quad z\in(1,2)\,,$$

$$\rho_1\bigg(h_t + h_x\ u_1|_{z=1} + h_y\ v_1|_{z=1} \bigg) = \frac{\rho_1}{h}\ \phi_z|_{z=1}\ ,$$

$$\frac{\rho_1}{h}\ \phi_z|_{z=0} = 0\,,$$

$$\rho_1\bigg(h\partial_x u_1 - zh_x\partial_z u_1 + h\partial_y v_1 - zh_y\partial_z v_1 + \frac{1}{h}\phi_{zz}\bigg) = 0, \quad z\in(0,1)\,,$$

$$\rho_1\bigg(\partial_t\Phi_1 + u_1|_{z=1}\ \phi_x|_{z=1} + v_1|_{z=1}\ \phi_y|_{z=1} + \int_0^1 z(\partial_x u_1\phi_z - \partial_z u_1\phi_x)\ dz$$
$$+ \int_0^1 z(\partial_y v_1\phi_z - \partial_z v_1\phi_y)\ dz - \frac{1}{2}\int_0^1 u_1^2 + v_1^2 + \frac{1}{h^2}\phi_z^2\ dz + gh - R_1\bigg)$$
$$= \rho_2\bigg(\partial_t\Psi_1 + u_2|_{z=1}\ \varphi_x|_{z=1} + v_2|_{z=1}\ \varphi_y|_{z=1} + gh - R_2$$
$$+ \int_1^2 (2-z)(\partial_x u_2\varphi_z - \partial_z u_2\varphi_x)\ dz + \int_1^2 (2-z)(\partial_y v_2\varphi_z - \partial_z v_2\varphi_y)\ dz$$
$$- \frac{1}{2}\int_1^2 u_2^2 + v_2^2 + \frac{1}{(H-h)^2}\varphi_z^2\ dz\bigg)\,,$$

$$\rho_2\bigg(\partial_t\Psi_2 + u_2|_{z=2}\,\partial_x\Psi_2 + v_2|_{z=2}\,\partial_y\Psi_2 - R_2 + gH$$
$$+ \int_1^2 (z-1)(\partial_x u_2\varphi_z - \partial_z u_2\varphi_x)\ dz + \int_1^2 (z-1)(\partial_y v_2\varphi_z - \partial_z v_2\varphi_y)\ dz$$
$$- \frac{1}{2}\int_1^2 u_2^2 + v_2^2 + \frac{1}{(H-h)^2}\varphi_z^2\ dz\bigg) = 0\,.$$

The first three correspond to the kinematic conditions at the top surface, at the interface between the fluids and the Laplace equation for the top fluid respectively. The following three are the kinematic conditions at the interface, the bed and Laplace's equation in the bottom fluid respectively. The variation in h gives the flattened version of the Bernoulli equation at the interface, and the H variations result in the flattened Bernoulli equation for the top free surface.

Appendix B. Details of the Jordan Chain Calculation

Here we provide the details for the Jordan chain theory that results in the dispersive term (38). The linear operator about the uniform flow state may be found as

$$\mathbf{L} = \begin{pmatrix} \rho_2(\omega_1\partial_{\theta_1}+\omega_2\partial_{\theta_2})\Psi_1 - \rho_1(\omega_1\partial_{\theta_1}+\omega_2\partial_{\theta_2})\Phi_1 + (\rho_2-\rho_1)gh \\ -\rho_2(\omega_1\partial_{\theta_1}+\omega_2\partial_{\theta_2})\Psi_2 - \rho_2 gH \\ \rho_1(\omega_1\partial_{\theta_1}+\omega_2\partial_{\theta_2})h - \frac{\rho_1}{h_0}\phi_z\Big|^{z=1} \\ \rho_2(\omega_1\partial_{\theta_1}+\omega_2\partial_{\theta_2})h + \frac{\rho_2}{H_0-h_0}\psi_z\Big|^{z=1} \\ \frac{\rho_1}{h_0}\phi_{zz} \\ \frac{\rho_2}{H_0-h_0}\psi_{zz} \\ \rho_2(\omega_1\partial_{\theta_1}+\omega_2\partial_{\theta_2})H - \frac{\rho_2}{H_0-h_0}\psi_z\Big|^{z=2} \\ \rho_1 h_0 u_1 \\ \rho_2(H_0-h_0)u_2 \\ \rho_1 h_0 v_1 \\ \rho_2(H_0-h_0)v_2 \end{pmatrix}$$

We can also determine the skew symmetric operator in the Jordan Chain relation as

$$\mathbf{J}(\widehat{Z}) =$$
$$\begin{pmatrix} 0 & 0 & \rho_1 k_1 & -\rho_2 k_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \rho_2 k_2 & 0 & 0 & 0 & 0 \\ -\rho_1 k_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho_2 k_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho_1 h_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho_2(H_0-h_0) & 0 & 0 \\ 0 & -\rho_2 k_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho_1 h_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_2(H_0-h_0) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Thus, the first four elements of the chain according to the theory in §2.3 are

$$\xi_1^1 = \widehat{Z}_{\theta_1} = \big(0,0,1,0,1,0,0,0,0,0,0\big)^T,$$

$$\xi_1^2 = \widehat{Z}_{\theta_2} = \big(0,0,0,1,0,1,1,0,0,0,0,\big)^T,$$

$$\xi_2^1 = \widehat{Z}_{k_1} = \big(\partial_{k_1}h_0,0,0,0,0,0,0,1,0,0,0\big)^T,$$

$$\xi_2^2 = \widehat{Z}_{k_2} = \big(\partial_{k_2}h_0,-\frac{k_2}{g},0,0,0,0,0,0,1,0,0\big)^T$$

The next element in the chain satisfies

$$\mathbf{L}\xi_3 = \mathbf{J}(\widehat{Z})(\zeta_i\xi_2^1 + \zeta_2\xi_2^2) = \begin{pmatrix} 0 \\ 0 \\ -\zeta_1\rho_1 k_1\partial_{k_1}h_0 - \zeta_2\rho_1 k_1\partial_{k_2}h_0 \\ \zeta_1\rho_2 k_2\partial_{k_1}k_0 + \zeta_2\rho_2 k_2\partial_{k_2}h_0 \\ -\zeta_1\rho_1 h_0 \\ -\zeta_2\rho_2(H_0 - h_0) \\ \zeta_2\rho_2\frac{k_2^2}{g} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

This system is solvable providing $\det\big[\mathrm{D}_{\mathbf{k}}\mathbf{B}\big] = 0$, and gives that

$$\xi_3 = \begin{pmatrix} 0 \\ 0 \\ -\zeta_1\frac{1}{2}h_0^2 \\ \zeta_2(H_0 - h_0)^2\big(\frac{3}{2} - F_2^2\big) \\ -\zeta_1\frac{1}{2}h_0^2 z^2 \\ \zeta_2(H_0 - h_0)^2 z\big(2 - F_2^2 - \frac{1}{2}z\big) \\ 2\zeta_2(H_0 - h_0)^2\big(1 - F_2^2\big) \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Finally we then solve

$$\mathbf{L}\xi_4 = \begin{pmatrix} -\frac{1}{2}\zeta_1\rho_1 k_1 h_0^2 - \zeta_2\rho_2 k_2(H_0 - h_0)\big(\frac{3}{2}(H_0 - h_0) - \frac{k_2^2}{g}\big) \\ 2\zeta_2\rho_2 k_2\zeta_2(H_0 - h_0)\big(1 - F_2^2\big) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -\frac{1}{2}\rho_1 h_0^3 z^2 \\ \zeta_2\rho_2(H_0 - h_0)^3 z\big(2 - F_2^2 - \frac{1}{2}z\big) \\ 0 \\ 0 \end{pmatrix},$$

giving that

$$\xi_4 = \begin{pmatrix} \frac{1}{g(\rho_1-\rho_2)}\left(\frac{1}{2}\zeta_1\rho_1 k_1 h_0^2 + \zeta_2\rho_2 k_2 (H_0-h_0)^2\left(\frac{3}{2}-F_2^2\right)\right) \\ -2\zeta_2 k_2 g^{-1}\zeta_2 (H_0-h_0)^2\left(1-F_2^2\right) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -\zeta_1\frac{1}{2}h_0^2 z^2 \\ \zeta_2 z(H_0-h_0)\left(2-F_2^2-\frac{1}{2}z\right) \\ 0 \\ 0 \end{pmatrix}$$

From this, we can determine the elements of the vector $\mathscr{K}$:

$$\begin{aligned}
\mathscr{K}_1 =& -\langle \widehat{Z}_{\theta_1}, \mathbf{J}\xi_4\rangle \\
=& \frac{1}{(1-r)}\left[\zeta_1\rho_1 h_0^3\left(\frac{1}{2}F_1^2 - \frac{1}{6}(1-r)\right) + \frac{\zeta_2\rho_2 k_1 k_2}{g}(H_0-h_0)^2\left(\frac{3}{2}-F_2^2\right)\right] \\
\mathscr{K}_2 =& -\langle \widehat{Z}_{\theta_2}, \mathbf{J}\xi_4\rangle \\
=& \frac{\rho_1\rho_2}{(\rho_1-\rho_2)}\Bigg[-\frac{1}{2g}\zeta_1 k_1 k_2 h_0^2 + \zeta_2 (H_0-h_0)^3\left((1-r)\left(\frac{11}{6}-\frac{3}{2}F_2^2\right)\right. \\
& \left. -2(1-r)F_2^2\left(1-F_2^2\right) - rF_2^2\left(\frac{3}{2}-F_2^2\right)\right)\Bigg].
\end{aligned}$$

The projection of $\mathscr{K}$ then gives, by utilising the criticality condition (37) several times to simplify,

$$\begin{aligned}
\boldsymbol{\zeta}^T\mathscr{K} &= \zeta_1\mathscr{K}_1 + \zeta_2\mathscr{K}_2 \\
&= g^2\rho_1^2\rho_2 h_0^2 (H_0-h_0)(1-r-F_1^2) \\
&\quad \times \left[\frac{1}{3}h_0^2 F_1^2 (1-F_2^2) + (H_0-h_0)^2 F_2^2 (1-F_1^2)\left(\frac{1}{3} - 2F_2^2 + F_2^4\right)\right].
\end{aligned}$$