Rigorous justification of the Whitham modulation theory for equations of NLS type

W.A. Clarke
School of Mathematics and Statistics
University of Sydney
wcla7359@uni.sydney.edu.au

R. Marangell
School of Mathematics and Statistics
University of Sydney
robert.marangell@sydney.edu.au

November 19, 2020

Abstract

We study the modulational stability of periodic travelling wave solutions to equations of nonlinear Schrödinger type. In particular, we prove that the characteristics of the quasi-linear system of equations resulting from a slow modulation approximation satisfy the same equation, up to a change in variables, as the normal form of the linearized spectrum crossing the origin. This normal form is taken from [LBJM2019], where Leisman et al. compute the spectrum of the linearized operator near the origin via an analysis of Jordan chains. We derive the modulation equations using Whitham's formal modulation theory, in particular the variational principle applied to an averaged Lagrangian. We use the genericity conditions assumed in the rigorous theory of [LBJM2019] to direct the homogenization of the modulation equations. As a result of the agreement between the equation for the characteristics and the normal form from the linear theory, we show that the hyperbolicity of the Whitham system is a necessary condition for modulational stability of the underlying wave.

1 Introduction

In this paper, we consider the modulational stability of periodic solutions to equations of nonlinear Schrödinger type

$$i\psi_t = \psi_{xx} + \zeta f(|\psi|^2)\psi \tag{1}$$

under perturbations in $L_2(\mathbb{R})$. The nonlinearity $f(|\psi|^2)$ is arbitrary, but assumed to be well-behaved (double-integrable). This equation has a rich history of study; the cubic nonlinear Schrödinger equation, $f(|\psi|^2) = \pm|\psi|^2$, describes the envelope of a slowly modulated carrier wave in a dispersive medium [SS1999, AS1981]. This equation, and others of NLS-type with higher order nonlinearities, arise in the study of a plethora of physical systems, including: water waves [Zakharov1968, HO1972]; nonlinear optics [Agrawal2013, HM2003]; plasma physics [Chen2016, LTE2019, MOMT1976]; and Bose-Einstein condensates [Gross1961, Pitaevskii1961].

Since the dynamics of equation (1) exhibit linear, nonlinear and modulatory behaviour, the literature includes analyses of linearized, orbital and modulational stability. If one chooses a suitable potential $f(|\psi|^2)$ such that equation (1) is integrable, then it is possible to give an explicit description of the spectrum. The cubic nonlinear Schrödinger equation is the best example of this [BDN2011, DS2017], however relying on integrability is not necessary [GLCT2017]. Rowlands determined the spectral stability of stationary periodic solutions of the cubic nonlinear Schrödinger equation subject to long-wavelength disturbances [Rowlands1974], and in so doing demonstrated modulational instability in the focusing case. Alfimov, Its and Kulagin [AIK1990] constructed the homoclinic orbit for an unstable, spatially periodic solution to the focusing nonlinear Schrödinger equation, essentially providing a nonlinear description of the modulational stability of this type of solution. Using the general methods of [GSS1987, GSS1990], Gallay and Hărăguş proved that quasiperiodic, small-amplitude solutions to the cubic nonlinear Schrödinger equation are orbitally stable within the class of solutions having the same period and Floquet multipler [GH2007A]. They further proved that these solutions are linearly stable under bounded perturbations in the defocusing case, but linearly unstable in the focusing case. In [GH2007B], Gallay and Hărăguş extend the orbital stability results in [GH2007A] to solutions of any amplitude.

This paper deals with modulational stability of equation (1), that is, the spectral stability subject to long-wavelength perturbations. Rigorously speaking, this amounts to expanding the spectrum of the linearized operator in a neighbourhood of the origin in the spectral plane. Whitham modulation theory [Whi1965A, Whi1965B, Whi1967, Whi1970, Whi1999] provides a formal procedure in which the modulational stability is computed by considering the hyperbolicity of a system of PDEs called the Whitham modulation equations. These modulation equations arise from an asymptotic expansion of the governing PDE via $(x,t) \mapsto (\epsilon x, \epsilon t)$ along with a WKB approximation of the solution. To $O(\epsilon)$, these equations are an homogeneous system of PDEs in terms of the slowly-varying parameters of the original PDE. Proving that the Whitham modulation theory accurately predicts the results of the rigorous analysis of the linearized spectrum is non-trivial, and to our knowledge is an open problem in the general case. Our present analysis is influenced by the examples of where this has been done, such as: the nonlinear Klein-Gordon equation [JMMP2014]; the generalized Korteweg-de Vries equation [JZ2010]; and systems of viscous conservation laws [Serre2005, OZ2006].

In addition to the WKB approximation and asymptotic expansion, Whitham proves that there are several other, equivalent methods for deriving the Whitham modulation equations [Whi1970, Whi1999]. In particular, these are: the variational principle applied to an averaged Lagrangian; and averaging and two-timing a system of conservation laws. As a result of unwieldy algebra that arises in the modulation equations, one often chooses the method which provides the simplest derivation of the Whitham modulation equations, thereafter making a change of variables in order to prove that the hyperbolicity of the Whitham system is equivalent to the linearized stability theory. When the modulational instability criterion from the linearized stability theory is given in terms of the derivatives of special quantities of the original PDE such as the period, mass and momentum, a useful technique is to introduce a classical action variable W, since the derivatives of W are related to these special quantities. This is the case for both the nonlinear Klein-Gordon equation [JMMP2014] and the generalized Korteweg-de Vries equation [BJ2010, JZ2010]. This approach relies on the symmetry properties of Lagrangian systems, and Whitham in fact derives the modulation equations using the averaged Lagrangian method in the cases of both of these PDEs (only the non-generalized case for KdV) [Whi1999]. Equation (1) admits a Lagrangian (cf. equation (22)), and so we follow the averaged Lagrangian method in order to more easily derive the Whitham modulation equations. If, however, there is not an obvious Lagrangian but the travelling wave ODE is integrable, one can compute the tangent space to the manifold of travelling wave solutions and construct the kernel of the linearized operator from a basis of the tangent space. In [Serre2005], Serre performs this exact procedure to prove that the hyperbolicity criterion of the Whitham modulation equations agrees with the rigorous modulational stability analysis (as determined from an Evans function expansion) of a system of scalar, viscous conservation laws [OZ2003], which is later extended to the multi-dimensional case [OZ2006].

There is a well-developed Whitham modulation theory for the nonlinear Schrödinger equation; Düll and Schneider have proven that the Whitham modulation equations are a valid approximation of spatial and temporal modulations of periodic wave solutions of the cubic nonlinear Schrödinger equation [DS2009]. In [Bridges2015, BR2019], the authors investigate the transition between a hyperbolic and elliptic system of Whitham modulation equations for cubic NLS in the single phase case and a coupled NLS system in the multiphase case. In [Kamchatnov2000], Kamchatnov provides techniques for calculating the Riemann invariants of the cubic nonlinear Schrödinger equation, from which the hyperbolicity of the Whitham modulation equations can be determined. Moreover, the Riemann invariants of a Whitham modulation system encode the asymptotic description of a dispersive shock wave; El and Hoefer provide an extensive discussion on this connection in [EH2016], as well as an analysis of dispersive shock waves in the case of cubic NLS and numerical simulations of dispersive shock wave behaviour for equation (1). El and Hoefer also note that applying the Madelung transformation to equation (1) yields the hydrodynamic system:

$$\begin{aligned} \rho_t + (\rho u)_x &= 0 \\ (\rho u)_t + \left(\rho u^2 + P(\rho)\right)_x &= \left(\frac{1}{4}\rho\,(\ln\rho)_{xx}\right)_x, \end{aligned}$$

with $P(\rho) = \int_0^\rho s f'(s) ds$. Serre's results [Serre2005] are not proven to extend to higher order dispersive terms such as $\frac{1}{4}\rho\,(\ln\rho)_{xx}$. To this end, we do not rely on Serre's techniques, but rather we provide a direct calculation of the characteristics of the Whitham system associated with equation (1).

The main result of this paper, Theorem 3, shows that the Whitham modulation theory for equation (1) correctly predicts the rigorous spectral stability results contained in [LBJM2019]. Leisman et. al. relate spectral instability of periodic travelling wave solutions of equation (1) near the origin of the spectral plane to the breakup of the generalized kernel of the operator of the linearized problem. The authors express the genericity conditions on the Jordan chains of the linearized operator in terms of matrices whose entries are moments of the travelling wave solutions. They use these matrices to derive a normal form for the spectrum of the linearized operator at the origin subject to both longitudinal and transverse perturbations. We give a brief summary of their results, before applying Whitham's averaged Lagrangian method to derive the Whitham modulation equations. We use the genericity conditions given in [LBJM2019] to homogenize the modulation equations, turning them into a quasi-linear system of four equations in four slowly-modulated parameters of the travelling wave solutions. The characteristics of this system are the zeros of a quartic equation which, upon changing variables, we identify as the normal form for the four continuous bands of spectrum emerging from the origin.

We have organized this paper into two main sections. In Section 2, we introduce the relevant spectral stability results of [LBJM2019], leading to a rigorous criterion for modulational instability in Corollary 1. We begin Section 3 with a general overview of Whitham's averaged Lagrangian approach to deriving the modulation equations. We then apply this theory to equation (1) in Section 3.1, from which we conclude in Corollary 2 that the Whitham theory criterion for modulational instability agrees exactly with the rigorous spectral analysis of Section 2. We conclude the paper with a discussion in Section 4 as well as, to the best of our knowledge, several open problems in the area. We include some necessary, though long calculations, as well as, we hope, some useful identities in appendices A to F.

2 Modulational stability from the linear theory

In this section, we follow the analysis of the modulational stability problem contained in [LBJM2019].

We seek travelling wave solutions of equation (1) in the form

$$\psi(x,t) = e^{i\omega t}\phi(x+ct).$$

Writing $y = x + ct$, we express ϕ in polar coordinates

$$\phi(y) = \exp\left(i\left(\theta_0 + \frac{cy}{2} + S(y+y_0)\right)\right) A(y+y_0). \tag{2}$$

Substituting equation (2) into equation (1) and equating real and imaginary parts we have:

$$2A_yS_y + AS_{yy} = 0 \tag{3}$$

$$A_{yy} = -\left(\omega + \frac{c^2}{4}\right)A + AS_y^2 - \zeta f(A^2)A. \tag{4}$$

Integrating equation (3) yields

$$S_y = \frac{\kappa}{A^2}, \tag{5}$$

which can be substituted into equation (4) and upon integrating we have:

$$A_y^2 = 2E - \left(\omega + \frac{c^2}{4}\right)A^2 - \frac{\kappa^2}{A^2} - \zeta F(A^2). \tag{6}$$

The symmetries of equation (1) allow us to eliminate several of the seven parameters $E, \omega, c, \kappa, \zeta, y_0, \theta_0$. In particular, equation (1) is invariant under the transformations (assume $\alpha \in \mathbb{R}$):

- $\psi(x,t) \mapsto \psi(x,t)e^{i\alpha}$ (phase invariance);
- $\psi(x,t) \mapsto \psi(x+\alpha,t)$ (translation invariance);
- $\psi(x,t) \mapsto \psi(x+\alpha t,t)e^{-i\left(\frac{\alpha}{2}x+\frac{\alpha^2}{4}t\right)}$ (Galilean invariance).

Phase and translation invariance means that we can eliminate θ_0 and y_0 from equation (2). Furthermore, by Galilean invariance we can reduce ψ to:

$$\psi(x,t) = e^{i\left(\omega+\frac{c^2}{4}\right)t}e^{iS(x)}A(x). \tag{7}$$

Both equation (7) and equation (6) suggest that we can absorb $\frac{c^2}{4}$ into ω, eliminating c from the equations. This leads us to make the following definition about the parameter space Ω:

Definition 1 (Definition 1 [LBJM2019]). *We define the parameter domain Ω as the open set of parameter values (E,κ,ω,ζ) such that:*

- $\kappa > 0$;
- *The function $P(A) = 2E - \omega A^2 - \frac{\kappa^2}{A^2} - \zeta F(A^2)$ has two positive, real, simple roots $a_-,\ a_+$ with $a_- < a_+$ and $P(A)$ is real and positive for $A \in (a_-, a_+)$.*

As explained in [LBJM2019, Remark 1], parameter values in Ω represent the most generic case and allow us to freely differentiate with respect to the parameters. The cases where $\kappa = 0$ or $R(A)$ has higher order roots correspond to degenerate cases which require their own analyses. We assume henceforth that our parameters are taking values in Ω.

We are interested in periodic solutions A to equations (5) and (6). We make the assumption $A(0) = a_-,\ A(\frac{T_*}{2}) = a_+$, which we can always guarantee by translating the wave. Under these conditions, the function ψ has the quasi-periodic boundary conditions:

$$\psi(T_*) = \psi(0)e^{i\eta} \tag{8}$$

$$\psi_y(T_*) = \psi(0)e^{i\eta}, \tag{9}$$

where the period T_* and the quasi-momentum $\eta = S(T_*) - S(0)$ are functions of the parameters E, κ, ω, ζ. These quantities, and the mass M, can be expressed as derivatives of the classical action:

$$K(E,\kappa,\omega,\zeta) = \int_0^{T_*} A_y^2 dy = 2\int_{a_-}^{a_+} \sqrt{2E - \omega A^2 - \frac{\kappa^2}{A^2} - \zeta F(A^2)}dA. \tag{10}$$

In particular, we have:

$$\begin{aligned} T_*(E,\kappa,\omega,\zeta) &= 2\int_{a_-}^{a_+} \frac{1}{A_y} dA = 2\int_{a_-}^{a_+} \frac{1}{\sqrt{2E - \omega A^2 - \frac{\kappa^2}{A^2} - \zeta F(A^2)}} dA \\ &= \frac{\partial K}{\partial E} \end{aligned} \tag{11}$$

$$\begin{aligned} \eta(E,\kappa,\omega,\zeta) &= S(T_*) - S(0) = 2\int_{a_-}^{a_+} \frac{\kappa}{A^2 A_y} dA \\ &= 2\int_{a_-}^{a_+} \frac{\kappa}{A^2\sqrt{2E - \omega A^2 - \frac{\kappa^2}{A^2} - \zeta F(A^2)}} dA \\ &= -\frac{\partial K}{\partial \kappa} \end{aligned} \tag{12}$$

$$\begin{aligned} M(E,\kappa,\omega,\zeta) &= 2\int_{a_-}^{a_+} \frac{A^2}{A_y} dA = 2\int_{a_-}^{a_+} \frac{A^2}{\sqrt{2E - \omega A^2 - \frac{\kappa^2}{A^2} - \zeta F(A^2)}} dA \\ &= -2\frac{\partial K}{\partial \omega}. \end{aligned} \tag{13}$$

We now consider the linearized problem under the perturbation:

$$\psi(y,t) = e^{i\omega t}\left(\phi(y) + \epsilon e^{i[S(y+y_0) + \frac{cy}{2} + \theta_0]} W(y,t)\right).$$

Since the phase $\exp(i[S(y+y_0) + \frac{cy}{2} + \theta_0])$ is also present in $\phi(y)$ (equation (2)), then we can consider W as a complex-valued perturbation of A. At $O(\epsilon)$ we have:

$$iW_t = W_{yy} + \omega W - S_y^2 W + \zeta f(A^2)W + 2\zeta f'(A^2)A^2 \mathrm{Re}(W) + i(S_{yy}W + 2S_y W_y). \tag{14}$$

Writing $W(y,t) = u(y,t) + iv(y,t)$ in equation (14), we have the two linearized equations

$$\begin{aligned} u_t &= v_{yy} + \omega v - S_y^2 v + \zeta f(A^2) v + S_{yy} u + 2S_y u_y \\ -v_t &= u_{yy} + \omega u - S_y^2 u + \zeta f(A^2) u + 2\zeta f'(A^2)A^2 u - S_{yy} v - 2S_y v_y, \end{aligned}$$

which can be collected into the equation

$$\begin{pmatrix} u \\ v \end{pmatrix}_t = \begin{pmatrix} \mathcal{K} & -\mathcal{L}_- \\ \mathcal{L}_+ & \mathcal{K} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{15}$$

where

$$\begin{aligned} \mathcal{K} &= S_{yy} + 2S_y \partial_y \\ \mathcal{L}_+ &= -\omega - \partial_{yy} + S_y^2 - \zeta f(A^2) - 2\zeta f'(A^2)A^2 \\ \mathcal{L}_- &= -\omega - \partial_{yy} + S_y^2 - \zeta f(A^2). \end{aligned}$$

We can further manipulate equation (15):

$$\begin{pmatrix} u \\ v \end{pmatrix}_t = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{L}_+ & \mathcal{K} \\ \mathcal{K}^T & \mathcal{L}_- \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{16}$$

$$= \mathcal{L} \begin{pmatrix} u \\ v \end{pmatrix}. \tag{17}$$

Since $\mathcal{L}$ is the product of a skew-symmetric operator and symmetric operator then equation (17) defines a Hamiltonian eigenvalue problem. From here, the idea is to analyze the Jordan chains of $\mathcal{L}$. In particular, we wish to find the genericity conditions on the Jordan structure of the linearized operator in terms of the derivatives of K. Having found these conditions, we can then perturb $\mathcal{L}$ and analyze the affect this has on the generalized kernels. The following result summarizes the genericity conditions.

Theorem 1 (Theorem 1 [LBJM2019]). *Assume that:*

- $(E, \kappa, \omega, \zeta) \in \Omega$;
- $F(x)$ *is linearly independent from* $1, x, \frac{1}{x}$;
- $A(y)$ *is non-constant.*

The generalized kernel of $\mathcal{L}$ (defined in equation (17)) generically takes the form of a direct sum of two Jordan chains of length two. Making the definitions

$$\sigma := \{T_*, \eta\}_{E,\kappa} = T_{*E}\eta_\kappa - T_{*\kappa}\eta_E$$

$$D := \begin{vmatrix} K_{\kappa\kappa} & K_{\kappa E} & K_{\kappa\omega} & T_* \\ K_{\kappa E} & K_{EE} & K_{E\omega} & 0 \\ K_{\kappa\omega} & K_{E\omega} & K_{\omega\omega} & 0 \\ T_* & 0 & 0 & -M \end{vmatrix},$$

then the genericity conditions on these Jordan chains are:

$$\sigma \neq 0 \tag{18}$$

$$D \neq 0. \tag{19}$$

Remark 1. We have chosen to omit several results from Theorem 1 in [LBJM2019], such as the explicit calculation of the Jordan chains of $\mathcal{L}$. This is because the Whitham theory in the later sections of this paper makes use of just the genericity conditions on these chains.

Remark 2. In the proof of Theorem 1, the authors calculate the determinant D in terms of T_*, M, η and the Poisson bracket quantities defined in appendix A. Using the Dodgson-Jacobi-Desnanot condensation identity, they calculate:

$$-\frac{\sigma^3}{4} D = \begin{vmatrix} a_2 & b_2 \\ b_2 & d_2 \end{vmatrix},$$

where a_2, b_2, d_2 are defined in appendix A.

We now examine the breakup of the Jordan chains of $\mathcal{L}$ under perturbations of the quasi-periodic boundary conditions. We start with the following proposition:

Proposition 1 (Proposition 1 [LBJM2019]). *Let $\mathcal{L}$ be an operator with compact resolvent. Suppose further that $\mathcal{L}$ has a d-dimensional kernel spanned by $\{\mathbf{u}_{2j}\}_{j=0}^{d-1}$ and a d-dimensional first generalized kernel spanned by $\{\mathbf{u}_{2j+1}\}_{j=0}^{d-1}$ satisfying $\mathcal{L}\mathbf{u}_{2j+1} = \mathbf{u}_{2j}$. Suppose similarly that $\mathcal{L}$ has a left basis satisfying $\mathbf{v}_{2j+1}\mathcal{L} = 0,\ \mathbf{v}_{2j}\mathcal{L} = \mathbf{v}_{2j+1}$. Consider a perturbation of the form $\mathcal{L} + \mu\mathcal{L}^{(1)} + \mu^2\mathcal{L}^{(2)}$, where $\mathcal{L}^{(1)}(\lambda - \mathcal{L})^{-1}$, $\mathcal{L}^{(1)}(\lambda - \mathcal{L}^{(1)})^{-1}\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$ are bounded operators. Suppose that the first order perturbation satisfies the conditions:*

$$\mathbf{v}_{2j+1}\mathcal{L}^{(1)}\mathbf{u}_{2k} = 0 \quad \forall j, k = 0, \dots, d-1.$$

Then, to leading order in μ, the $2d$-dimensional generalized kernel breaks up into $2d$ eigenspaces, with the eigenvalues given by

$$\lambda(\mu) = \lambda_1\mu + O(\mu^2),$$

where λ_1 is a root of the polynomial

$$\det(\lambda_1^2\mathbf{M}^{(2)} + \lambda_1\mathbf{M}^{(1)} + \mathbf{M}^{(0)}) = 0,$$

with the $d \times d$ *matrices* $\mathbf{M}^{(i)}$ *are defined as:*

$$
\begin{aligned}
\mathbf{M}^{(2)}_{j,k} &= \mathbf{v}_{2j+1}\mathbf{u}_{2k+1} \\
\mathbf{M}^{(1)}_{j,k} &= -\mathbf{v}_{2j}\mathcal{L}^{(1)}\mathbf{u}_{2k} - \mathbf{v}_{2j+1}\mathcal{L}^{(1)}\mathbf{u}_{2k+1} \\
\mathbf{M}^{(0)}_{j,k} &= \mathbf{v}_{2j+1}\mathcal{L}^{(1)}\mathcal{L}^{-1}\mathcal{L}^{(1)}\mathbf{u}_{2k} - \mathbf{v}_{2j+1}\mathcal{L}^{(2)}\mathbf{u}_{2k}.
\end{aligned}
$$

Proposition 1 allows us to relate the eigenvalues λ near the origin to the break up of the generalized kernels of $\mathcal{L}$ under small perturbations. We will first deal with the case of longitudinal perturbations, that is, the 1D NLS equation equation (1). In particular, when we substitute

$$
\begin{pmatrix} u \\ v \end{pmatrix} = e^{\lambda t}\mathbf{u}(y)
$$

into the linearized problem equation (17) we have:

$$
\mathcal{L}\mathbf{u} = \lambda\mathbf{u}. \tag{20}
$$

Equation (20) can be written as a system of four first order ODEs in y with T_*-periodic coefficients, and so by applying Floquet theory we see that the spectrum of $\mathcal{L}$ is the union over all $\mu \in (-\frac{\pi}{T_*}, \frac{\pi}{T_*}]$ of the spectrum of $\mathcal{L}$ with quasi-periodic boundary conditions:

$$
\begin{aligned}
\mathbf{u}(T_*) &= e^{i\mu T_*}\mathbf{u}(0) \\
\mathbf{u}_y(T_*) &= e^{i\mu T_*}\mathbf{u}_y(0).
\end{aligned}
$$

If instead we write $\mathbf{u} = e^{i\mu y}\mathbf{v}$, then $\mathbf{v}$ has periodic boundary conditions, and the μ-dependency is instead captured by a μ-dependent operator:

$$
\begin{aligned}
&\mathcal{L}(\mu)\mathbf{v} = \lambda\mathbf{v} \\
&\mathcal{L}(\mu) := \mathcal{L} + 2i\mu \begin{pmatrix} S_y & \partial_y \\ -\partial_y & S_y \end{pmatrix} + \mu^2 \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.
\end{aligned}
$$

Taking

$$
\begin{aligned}
\mathcal{L}^{(1)} &= 2i \begin{pmatrix} S_y & \partial_y \\ -\partial_y & S_y \end{pmatrix} \\
\mathcal{L}^{(2)} &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}
\end{aligned}
$$

allows us to apply proposition 1, leading to the following theorem.

Theorem 2 (Corollary 1 [LBJM2019])**.** *Suppose that* $\mathcal{L}$ *is defined as in equation* (17)*, and further suppose that the genericity conditions of Theorem 1 apply. For small values of the Floquet exponent* μ*, the normal form for the four continuous bands of spectrum emerging from the origin in the spectral plane is:*

$$
\det\left(\lambda^2 \begin{pmatrix} a_2 & b_2 \\ b_2 & d_2 \end{pmatrix} + \lambda\mu \begin{pmatrix} a_1 & b_1 \\ b_1 & d_1 \end{pmatrix} + \mu^2 \begin{pmatrix} a_0 & b_0 \\ b_0 & d_0 \end{pmatrix}\right) + O(5) = 0, \tag{21}
$$

where $O(5)$ *consists of terms* $\lambda^i\mu^j$ *with* $i, j > 0$ *and* $i + j \geq 5$*. The matrix entries are given in appendix A.*

Corollary 1 (Modulational instability criterion)**.** *If any of the roots* λ *of equation* (21) *are not purely imaginary, then the periodic travelling wave* ψ *about which we have linearized is modulationally unstable.*

3 Whitham modulation theory

Our goal in this section is to provide a formal Whitham theory calculation which results in the same criterion for modulational instability as the rigorous results in the previous section, in particular Theorem 2 and Corollary 1. We want to show that the equation for the characteristics of the Whitham modulation equations is equivalent to the normal form equation (21). There are several ways to derive the Whitham modulation equations; we choose to use the averaged Lagrangian formulation and the variational principle from Whitham's work [Whi1965B, Whi1970, Whi1999], since the averaged Lagrangian closely resembles the classical action K (cf. equation (10)) from the rigorous linear theory.

As a side note, many of the symbols of the previous section will be re-used but with the subscript $*$. These subscripted symbols are seen as completely new, and not connected to their un-subscripted counterparts. We start with the Lagrangian of equation (1):

$$L(\psi_t, \psi_x, \psi, \overline{\psi}_t, \overline{\psi}_x, \overline{\psi}) = i(\overline{\psi}\psi_t - \psi\overline{\psi}_t) + 2|\psi_x|^2 - 2\zeta F(|\psi|^2), \tag{22}$$

which satisfies the Euler-Lagrange equations:

$$\begin{aligned} L_\psi - \partial_t L_{\psi_t} - \partial_x L_{\psi_x} = 0 \\ L_{\overline{\psi}} - \partial_t L_{\overline{\psi}_t} - \partial_x L_{\overline{\psi}_x} = 0. \end{aligned}$$

We seek solutions to equation (1) of the form:

$$\psi(x,t) = r(x,t)e^{i\varphi(x,t)}.$$

Upon subsitution, we have (for imaginary and real parts respectively)

$$r_t = 2\varphi_x r_x + \varphi_{xx} r \tag{23}$$

$$-\varphi_t r = r_{xx} - \varphi_x^2 r + \zeta f(r^2)r, \tag{24}$$

with the Lagrangian:

$$L(r_x, r, \varphi_t, \varphi_x) = -2\varphi_t r^2 + 2(r_x)^2 + 2(\varphi_x)^2 r^2 - 2\zeta F(r^2). \tag{25}$$

We now seek periodic solutions r, φ, with period normalized to 2π. Introducing the variable $\theta = kx - \omega_* t$, we let

$$\begin{aligned} r(x,t) &= R(\theta) \\ \varphi(x,t) &= \beta x - \gamma_* t + \Phi(\theta). \end{aligned}$$

The pseudo-phase component $\beta x - \gamma_* t$ is a necessary generalization since only the derivatives of φ appear in L (cf. equation (25)). Consequently, the pseudo-phase ensures that the quantities φ_x and φ_t have mean β and $-\gamma_*$ respectively when averaged over one period. This is reminiscent of the quasi-periodic boundary conditions set for ψ in equation (8). Substituting into equations (23) to (25) yields:

$$-\omega_* R_\theta = 2kR_\theta(\beta + k\Phi_\theta) + k^2\Phi_{\theta\theta}R \tag{26}$$

$$R(\gamma_* + \omega_*\Phi_\theta) = k^2 R_{\theta\theta} - R(\beta + k\Phi_\theta)^2 + \zeta f(R^2)R \tag{27}$$

$$L(kR_\theta, R, -\gamma - \omega_*\Phi_\theta, \beta + k\Phi_\theta) = 2R^2(\gamma_* + \omega_*\Phi_\theta) + 2k^2R_\theta^2 + 2R^2(\beta + k\Phi_\theta)^2 - 2\zeta F(R^2). \tag{28}$$

For a slow modulation, we consider the the parameters $\omega_*,\ k,\ \gamma_*,\ \beta$ to be functions of $X = \epsilon x$ and $T = \epsilon t$. We write:

$$\theta = \epsilon^{-1}\Theta(X,T), \quad \beta x - \gamma_* t = \widetilde{\theta} = \epsilon^{-1}\widetilde{\Theta}(X,T), \tag{29}$$

and define:

$$-\omega_*(X,T) = \Theta_T, \quad k(X,T) = \Theta_X, \quad -\gamma_*(X,T) = \widetilde{\Theta}_T, \quad \beta(X,T) = \widetilde{\Theta}_X. \tag{30}$$

Our functions R and Φ are also written in these variables as:

$$R = R(\theta, X, T; \epsilon), \quad \Phi = \Phi(\theta, X, T; \epsilon).$$

The fact that these functions are evolving on both slow and fast scales is called two-timing, and is closely examined in [Whi1970, Whi1999]. When averaging the Lagrangian, we consider R and Φ to be functions of three independent variables $\theta,\ X,\ T$, even though θ is a function of X and T in equation (29). This extra flexibility ensures that secular terms in the asymptotic expansions are suppressed, and in conjunction with the variational principle this is equivalent to a WKB approximation [Whi1999]. According to Whitham's theory [Whi1999, Whi1970], the leading order modulation equations can be derived from the variational principle:

$$\delta \int\int \frac{1}{2\pi}\int_0^{2\pi} L(kR_\theta, R, -\gamma_* - \omega_*\Phi_\theta, \beta + k\Phi_\theta)\,d\theta dX dT = 0. \tag{31}$$

For functions h which vanish on the (θ, X, T) boundary, equation (31) means that, for variations in R (denoted δR), we have:

$$\begin{aligned} \delta R: \quad & \int\int \frac{1}{2\pi}\int_0^{2\pi} L(k(R+h)_\theta, R+h, -\gamma_* - \omega_*\Phi_\theta, \beta + k\Phi_\theta)\,d\theta dX dT = 0 \\ & \implies \int\int \frac{1}{2\pi}\int_0^{2\pi} h_\theta L_{R_\theta} + hL_R + O(h^2)\,d\theta dX dT = 0. \end{aligned}$$

Integrating by parts, we end up with

$$\int\int \frac{1}{2\pi}\int_0^{2\pi} h(L_R - \partial_\theta L_{R_\theta})d\theta dX dT = 0.$$

Since this is true for all such h, we conclude that:

$$\partial_\theta L_{R_\theta} - L_R = 0. \tag{32}$$

Equation (32) is simply equation (27), one of the Euler-Lagrange equations for L (cf. equation (28)). The other Euler-Lagrange equation comes from considering variations in Φ, which yields:

$$\partial_\theta L_{\Phi_\theta} = 0. \tag{33}$$

Equation (33) has an immediate first integral:

$$B(X,T) = L_{\Phi_\theta}, \tag{34}$$

so B is constant with respect to θ, however it is now added to the ensemble of slowly-varying parameters. Equation (33) is in fact a multiple of equation (26) (the factor is $4R$), so finding B is equivalent to finding a first integral of equation (26). Equation (32) also admits an integral, since this equation is an ODE in the variable θ. To see this, we multiply equation (32) by R_θ:

$$\begin{aligned} R_\theta\partial_\theta L_{R_\theta} - R_\theta L_R &= 0 \\ \implies \partial_\theta(R_\theta L_{R_\theta}) - R_{\theta\theta}L_{R_\theta} - R_\theta L_R &= 0 \\ \implies \partial_\theta(R_\theta L_{R_\theta}) - \partial_\theta L + B\Phi_{\theta\theta} &= 0. \end{aligned}$$

Note that we have used equation (33) to write the θ-derivative of L as:

$$\partial_\theta L = R_{\theta\theta}L_{R_\theta} + R_\theta L_R + B\Phi_{\theta\theta}.$$

Integrating, we have:

$$R_\theta L_{R_\theta} + B\Phi_\theta - L = A_*(X,T), \tag{35}$$

for a slowly-varying parameter A_*. For variations in Θ, we consider $\Theta(X,T) + h(X,T)$ for an appropriate function h vanishing on the (X,T) boundary. This means that k will be replaced by $k + h_X$, and similarly ω_* by $\omega - h_T$:

$$\delta\Theta:\quad \int\int \frac{1}{2\pi}\int_0^{2\pi} L((k+h_X)R_\theta, R, -\gamma - (\omega_* - h_T)\Phi_\theta, \beta + (k+h_X)\Phi_\theta)d\theta dX dT = 0. \tag{36}$$

Writing the averaged Lagrangian as

$$\overline{L}(\omega_*, k, \gamma_*, \beta) = \frac{1}{2\pi}\int_0^{2\pi} L d\theta, \tag{37}$$

equation (36) implies:

$$\int\int h_X\overline{L}_k - h_T\overline{L}_{\omega_*} + O(h^2)dXdT = 0. \tag{38}$$

From equation (38), we have:

$$\int\int h(\partial_T\overline{L}_{\omega_*} - \partial_X\overline{L}_k)dXdT = 0,$$

which is true for all such h, and so we conclude that:

$$\partial_T\overline{L}_{\omega_*} - \partial_X\overline{L}_k = 0. \tag{39}$$

Equation (39) is called a modulation equation since the quantities involved are functions of the slowly-modulated variables X and T. Similarly for variations $\delta\widetilde{\Theta}$, we have another modulation equation:

$$\partial_T\overline{L}_{\gamma_*} - \partial_X\overline{L}_\beta = 0. \tag{40}$$

The last two modulation equations are associated with the conservation of waves:

$$\partial_T k + \partial_X\omega_* = 0 \tag{41}$$

$$\partial_T\beta + \partial_X\gamma_* = 0, \tag{42}$$

which is a consequence of requiring the mixed partial derivatives of $\Theta(X,T)$ and $\widetilde{\Theta}(X,T)$ to be equal. At this point, we have derived four modulation equations in the four parameters $(\omega_*, k, \gamma_*, \beta)$. In order to evaluate the averaged Lagrangian equation (37), we need to solve equations (34) and (35) for R, Φ. Instead, Whitham proposes using the first integrals $A_*(X,T)$ and $B_*(X,T)$ to simplify the form of the averaged Lagrangian [Whi1999]. We can avoid directly computing R, Φ, and moreover we can absorb the dispersion relations into one variational principle. To this end, we consider a Hamiltonian transformation with the variables:

$$\Pi_1 = L_{R_\theta}, \quad \Pi_2 = L_{\Phi_\theta} = B. \tag{43}$$

These are generalized momenta, and so we apply a Legendre transform to L in order to eliminate R_θ and Φ_θ from the equations:

$$H(\Pi_1, \Pi_2, R, \Phi; \omega_*, k, \gamma_*, \beta) = R_\theta \Pi_1 + \Phi_\theta \Pi_2 - L. \tag{44}$$

We see from the transformation that

$$R_\theta = \partial_{\Pi_1} H, \quad \Phi_\theta = \partial_{\Pi_2} H, \tag{45}$$

and it follows from equation (32) and equation (33) respectively that

$$\partial_\theta \Pi_1 = -\partial_R H, \quad \partial_\theta \Pi_2 = -\partial_\Phi H = 0. \tag{46}$$

Equations (45) and (46) are Hamilton's equations. We can now write the earlier variational principle equation (31) as:

$$\delta \int \int \overline{L} dX dT = 0, \tag{47}$$

where

$$\overline{L} = \frac{1}{2\pi} \int_0^{2\pi} (R_\theta \Pi_1 + B\Phi_\theta - H) d\theta \tag{48}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} (R_\theta \Pi_1 - H) d\theta, \tag{49}$$

since Φ is 2π-periodic. The averaged Lagrangian is now a function of H, B and the previous four parameters. We note that Hamilton's equations (45) and (46) follow from the independent variations $\delta\Pi_1, \delta R, \delta\Pi_2, \delta\Phi$ in equation (48). We use this extension as Whitham describes in [Whi1970, Whi1999]. Next, we observe that equation (44) is the same as equation (35), so we identify:

$$H(\Pi_1, \Pi_2, R; \omega_*, k, \gamma_*, \beta) = A_*(X,T). \tag{50}$$

The stationary values of equation (48) also satisfy equation (50), meaning that we can restrict the stationary values of equation (48) to the class of functions $R,\ \Phi,\ \Pi_1,\ \Pi_2$ which satisfy equation (50). Importantly, this is the only restriction we make; using the dispersion relation or any information about the forms of the solutions Π_1 and R_θ (equation (45)) would result in $\overline{L}$ having no variation. We relabel $H(\Pi_1, \Pi_2, R, \Phi; \omega_*, k, \gamma_*, \beta)$ as $H(X,T)$ in equation (48), and we solve equation (50) for $\Pi_1(R; H, \omega_*, k, B, \gamma_*, \beta)$ which yields:

$$\mathscr{L}(H, \omega_*, k, B, \gamma_* \beta) = \frac{1}{2\pi} \oint \Pi_1 dR - H, \tag{51}$$

where the integral is taken around the orbit of R. The variational principle equation (47) can now be written as

$$\delta \int \int \mathscr{L}(H, \omega_*, k, B, \gamma_* \beta) dX dT = 0. \tag{52}$$

As mentioned earlier, we have now added the two parameters H and B, however we have exchanged the variational equations $\delta R, \delta\Pi_1$ and $\delta\Phi, \delta\Pi_2$ for δH and δB:

$$\delta H : \quad \mathscr{L}_H = 0 \tag{53}$$

$$\delta B : \quad \mathscr{L}_B = 0. \tag{54}$$

Equations (53) and (54) are relations between the parameters of the periodic wavetrain, and so they are in fact the dispersion relations. Our aim is to use these equations in order to eliminate two of the six parameters from the four modulation equations, resulting in a homogeneous system of first-order PDEs. Using equations (41), (42) and (52), we now have a complete picture of the Whitham theory:

$$\partial_T \mathscr{L}_{\omega_*} - \partial_X \mathscr{L}_k = 0 \tag{55}$$

$$\partial_T \mathscr{L}_{\gamma_*} - \partial_X \mathscr{L}_\beta = 0 \tag{56}$$

$$\partial_T k + \partial_X \omega_* = 0 \tag{57}$$

$$\partial_T \beta + \partial_X \gamma_* = 0. \tag{58}$$

3.1 Whitham theory applied to NLS

In this subsection, we apply the more general theory and observations from section 3 to equation (1). We provide a direct computation of the characteristics of the Whitham modulation equations, from which we conclude that the modulational instability criterion from the Whitham theory agrees with the spectral analysis in section 2. For the sake of brevity, we include more detailed calculations in appendices B to F.

We calculate $\mathscr{L}$ in terms of the parameters $H, \omega_*, k, B, \gamma_*\beta$ using equations (26) to (28). Firstly, $\Pi_2 = B$ amounts to taking an integral of equation (26), which yields:

$$B = 2\omega_* R^2 + 4kR^2(\beta + k\Phi_\theta). \tag{59}$$

Next, we calculate H using equation (44), eliminating Φ_θ and R_θ via equation (59) and $\Pi_1 = L_{R_\theta} = 4k^2 R_\theta$:

$$\begin{aligned} H &= \frac{\Pi_1^2}{4k^2} + B\Phi_\theta - \frac{\Pi_1^2}{8k^2} - 2R^2(\gamma_* + \omega_*\Phi_\theta) - 2R^2(\beta + k\Phi_\theta)^2 + 2\zeta F(R^2) \\ &= \frac{\Pi_1^2}{8k^2} - \frac{\beta B}{k} - \frac{\omega_* B}{2k^2} + 2R^2\left(\frac{\beta\omega_*}{k} + \frac{\omega_*^2}{4k^2} - \gamma_*\right) + \frac{B^2}{8k^2R^2} + 2\zeta F(R^2) \\ \implies \frac{\Pi_1^2}{8k^2} &= H + \frac{\beta B}{k} + \frac{\omega_* B}{2k^2} - 2R^2\left(\frac{\beta\omega_*}{k} + \frac{\omega_*^2}{4k^2} - \gamma_*\right) - \frac{B^2}{8k^2R^2} - 2\zeta F(R^2). \end{aligned}$$

This can also be written as:

$$2k^2R_\theta^2 = H + \frac{\beta B}{k} + \frac{\omega_* B}{2k^2} - 2R^2\left(\frac{\beta\omega_*}{k} + \frac{\omega_*^2}{4k^2} - \gamma_*\right) - \frac{B^2}{8k^2R^2} - 2\zeta F(R^2), \tag{60}$$

which we identify as the first integral of equation (27) once Φ_θ has been eliminated, i.e:

$$k^2 R_{\theta\theta} = -R\left(\frac{\beta\omega_*}{k} + \frac{\omega_*^2}{4k^2} - \gamma_*\right) + \frac{B^2}{16k^2R^3} - \zeta f(R^2)R. \tag{61}$$

The averaged Lagrangian is hence:

$$\mathscr{L} = \frac{k\sqrt{2}}{\pi}\oint\sqrt{H + \frac{\beta B}{k} + \frac{\omega_* B}{2k^2} - 2R^2\left(\frac{\beta\omega_*}{k} + \frac{\omega_*^2}{4k^2} - \gamma_*\right) - \frac{B^2}{8k^2R^2} - 2\zeta F(R^2)dR} - H. \tag{62}$$

Following Whitham's examples [Whi1965A, Whi1999], we introduce a function W which is essentially the classical action:

$$W(H, \omega_*, k, B, \gamma_*, \beta) = \frac{1}{2\pi}\oint\frac{\Pi_1}{2k}dR \tag{63}$$

$$= \frac{\sqrt{2}}{2\pi}\oint\sqrt{H + \frac{\beta B}{k} + \frac{\omega_* B}{2k^2} - 2R^2\left(\frac{\beta\omega_*}{k} + \frac{\omega_*^2}{4k^2} - \gamma_*\right) - \frac{B^2}{8k^2R^2} - 2\zeta F(R^2)dR}. \tag{64}$$

This allows us to write equation (62) as:

$$\mathscr{L} = 2kW - H. \tag{65}$$

For a generic evaluation of W from equation (64), we require assumptions similar to those listed in definition 1. We derive these assumptions by equating our periodic solution $\psi(x,t) = R(\theta)\exp(i(\beta x - \gamma_* t + \Phi(\theta)))$ with the periodic solution equation (2), which yields (once translation invariance and phase invariance are taken into account):

$$R(\theta) = A(y) \tag{66}$$

$$\beta x - \gamma_* t + \Phi(\theta) = \left(\omega - \frac{c^2}{4}\right)t + \frac{cy}{2} + S(y) \tag{67}$$

Note that we have redefined ω as $\omega - \frac{c^2}{4}$ as suggested by equation (7). Taking x and t derivatives of equation (66), we have:

$$kR_\theta = A_y, \quad -\omega_* R_\theta = cA_y \tag{68}$$

$$\implies c = -\frac{\omega_*}{k}. \tag{69}$$

Similarly for equation (67) we have:

$$\beta + k\Phi_\theta = \frac{c}{2} + S_y, \quad -\gamma_* - \omega_*\Phi_\theta = \omega - \frac{c^2}{4} + \frac{c^2}{2} + cS_y \tag{70}$$

$$\implies \omega = \frac{\beta\omega_*}{k} + \frac{\omega_*^2}{4k^2} - \gamma_*. \tag{71}$$

Substituting equations (68) to (70) into equation (5) yields:

$$\beta + \frac{\omega_*}{2k} + k\Phi_\theta = \frac{\kappa}{R^2},$$

and recalling equation (59) we identify:

$$\kappa = \frac{B}{4k}. \tag{72}$$

Finally, using equations (66), (68), (69), (71) and (72), we see that equation (60) is twice equation (6), with:

$$E = \frac{1}{4}\left(H + \frac{\beta B}{k} + \frac{\omega_* B}{2k^2}\right). \tag{73}$$

Equations (69) and (71) to (73) now express the parameters of the linear theory in terms of our Whitham parameters. Recalling definition 1, we have that the parameters $(H, \omega_*, k, B, \gamma_*, \beta)$ exist such that:

- $\frac{B}{k} > 0$;
- The function $P(R) = H + \frac{\beta B}{k} + \frac{\omega_* B}{2k^2} - 2R^2\left(\frac{\beta\omega_*}{k} + \frac{\omega_*^2}{4k^2} - \gamma_*\right) - \frac{B^2}{8k^2R^2} - 2\zeta F(R^2)$ has two positive, real, simple roots $R_- = a_-$ and $R_+ = a_+$, with $R_- < R_+$ and $P(R)$ real and positive for $R \in (R_-, R_+)$.

We make the change of variables:

$$U = \frac{\omega_*}{k}$$
$$J = \frac{B}{k}.$$

This eliminates the explicit dependence of $P(R)$ on k, which will be advantageous when we calculate the characteristics of the modulation equations. Also note that the two roots $R_-,\ R_+$ are now functions of $H, U, J, \gamma, \beta, \zeta$. Updating W in equation (64), we have:

$$W(H,U,J,\gamma,\beta,\zeta) = \frac{\sqrt{2}}{\pi}\int_{R_-}^{R_+}\sqrt{H + \beta J + \frac{UJ}{2} - 2R^2\left(\beta U + \frac{U^2}{4} - \gamma_*\right) - \frac{J^2}{8R^2} - 2\zeta F(R^2)}dR. \tag{74}$$

Moreover, we can relate the classical action from equation (10) with W:

$$K = \pi W. \tag{75}$$

We now calculate the derivatives of W with respect to the parameters:

$$W_H = \frac{\sqrt{2}}{\pi}\int_{R_-}^{R_+}\frac{1}{2\sqrt{H + \beta J + \frac{UJ}{2} - 2R^2\left(\beta U + \frac{U^2}{4} - \gamma_*\right) - \frac{J^2}{8R^2} - 2\zeta F(R^2)}}dR \tag{76}$$

$$W_J = W_H\left(\beta + \frac{U}{2}\right) - \eta_* \tag{77}$$

$$\eta_* := \frac{\sqrt{2}}{\pi}\int_{R_-}^{R_+}\frac{J}{8R^2\sqrt{H + \beta J + \frac{UJ}{2} - 2R^2\left(\beta U + \frac{U^2}{4} - \gamma_*\right) - \frac{J^2}{8R^2} - 2\zeta F(R^2)}}dR \tag{78}$$

$$M_* := W_{\gamma_*} = \frac{\sqrt{2}}{\pi}\int_{R_-}^{R_+}\frac{R^2}{\sqrt{H + \beta J + \frac{UJ}{2} - 2R^2\left(\beta U + \frac{U^2}{4} - \gamma_*\right) - \frac{J^2}{8R^2} - 2\zeta F(R^2)}}dR \tag{79}$$

$$W_U = \frac{J}{2}W_H - M_*\left(\beta + \frac{U}{2}\right) \tag{80}$$

$$W_\beta = JW_H - UM_* \tag{81}$$

$$W_\zeta = -\frac{\sqrt{2}}{\pi}\int_{R_-}^{R_+}\frac{F(R^2)}{\sqrt{H + \beta J + \frac{UJ}{2} - 2R^2\left(\beta U + \frac{U^2}{4} - \gamma_*\right) - \frac{J^2}{8R^2} - 2\zeta F(R^2)}}dR. \tag{82}$$

We choose the notation η_* and M_* because these integrals are closely related to the η and M from the linear theory. We give the relations in appendix B. Using equation (65), we have:

$$\begin{aligned}\mathscr{L}_{\omega_*} &= 2W_U \\ \mathscr{L}_k &= 2W - 2UW_U - 2JW_J \\ \mathscr{L}_{\gamma_*} &= 2kW_{\gamma_*} \\ \mathscr{L}_\beta &= 2kW_\beta\end{aligned}$$

We are now in a position to write the modulation equations in terms of the derivatives of W. For equations (55) and (56) we have:

$$\partial_T W_U + U\partial_X W_U + J\partial_X W_J + W_U U_X + W_J J_X - \partial_X W = 0 \tag{83}$$

$$\partial_T \left(2kW_{\gamma_*}\right) - \partial_X \left(2kW_\beta\right) = 0. \tag{84}$$

We now seek to use the variational equations equations (53) and (54) to eliminate two parameters from the Whitham system. Using equation (53) with equation (65) we have:

$$\mathscr{L}_H = 0 \tag{85}$$

$$\implies W_H = \frac{1}{2k}. \tag{86}$$

Similarly for equation (54):

$$\mathscr{L}_B = 0 \tag{87}$$

$$\implies W_J = 0. \tag{88}$$

Since W and its derivatives have no explicit dependence on k, then equation (86) is in fact the dispersion relation for k in terms of the other parameters $H, U, J, \gamma_*, \beta, \zeta$. Substituting equation (86) into the modulation equations allows us to eliminate k entirely. This is particularly relevant for equation (57):

$$\begin{aligned}\partial_T k + \partial_X(kU) = 0 \\ \implies \partial_T\left(\frac{1}{2W_H}\right) + \partial_X\left(\frac{U}{2W_H}\right) = 0,\end{aligned}$$

which simplifies to:

$$\partial_T W_H + U\partial_X W_H - W_H U_X = 0.$$

Equation (83) is already free of k, but equation (84) can be simplified using equations (79), (81) and (86):

$$W_H\partial_T M_* - M_*\partial_T W_H + UW_H\partial_X M_* - UM_*\partial_X W_H + W_H M_* U_X - W_H^2 J_X = 0.$$

We also observe that equation (88) can eliminate terms in equation (83). The four modulation equations are now:

$$\beta_T + \gamma_{*X} = 0 \tag{89}$$

$$\partial_T W_H + U\partial_X W_H - W_H U_X = 0 \tag{90}$$

$$\partial_T W_U + U\partial_X W_U + W_U U_X - \partial_X W = 0 \tag{91}$$

$$W_H\partial_T M_* - M_*\partial_T W_H + UW_H\partial_X M_* - UM_*\partial_X W_H + W_H M_* U_X - W_H^2 J_X = 0. \tag{92}$$

The derivatives of W_J with respect to the parameters H and J do no vanish simultaneously, so we can apply the implicit function theorem in order to eliminate another parameter. Recalling the genericity condition $\sigma \neq 0$ from Theorem 1, we have that $\sigma_* = \{W_H, W_J\}_{J,H}$ defined in appendix C is non-zero, which implies that at least one of W_{HJ} and W_{JJ} are non-zero. We first assume that $W_{HJ} \neq 0$. By the implicit function theorem, there exists a continuously differentiable function g defined on the appropriate parameter space such that:

$$H = g(U, J, \gamma_*, \beta). \tag{93}$$

Next we take derivatives of the equation $W_J = 0$:

$$\begin{aligned}\partial_J W_J(g, U, J, \gamma_*, \beta) = 0 \\ \implies g_J W_{HJ} + W_{JJ} = 0\end{aligned}$$

Similarly for the other derivatives:

$$g_J W_{HJ} = -W_{JJ} \tag{94}$$
$$g_U W_{HJ} = -W_{UJ} \tag{95}$$
$$g_{\gamma_*} W_{HJ} = -W_{\gamma_* J} \tag{96}$$
$$g_\beta W_{HJ} = -W_{\beta J}. \tag{97}$$

Equations (94) to (97) allow us to expand the X and T derivatives of W_H and M_* in equations (90) to (92) by using the chain rule. For example:

$$\begin{aligned}\partial_T W_H &= g_T W_{HH} + J_T W_{HJ} + U_T W_{HU} + \gamma_{*T} W_{H\gamma_*} + \beta_T W_{H\beta} \\ &= (g_J J_T + g_U U_T + g_{\gamma_*}\gamma_{*T} + g_\beta \beta_T) W_{HH} + J_T W_{HJ} + U_T W_{HU} + \gamma_{*T} W_{H\gamma_*} + \beta_T W_{H\beta} \\ &= (g_J W_{HH} + W_{HJ}) J_T + (g_U W_{HH} + W_{HU}) U_T + (g_{\gamma_*} W_{HH} + W_{H\gamma_*})\gamma_{*T} + (g_\beta W_{HH} + W_{H\beta})\beta_T.\end{aligned}$$

If we multiply the derivative by W_{HJ}, we have:

$$\begin{aligned}W_{HJ}\partial_T W_H &= (W_{HJ}^2 - W_{HH}W_{JJ})J_T + (W_{HU}W_{HJ} - W_{HH}W_{UJ})U_T + (W_{H\gamma_*}W_{HJ} - W_{HH}W_{\gamma_* J})\gamma_{*T} \\ &\quad + (W_{H\beta}W_{HJ} - W_{HH}W_{\beta J})\beta_T \\ &= \{W_H, W_J\}_{J,H} J_T + \{W_H, W_U\}_{J,H} U_T + \{W_H, W_{\gamma_*}\}_{J,H}\gamma_{*T} + \{W_H, W_\beta\}_{J,H}\beta_T.\end{aligned}$$

Using appendix C, we can identify

$$\begin{aligned}\{W_H, W_J\}_{J,H} &= \sigma_* \\ \{W_H, W_{\gamma_*}\}_{J,H} &= -\{W_H, W_J\}_{H,\gamma_*} = -\rho_*,\end{aligned}$$

where we have used the properties of Poisson brackets quoted in appendix C.

Remark 3. An important consideration when computing these Poisson brackets is the order of differentiation and substitution. The correct order is to compute the Poisson bracket and then substitute $H = g(U, J, \gamma_*, \beta)$ into the expression, rather than applying the chain rule within a Poisson bracket. This is because the Poisson brackets produce identities that hold regardless of the equations tying the parameters together; we are using them as a convenient notation for leveraging the symmetry of the mixed partial derivatives of W. As an example, we make the substitution $H = g(U, J, \gamma_*, \beta)$ into the identity $\{W_H, W_{\gamma_*}\}_{J,H} = -\{W_H, W_J\}_{H,\gamma_*} = -\rho_*$ instead of computing $\partial_J W_H(g, U, J, \gamma_*, \beta)$ and other such derivatives.

For $\{W_H, W_U\}_{J,H}$, we have:

$$\begin{aligned}\{W_H, W_U\}_{J,H} &= \left\{W_H, \frac{JW_H}{2} - \left(\beta + \frac{U}{2}\right) W_{\gamma_*}\right\}_{J,H} \\ &= \{W_H, \frac{JW_H}{2}\}_{J,H} - \left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,H} \\ &= \nu_* + \left(\beta + \frac{U}{2}\right)\rho_*,\end{aligned}$$

and similarly we calculate

$$\{W_H, W_\beta\}_{J,H} = 2\nu_* + U\rho_*,$$

allowing us to write $W_{HJ}\partial_T W_H$ in the following way:

$$W_{HJ}\partial_T W_H = \begin{pmatrix}\sigma_* & -\rho_* & \nu_* + \left(\beta + \frac{U}{2}\right)\rho_* & 2\nu_* + U\rho_*\end{pmatrix}\begin{pmatrix}J_T \\ \gamma_{*T} \\ U_T \\ \beta_T\end{pmatrix}. \tag{98}$$

Repeating the process of expressing the derivatives in the modulation equations as products of vectors as in equation (98), we can write the modulation equations as a quasi-linear first order system:

$$A\begin{pmatrix}J_T \\ \gamma_{*T} \\ U_T \\ \beta_T\end{pmatrix} + a\begin{pmatrix}J_X \\ \gamma_{*X} \\ U_X \\ \beta_X\end{pmatrix} = 0 \tag{99}$$

with

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix} \tag{100}$$

$$a = \begin{pmatrix} 0 & 1 & 0 & 0 \\ UA_{21} & UA_{22} & UA_{23} - 2\tau & UA_{24} \\ UA_{31} + a_{31} & UA_{32} + a_{32} & UA_{33} + a_{33} & UA_{34} + a_{34} \\ UA_{41} - 2\tau W_H & UA_{42} & UA_{43} + 2\tau M & UA_{44} \end{pmatrix} \tag{101}$$

and the coefficients (calculated in appendix D)

$$\begin{aligned}
A_{21} &= \sigma_* \\
A_{22} &= -\rho_* \\
A_{23} &= \nu_* + \rho_* \left(\beta + \frac{U}{2}\right) \\
A_{24} &= 2\nu_* + U\rho_* \\
A_{31} &= \tau_* + \frac{J\sigma_*}{2} + \left(\beta + \frac{U}{2}\right)\Gamma \\
A_{32} &= -\frac{J}{2}\rho_* - \left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*} \\
A_{33} &= \frac{J}{2}\nu_* + \frac{W_H M_H}{2}\left(\beta + \frac{U}{2}\right) - \frac{M_* W_{HJ}}{2} + \frac{J}{2}\rho_*\left(\beta + \frac{U}{2}\right) + \left(\beta + \frac{U}{2}\right)^2 \{W_H, W_{\gamma_*}\}_{J,\gamma_*} \\
A_{34} &= J\nu_* + W_H M_{*H}\left(\beta + \frac{U}{2}\right) - M_* W_{HJ} + \frac{J}{2}U\rho_* + U\left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*} \\
A_{41} &= -\Gamma W_H - \sigma_* M_* \\
A_{42} &= W_H\{W_H, W_{\gamma_*}\}_{J,\gamma_*} + \rho_* M_* \\
A_{43} &= -\frac{W_H^2 M_{*H}}{2} - W_H\left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*} - M_*(\nu_* + \rho_*\left(\beta + \frac{U}{2}\right)) \\
A_{44} &= -W_H^2 M_{*H} - UW_H\{W_H, W_{\gamma_*}\}_{J,\gamma_*} - M_*(2\nu_* + U\rho_*) \\
a_{31} &= W_H W_{JJ} \\
a_{32} &= W_H M_{*J} - W_{HJ} M_* \\
a_{33} &= \frac{W_H^2}{2} + J\tau_* - W_H M_{*J}\left(\beta + \frac{U}{2}\right) \\
a_{34} &= W_H^2 + UM_* W_{HJ} - UW_H M_{*J}.
\end{aligned}$$

If any of the characteristics of the system in equation (99) are complex, then according to Whitham's theory [Whi1999], the system is modulationally unstable. Our task is to find the characteristics of equation (99) by solving for X', T' in:

$$\det(AX' - aT') = 0, \tag{102}$$

where $X' = \frac{dX}{ds}$, $T' = \frac{dT}{ds}$ along a characteristic curve in the (X, T)-plane parametrized by s. The degenerate case $X' = T' = 0$ in equation (102) corresponds to the non-existence of c_1, c_2 not both zero such that

$$\det(c_1 A + c_2 a) \neq 0. \tag{103}$$

To demonstrate that our system indeed satisfies equation (103), consider:

$$\begin{aligned}
\det A &= -W_{HJ}\tau_* \left(M_*(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}) - \frac{W_H^2}{2}\{W_H, W_{\gamma_*}\}_{H,\gamma_*}\right) \\
&= -\frac{W_{HJ}\tau_*}{2^9\pi^4} D \\
&\neq 0,
\end{aligned}$$

where D is the non-vanishing determinant defined in Theorem 1, which has been calculated in appendix C as equation (124). Hence $c_2 = 0$ and any $c_1 \neq 0$ satisfies equation (103), so we do not need to consider the degenerate case. We now focus on computing:

$$\det(AX' - aT') = \begin{vmatrix} 0 & -T' & 0 & X' \\ A_{21}(X'-UT') & A_{22}(X'-UT') & A_{23}(X'-UT')+2\tau_* T' & A_{24}(X'-UT') \\ A_{31}(X'-UT')-a_{31}T' & A_{32}(X'-UT')-a_{32}T' & A_{33}(X'-UT')-a_{33}T' & A_{34}(X'-UT')-a_{34}T' \\ A_{41}(X'-UT')+2\tau_* W_H T' & A_{42}(X'-UT') & A_{43}(X'-UT')-2\tau_* M T' & A_{44}(X'-UT') \end{vmatrix}. \quad (104)$$

Equation (104) can be written in block form:

$$\det(AX' - aT') = \begin{vmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{vmatrix}, \quad (105)$$

with

$$P_{11} = \begin{pmatrix} 0 & -T' \\ A_{21}(X' - UT') & A_{22}(X' - UT') \end{pmatrix}$$
$$P_{12} = \begin{pmatrix} 0 & X' \\ A_{23}(X' - UT') + 2\tau_* T' & A_{24}(X' - UT') \end{pmatrix}$$
$$P_{21} = \begin{pmatrix} A_{31}(X' - UT') - a_{31}T' & A_{32}(X' - UT') - a_{32}T' \\ A_{41}(X' - UT') + 2\tau_* W_H T' & A_{42}(X' - UT') \end{pmatrix}$$
$$P_{22} = \begin{pmatrix} A_{33}(X' - UT') - a_{33}T' & A_{34}(X' - UT') - a_{34}T' \\ A_{43}(X' - UT') - 2\tau_* M T' & A_{44}(X' - UT') \end{pmatrix}.$$

Lemma 1. *Under the assumptions of Theorem 1, the matrix*

$$P_{11} = \begin{pmatrix} 0 & -T' \\ A_{21}(X' - UT') & A_{22}(X' - UT') \end{pmatrix}$$

is invertible.

Proof. We note that

$$\det P_{11} = A_{21}T'(X' - UT') = \sigma_* T'(X' - UT'). \quad (106)$$

From Theorem 1, $\sigma_* \neq 0$, so P_{11} is invertible if and only if $T' \neq 0$ and $X' \neq UT'$. If $T' = 0$, then

$$\begin{aligned} \det(AX' - aT') &= X'^4 \det A \\ &= \frac{-W_{HJ}\tau_*}{2^9\pi^4} X'^4 D. \end{aligned}$$

By assumption, $W_{HJ} \neq 0$, $W_H \neq 0$ (this corresponds to the underlying periodic wave having non-zero period) and hence $\tau_* = \frac{1}{2} W_H W_{HJ} \neq 0$. Moreover, $D \neq 0$ from Theorem 1, hence for X', T' to solve equation (102) with $T' = 0$ we have:

$$\det(AX' - aT') = 0 \implies T' = X' = 0,$$

which we exclude as a trivial solution. In fact, if $T'(s_*) = 0$ for some value of $s = s_*$, we have that $X'(s_*) = 0$ as well, meaning that the characteristic curve $(X(s), T(s))$ terminates at $s = s_*$. This is not possible, since the characteristic curves are defined for all $X, T \in \mathbb{R}$, so in fact T' can never vanish. Next, if $X' = UT'$, then:

$$\begin{aligned} \det(AX' - aT') &= \begin{vmatrix} 0 & -T' & 0 & UT' \\ 0 & 0 & 2\tau_* T' & 0 \\ -a_{31}T' & -a_{32}T' & -a_{33}T' & -a_{34}T' \\ 2\tau_* W_H T' & 0 & -2\tau_* M T' & 0 \end{vmatrix} \\ &= -2\tau_* T' \begin{vmatrix} 0 & -T' & UT' \\ -a_{31}T' & -a_{32}T' & -a_{34}T' \\ 2\tau_* W_H T' & 0 & 0 \end{vmatrix} \\ &= -4\tau_*^2 W_H T'^2 (a_{34}T'^2 + a_{32}UT'^2) \\ &= -4\tau_*^2 W_H T'^4 \left(W_H^2 + UM_* W_{HJ} - UW_H M_{*J} + U(W_H M_{*J} - W_{HJ} M_*)\right) \\ &= -4\tau_*^2 W_H^3 T'^4. \end{aligned}$$

Only T' is able to vanish, hence if $X' = UT'$ then

$$\det(AX' - aT') = 0 \implies T' = X' = 0.$$

There is no non-trivial solution (X', T') of equation (102) for which P_{11} is singular, which proves the lemma. □

Lemma 1 allows us to use the Schur determinant formula (see, for example, [Zhang2006]) in equation (105), in particular:

$$\det(AX' - aT') = \det(P_{11})\det(P_{22} - P_{21}P_{11}^{-1}P_{12}). \tag{107}$$

Direct calculation of $P_{22} - P_{21}P_{11}^{-1}P_{12}$ yields:

$$P_{22} - P_{21}P_{11}^{-1}P_{12} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix},$$

where

$$m_{11} = A_{33}(X' - UT') - a_{33}T' - \frac{1}{\sigma_*(X' - UT')}(A_{23}(X' - UT') + 2\tau_* T')(A_{31}(X' - UT') - a_{31}T') \tag{108}$$

$$m_{12} = A_{34}(X' - UT') - a_{34}T' + \frac{X'}{T'}(A_{32}(X' - UT') - a_{32}T') - \frac{A_{24}}{\sigma_*}(A_{31}(X' - UT') - a_{31}T') \tag{109}$$

$$- \frac{A_{22}X'}{\sigma_* T'}(A_{31}(X' - UT') - a_{31}T') \tag{110}$$

$$m_{21} = A_{43}(X' - UT') - 2\tau_* MT' - \frac{1}{\sigma_*(X' - UT')}(A_{23}(X' - UT') + 2\tau_* T')(A_{41}(X' - UT') + 2\tau_* W_H T') \tag{111}$$

$$m_{22} = \frac{1}{T'}(X' - UT')(A_{42}X' + A_{44}T') - \frac{A_{24}}{\sigma_*}(A_{41}(X' - UT') + 2\tau_* W_H T') - \frac{A_{22}X'}{\sigma_* T'}(A_{41}(X' - UT') + 2\tau_* W_H T'). \tag{112}$$

We make the substitution $\lambda = X' - UT'$ and $\mu = iT'$ (justified by the end result) and note from equation (107) that:

$$\begin{aligned} \det(AX' - aT') &= \sigma_*(-i\lambda\mu)\begin{vmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{vmatrix} \\ &= \frac{1}{\sigma_* W_H}\begin{vmatrix} \sigma_* \lambda W_H m_{11} & -i\sigma_* \mu W_H m_{12} \\ \sigma_* \lambda m_{21} & -i\sigma_* \mu m_{22} \end{vmatrix}, \end{aligned}$$

which can be simplified to:

$$\det(AX' - aT') = \frac{1}{\sigma_* W_H}\begin{vmatrix} a'_{11}\lambda^2 + b'_{11}\lambda\mu + c'_{11}\mu^2 & a'_{12}\lambda^2 + b'_{12}\lambda\mu + c'_{12}\mu^2 \\ a'_{21}\lambda^2 + b'_{21}\lambda\mu + c'_{21}\mu^2 & a'_{22}\lambda^2 + b'_{22}\lambda\mu + c'_{22}\mu^2 \end{vmatrix}. \tag{113}$$

We calculate the coefficients in appendix E. To list them, we have:

$$\begin{aligned}
a'_{11} &= 2\tau_*\left(\beta + \frac{U}{2}\right)^2(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}) - 2\tau_*\rho_* W_H\left(\beta + \frac{U}{2}\right) - \sigma_*\tau_* M_* - \nu_*\tau_* W_H \\
b'_{11} &= 2i\tau_* W_H(2\tau_* + 2\Gamma\left(\beta + \frac{U}{2}\right) + J\sigma_*) \\
c'_{11} &= -2\tau_* W_H^2 W_{JJ} \\
a'_{12} &= \rho_*\tau_* W_H - 2\tau_*\left(\beta + \frac{U}{2}\right)(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}) \\
b'_{12} &= 2i\tau_* W_H(\nu_* - \Gamma + \rho_*\left(\beta + \frac{U}{2}\right)) \\
c'_{12} &= 4\tau_*^2 W_H \\
a'_{21} &= \rho_*\tau_* W_H - 2\tau_*\left(\beta + \frac{U}{2}\right)(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}) = a'_{12} \\
b'_{21} &= 2i\tau_* W_H(\nu_* - \Gamma + \rho_*\left(\beta + \frac{U}{2}\right)) = b'_{12} \\
c'_{21} &= 4\tau_*^2 W_H = c'_{12} \\
a'_{22} &= 2\tau_*(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}) \\
b'_{22} &= -4i\tau_*\rho_* W_H \\
c'_{22} &= 4\tau_*\nu_* W_H.
\end{aligned}$$

We note that the matrix in equation (113) is symmetric, which simplifies our calculations. Now we perform row and column operations to equation (113), which will leave the determinant unchanged. Adding $(\beta + \frac{U}{2})$ times the second row to the first row, and then $(\beta + \frac{U}{2})$ times the second column to the first column, we have the following result:

$$\det(AX' - aT') = \frac{1}{\sigma_* W_H} \begin{vmatrix} Q_{11}(\lambda,\mu) & Q_{12}(\lambda,\mu) \\ Q_{21}(\lambda,\mu) & Q_{22}(\lambda,\mu) \end{vmatrix}. \tag{114}$$

The polynomials are calculated in appendix E. We list them as:

$$\begin{aligned} Q_{11}(\lambda,\mu) &= \frac{\tau_*}{2^{15}\pi^5\sigma_*}(d_2\lambda^2 + d_1\lambda\mu + d_0\mu^2) \\ Q_{12}(\lambda,\mu) &= \frac{\tau_*}{2^{15}\pi^5\sigma_*}(b_2\lambda^2 + b_1\lambda\mu + b_0\mu^2) \\ Q_{21}(\lambda,\mu) &= Q_{12}(\lambda,\mu) \\ Q_{22}(\lambda,\mu) &= \frac{\tau_*}{2^{15}\pi^5\sigma_*}(a_2\lambda^2 + a_1\lambda\mu + a_0\mu^2) \end{aligned}$$

We can swap the two columns and then the two rows without changing the determinant:

$$\det(AX' - aT') = \frac{1}{\sigma_* W_H} \begin{vmatrix} Q_{22}(\lambda,\mu) & Q_{21}(\lambda,\mu) \\ Q_{12}(\lambda,\mu) & Q_{11}(\lambda,\mu) \end{vmatrix} \tag{115}$$

$$= \frac{\tau_*^2}{2^{30}\pi^{10}\sigma_*^3 W_H} \begin{vmatrix} a_2\lambda^2 + a_1\lambda\mu + a_0\mu^2 & b_2\lambda^2 + b_1\lambda\mu + b_0\mu^2 \\ b_2\lambda^2 + b_1\lambda\mu + b_0\mu^2 & d_2\lambda^2 + d_1\lambda\mu + d_0\mu^2 \end{vmatrix} \tag{116}$$

$$= \frac{W_H W_{HJ}^2}{2^{32}\pi^{10}\sigma_*^3} \det\left(\lambda^2 \begin{pmatrix} a_2 & b_2 \\ b_2 & d_2 \end{pmatrix} + \lambda\mu \begin{pmatrix} a_1 & b_1 \\ b_1 & d_1 \end{pmatrix} + \mu^2 \begin{pmatrix} a_0 & b_0 \\ b_0 & d_0 \end{pmatrix}\right). \tag{117}$$

This concludes the calculation of $\det(AX' - aT')$ when $W_{HJ} \neq 0$. For completeness, we must carry out the same calculation starting with the assumption that $W_{JJ} \neq 0$ and $W_{HH} \neq 0$, which is the other case arising from the genericity condition $\sigma_* \neq 0$. This second calculation is unremarkable, so we give the final result here, but provide some working in appendix F. In this case, we have, almost exactly as before:

$$\det(AX' - aT') = -\frac{W_H W_{JJ}^2}{2^{32}\pi^{10}\sigma_*^3} \det\left(\lambda^2 \begin{pmatrix} a_2 & b_2 \\ b_2 & d_2 \end{pmatrix} + \lambda\mu \begin{pmatrix} a_1 & b_1 \\ b_1 & d_1 \end{pmatrix} + \mu^2 \begin{pmatrix} a_0 & b_0 \\ b_0 & d_0 \end{pmatrix}\right), \tag{118}$$

with the difference being the non-zero factor W_{JJ}^2. This leads us to the following theorem.

Theorem 3. *Suppose that the assumptions of Theorem 1 hold, that is, $\sigma \neq 0$ and $D \neq 0$. Further, assume that $T_* = 4\pi W_H \neq 0$. Then, the equation for the characteristics of the Whitham modulation equations associated with the nonlinear Schrödinger equation* (1) *is equivalent to the normal form for the continuous bands of spectrum emerging from the origin in the spectral plane, given in Theorem 2.*

Proof. To solve for the characteristics of the Whitham system, we substitute into equation (102) the appropriate expression for $\det(AX' - aT')$ given in either equation (117) or equation (118). In both cases, we can divide $\det(AX' - aT')$ by the constants we have assumed to be non-zero, so X' and T' solve the quartic

$$\det\left(\lambda^2 \begin{pmatrix} a_2 & b_2 \\ b_2 & d_2 \end{pmatrix} + \lambda\mu \begin{pmatrix} a_1 & b_1 \\ b_1 & d_1 \end{pmatrix} + \mu^2 \begin{pmatrix} a_0 & b_0 \\ b_0 & d_0 \end{pmatrix}\right) = 0,$$

with

$$\lambda = X' - UT', \quad \mu = iT'. \tag{119}$$

This is exactly the normal form given in Theorem 2. □

Remark 4. The substitution $\mu = iT'$ implies that $T' \in i\mathbb{R}$, however this is not required for (X', T') to be a solution to equation (102). Rather, we can safely make the restriction that $T' \in i\mathbb{R}$ by multiplying equation (102) by a phase function $e^{4if(s)}$ so that T' becomes purely imaginary and $X' \mapsto X'e^{if(s)}$ is possibly complex.

Corollary 2. *Whitham modulation theory predicts the same criterion for modulational instability as the linear theory given in Corollary 1, that is, the existence of a complex characteristic of the system of modulation equations (equation* (99)*) corresponds to a root λ of the normal form equation* (21) *with $\mathrm{Re}(\lambda) \neq 0$, indicating modulational instability of the underlying periodic, travelling wave.*

Proof. Noting that $\det A \neq 0$ and $T'(s) \neq 0$, we may rewrite equation (102) as:

$$\det\left(\frac{X'}{T'} - A^{-1}a\right) = 0. \tag{120}$$

Since $T'(s) \neq 0$ for any s, the characteristic curves may instead be parametrized by T, so that $\frac{dX}{dT} = \frac{X'(s)}{T'(s)}$. Hence the eigenvalues $\frac{X'(s)}{T'(s)}$ of $A^{-1}a$ define the characteristic curves, which is clear once equation (99) is instead written as:

$$\begin{pmatrix} J_T \\ \gamma_{*T} \\ U_T \\ \beta_T \end{pmatrix} + A^{-1}a \begin{pmatrix} J_X \\ \gamma_{*X} \\ U_X \\ \beta_X \end{pmatrix} = 0,$$

and similarly for equation (125). According to the Whitham theory [Whi1999], modulational instability occurs when one of the characteristics of a Whitham system is complex. From equation (120), this is equivalent to $\frac{X'}{T'} \in \mathbb{C}$. We have

$$\begin{aligned} \frac{X'}{T'} &= \frac{\lambda - iU\mu}{-i\mu} \\ &= \frac{i\lambda}{\mu} + U. \end{aligned}$$

We note that $\mu \in \mathbb{R}$ since it is the Floquet exponent of the periodic solutions to equation (20), and from remark 4 we know that this does not restrict the class of characteristic curves. Hence we have that

$$\frac{X'}{T'} \in \mathbb{C} \iff \mathrm{Re}(\lambda) \neq 0,$$

which agrees with Corollary 1. □

Remark 5. The transformation from the characteristic variables of the Whitham system (X', T') to the spectral variable λ and Floquet exponent μ may have greater significance in the scope of rigorously proving the agreement of Whitham modulation theory with linear stability at the origin. In fact, the transformation in all previous examples [Serre2005, OZ2006, JZ2010, JMMP2014] is:

$$\lambda = X' - cT', \quad \mu = -\frac{iT'}{T_*},$$

where T_* is the period of the underlying wave. To explain the factor of $\frac{1}{T_*}$, the cited papers all consider the Floquet multiplier to have the form $e^{i\mu}$, whereas we have chosen to consider $e^{i\mu T_*}$ in keeping with [LBJM2019].

4 Discussion and open problems

In this paper, we show that the formal Whitham modulation theory correctly predicts the modulational instability of periodic, travelling wave solutions of the general nonlinear Schrödinger equation (1) as prescribed by the rigorous spectral analysis at the origin in [LBJM2019]. Applying the variational principle to the averaged Lagrangian allows us to derive four modulation equations, which we then homogenize using the genericity conditions described in [LBJM2019]. This results in two cases depending on which slowly-varying parameters we eliminate, however the calculations are essentially the same. Finally, we compute a quartic equation for the characteristics of the homogenized modulation equations from the determinant of the quasi-linear system equation (99). By invoking various determinant identities inspired by the proof of [LBJM2019, Proposition 1] and also a change of variables, we deduce that the characteristics of the Whitham system satisfy the same quartic equation as the normal form for the four continuous bands of spectrum at the origin [LBJM2019].

Leisman et al. also provide a modulational instability criterion for transverse perturbations, where they consider the stability of periodic solutions of equation (1) to the two-dimensional equation:

$$i\psi_t = \psi_{xx} \pm \psi_{zz} + \zeta f(|\psi|^2)\psi.$$

The extension is neat from the perspective of the linear theory, however we believe extending the Whitham theory would involve a two-phase approach, making the homogenization process considerably more difficult. We have decided that this lies outside the scope of this paper.

One interesting future direction for our results would be to compute Riemann invariants and investigate the behaviour of dispersive shock waves for suitable one-dimensional potentials $f(|\psi|^2)$. This would involve diagonalizing the matrix $A^{-1}a$ from the quasi-linear system equation (99).

Another obvious extension would be to consider a coupled general nonlinear Schrödinger system:

$$
\begin{aligned}
i\psi_{1t} &= \psi_{1xx} + \zeta_1 f(\zeta_1|\psi_1|^2, \zeta_2|\psi_2|^2)\psi_1 \\
i\psi_{2t} &= \psi_{2xx} + \zeta_2 f(\zeta_1|\psi_1|^2, \zeta_2|\psi_2|^2)\psi_2,
\end{aligned}
$$

which has the Lagrangian:

$$
L = i(\overline{\psi_1}\psi_{1t} - \psi_1\overline{\psi}_{1t}) + i(\overline{\psi_2}\psi_{2t} - \psi_2\overline{\psi}_{2t}) + 2|\psi_{1x}|^2 + 2|\psi_{2x}|^2 - 2F(\zeta_1|\psi_1|^2, \zeta_2|\psi_2|^2).
$$

More generally, we could couple n general nonlinear Schrödinger equations together, which yields

$$
i\partial_t \boldsymbol{\psi} = \partial_x^2 \boldsymbol{\psi} + \nabla_{\boldsymbol{\psi}} F(|\psi_1|^2, \ldots, |\psi_n|^2) \tag{121}
$$

$$
L = i(\overline{\boldsymbol{\psi}} \cdot \partial_t \boldsymbol{\psi} - \boldsymbol{\psi} \cdot \partial_t \overline{\boldsymbol{\psi}}) + 2|\partial_x \boldsymbol{\psi}|^2 - 2F(|\psi_1|^2, \ldots, |\psi_n|^2), \tag{122}
$$

where F is a scalar potential. We believe that the Whitham theory for these systems would closely resemble the procedure for equation (1), however difficulty may arise in the homogenization process. In particular, one must find suitable genericity conditions to eliminate enough slowly-varying parameters.

Finally, we cite the open problem of proving that the modulational instability criterion derived from the hyperbolicity of a Whitham modulation system coincides with spectral instability of a periodic travelling wave solution of some general class of PDEs. The difficulty in developing a general proof lies in the homogenization of the Whitham equations. This process is guided by genericity conditions which are determined from the underlying PDE in the course of the rigorous linearized spectral analysis; this explains the disparity amongst the analyses of [Serre2005, OZ2006, JZ2010, JMMP2014] and this paper. On the other hand, it is promising that the transformation between the characteristics of the Whitham system and the spectral variables (equation (119)) appears in the cited examples where the Whitham theory has been rigorously verified, perhaps offering another avenue for research.

5 Acknowledgements

Both authors acknowledge the support of the Australian Research Council under grant DP200102130.

Appendix A Poisson brackets from the linear theory

We first define the Poisson brackets:

$$
\begin{aligned}
&\gamma = \{T_*, \eta\}_{\kappa,\omega} \quad \rho = \{T_*, \eta\}_{\omega,E} \quad \tau = \frac{T_* T_{*\kappa}}{2} \\
&\nu = -\frac{T_* T_{*E}}{2} \quad \xi = \{T_*, \eta\}_{\kappa,\zeta} \quad \chi = \{T_*, \eta\}_{\zeta,E}.
\end{aligned}
$$

With these definitions, we can list the matrix entries in equation (21):

$$
\begin{aligned}
a_2 &= -\frac{\sigma}{2}(\gamma M_E + \rho M_\kappa + \sigma M_\omega) \\
b_2 &= -\frac{\sigma\rho T_*}{2} = -\sigma(\tau M_E + \nu M_\kappa) \\
d_2 &= -\frac{\sigma}{2}(\nu T_* + \frac{\sigma}{2} M) \\
a_1 &= 2i\sigma\rho T_* \\
b_1 &= i\sigma T_*(\nu + \gamma) \\
d_1 &= i\sigma T_*(2\tau + \sigma\kappa) \\
a_0 &= 2\sigma\nu T_* \\
b_0 &= 2\sigma\tau T_* \\
d_0 &= 2\sigma T_*(\omega\gamma - \zeta\xi - E\sigma)
\end{aligned}
$$

Appendix B Relations between K and W

In this appendix, we give the relations between the derivatives of K and the derivatives of W, defined in equations (10) and (64) respectively. This is carried out by taking a derivative of equation (75) with respect to one of the parameters $H, U, J, \gamma, \beta, \zeta$ and then applying the chain rule, simplifying the results using equations (71) to (73). Finally, we use equations (11) to (13) to eliminate K. For the period $T = K_E$ in the linear theory:

$$\begin{aligned}
T_* &= 4\pi W_H \\
T_{*E} &= 16\pi W_{HH} \\
T_{*\kappa} &= 16\pi(W_{HJ} - (\beta + \frac{U}{2})W_{HH}) = -16\pi\eta_{*H} \\
T_{*\omega} &= -4\pi M_H.
\end{aligned}$$

Next, for η we have:

$$\begin{aligned}
\eta &= 4\pi\eta_* \\
\eta_E &= 16\pi\eta_{*H} \\
\eta_\kappa &= 16\pi(\eta_{*J} - (\beta + \frac{U}{2})\eta_{*H}) \\
\eta_\omega &= -4\pi\eta_{*\gamma_*}.
\end{aligned}$$

Finally for M:

$$\begin{aligned}
M &= 2\pi M_* \\
M_E &= 8\pi M_{*H} \\
M_\kappa &= 8\pi(M_{*J} - (\beta + \frac{U}{2})M_{*H}) \\
M_\omega &= -2\pi M_{*\gamma_*}.
\end{aligned}$$

Since ζ is the same parameter for both cases, we have that:

$$K_\zeta = \pi W_\zeta.$$

Appendix C Poisson brackets for the modulation equations

Similarly to the Poisson brackets in appendix A, we make the following definitions for Poisson brackets for the Whitham system:

$$\begin{aligned}
\sigma_* &= \{W_H, \eta_*\}_{H,J} = \{W_H, W_J\}_{J,H} \\
\rho_* &= \{W_H, \eta_*\}_{\gamma_*,H} = \{W_H, W_J\}_{H,\gamma_*} \\
\Gamma &= \{W_H, \eta_*\}_{J,\gamma_*} = \{W_H, W_J\}_{\gamma_*,J} \\
\nu_* &= -\frac{1}{2}W_H W_{HH} \\
\tau_* &= \frac{1}{2}W_H W_{HJ} \\
\xi_* &= \{W_H, \eta_*\}_{J,\zeta} = \{W_H, W_J\}_{\zeta,J} \\
\chi_* &= \{W_H, \eta_*\}_{\zeta,H} = \{W_H, W_J\}_{H,\zeta}.
\end{aligned}$$

The reason we have included the Poisson brackets in terms of W_J as well as η_* is because it is more straightforward to use the symmetry of mixed partial derivatives of W when using W_J.

We make extensive use of several properties of Poisson brackets. Let P, Q, S be functions of w, x, y, z with symmetric mixed partial derivatives. Then:

- $\{P, Q\}_{w,x} = -\{P, Q\}_{x,w} = -\{Q, P\}_{w,x}$;
- $\{P_w, Q_x\}_{y,z} = \{P_y, Q_z\}_{w,x}$;
- $\{P, P\}_{w,x} = 0$;
- $\{PQ, S\}_{w,x} = P\{Q, S\}_{w,x} + Q\{P, S\}_{w,x}$;

- $\{\alpha P, Q\}_{w,x} = \alpha\{P, Q\}_{w,x}$ for a constant α.

The above definitions allow us to express the equivalent Poisson brackets in appendix A in terms of these newly-defined Poisson brackets for the Whitham parameters:

$$\begin{aligned}
&\sigma = 256\pi^2\sigma_*\\
&\rho = -64\pi^2\rho_*\\
&\gamma = -64\pi^2(\Gamma + \left(\beta + \frac{U}{2}\right)\rho_*)\\
&\nu = 64\pi^2\nu_*\\
&\tau = 64\pi^2(\tau_* + \left(\beta + \frac{U}{2}\right)\nu_*)\\
&\chi = 64\pi^2\chi_*\\
&\xi = 64\pi^2(\xi_* + \left(\beta + \frac{U}{2}\right)\chi_*)
\end{aligned}$$

Ultimately, we use these expressions to express the matrix entries of equation (21) (given in appendix A) in terms of the Whitham parameters:

$$\begin{aligned}
a_2 &= 2^{16}\pi^5\sigma_*(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*})\\
b_2 &= 2^{15}\pi^5\rho_*\sigma_* W_H = -2^{16}\sigma_*(\tau_* M_{*H} + \nu_* M_{*J})\\
d_2 &= -2^{15}\pi^5\sigma_*(\nu_* W_H + \sigma_* M_*)\\
a_1 &= -2^{17}i\pi^5\sigma_*\rho_* W_H\\
b_1 &= 2^{16}i\pi^5\sigma_* W_H(\nu_* - \Gamma - \left(\beta + \frac{U}{2}\right)\rho_*)\\
d_1 &= 2^{16}i\pi^5\sigma_* W_H(2\tau_* + 2\left(\beta + \frac{U}{2}\right)\nu_* + J\sigma_*)\\
a_0 &= 2^{17}\pi^5\sigma_*\nu_* W_H\\
b_0 &= 2^{17}\pi^5\sigma_* W_H(\tau_* + \left(\beta + \frac{U}{2}\right)\nu_*)
\end{aligned}$$

For the last entry d_0, we use the identity:

$$W = 2W_H\left(H + \beta J + \frac{UJ}{2}\right) - 2M_*\left(\beta U + \frac{U^2}{4} - \gamma_*\right) - J\eta_* - 2\zeta W_\zeta. \tag{123}$$

In particular, taking Poisson brackets of equation (123) with W_J and derivatives with respect to H and J, and using the fact that $W_J = 0$ in equation (88) yields:

$$W_H\eta_{*J} = -\sigma_*\left(2H + \beta J + \frac{UJ}{2}\right) - 2\Gamma\left(\beta U + \frac{U^2}{4} - \gamma_*\right) - 2\zeta\xi_*.$$

Similarly, we have for a Poisson bracket of equation (123) with W_H and derivatives H and J:

$$W_H\eta_{*H} = 2\rho_*\left(\beta U + \frac{U^2}{4} - \gamma_*\right) + J\sigma_* + 2\zeta\chi_*.$$

Hence

$$\begin{aligned}
d_0 &= 2\sigma T(\omega\gamma - \zeta\xi - E\sigma)\\
&= 2^{17}\pi^5\sigma_* W_H(-\left(\beta U + \frac{U^2}{4} - \gamma_*\right)\Gamma - \zeta\xi_* - \left(H + \beta J + \frac{UJ}{2}\right)\sigma_* - \left(\beta + \frac{U}{2}\right)(\rho_*\left(\beta + \frac{U^2}{4} - \gamma_*\right) + \zeta\chi_*))\\
&= 2^{16}\pi^5\sigma_* W_H^2(\eta_{*J} - \eta_{*H}\left(\beta + \frac{U}{2}\right)).
\end{aligned}$$

Now writing η_{*H}, η_{*J} as the derivatives of W_J using its definition in equation (77), we see that

$$
\begin{aligned}
d_0 &= 2^{16}\pi^5\sigma_* W_H^2 (2W_{HJ}\left(\beta + \frac{U}{2}\right) - W_{JJ} - \left(\beta + \frac{U}{2}\right)^2 W_{HH}) \\
&= 2^{16}\pi^5\sigma_* W_H (4\tau_*\left(\beta + \frac{U}{2}\right) + 2\nu_*\left(\beta + \frac{U}{2}\right)^2 - W_H W_{JJ}).
\end{aligned}
$$

We also provide a calculation of the determinant D from theorem 1 in terms of the Whitham parameters:

$$
\begin{aligned}
D &= -\frac{4}{\sigma^3}(a_2 d_2 - b_2^2) \\
&= -\frac{4}{2^{24}\pi^6\sigma_*^3}\left((2^{16}\pi^5\sigma_*(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma}))(-2^{15}\pi^5\sigma_*(\nu_* W_H + \sigma_* M_*)) - (2^{16}\pi^5\sigma_*(\tau_* M_{*H} + \nu_* M_{*J}))^2\right) \\
&= -\frac{2^9\pi^4}{\sigma_*}\left(-(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*})(\nu_* W_H + \sigma_* M_*) - 2(\tau_* M_{*H} + \nu_* M_{*J})^2\right) \\
&= \frac{2^9\pi^4}{\sigma_*}\left(\sigma_* M_* \left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}\right) + \nu_* W_H \left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}\right) + \frac{W_H^2}{2}\left(W_{HJ} M_{*H} - W_{HH} M_{*J}\right)^2\right) \\
&= 2^9\pi^4\Bigg(M_* \left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}\right) \\
&+ \frac{W_H^2}{2\sigma_*}\left(-W_{HH}\left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}\right) + W_{HJ}^2 M_{*H}^2 - 2W_{HJ}W_{HH}M_{*H}M_{*J} + W_{HH}^2 M_{*J}^2\right)\Bigg) \\
&= 2^9\pi^4\Bigg(M_* \left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma}\right) - \frac{W_H^2}{2\sigma_*}\Big(\sigma_* W_{HH} M_{*\gamma} + W_{HH} M_{*H}(M_{*H} W_{JJ} - W_{HJ} M_{*J}) \\
&+ W_{HH} M_{*J}(W_{HH} M_{*J} - W_{HJ} M_{*H}) - W_{HJ}^2 M_{*H}^2 + 2W_{HJ}W_{HH}M_{*H}M_{*J} - W_{HH}^2 M_{*J}^2\Big)\Bigg) \\
&= 2^9\pi^4\left(M_* \left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}\right) - \frac{W_H^2}{2\sigma_*}\left(\sigma_* W_{HH} M_{*\gamma_*} - M_{*H}^2 (W_{HJ}^2 - W_{HH} W_{JJ})\right)\right) \\
&= 2^9\pi^4\left(M_* \left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}\right) - \frac{W_H^2}{2\sigma_*}\left(\sigma_* W_{HH} M_{*\gamma_*} - \sigma_* M_{*H}^2\right)\right)
\end{aligned}
$$

from which we have:

$$
D = 2^9\pi^4\left(M_* \left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}\right) - \frac{W_H^2}{2}\{W_H, W_{\gamma_*}\}_{H,\gamma_*}\right). \tag{124}
$$

Appendix D Calculating matrix elements

In this appendix, we provide some calculations for the matrix elements of equations (100) and (101). We already calculated $A_{21}, A_{22}, A_{23}, A_{24}$ in equation (98). For equation (91), we compute:

$$\begin{aligned}
W_{HJ}\partial_T W_U &= \{W_U, W_J\}_{J,H} J_T + \{W_U, W_J\}_{\gamma_*,H}\gamma_{*T} + \{W_U, W_J\}_{U,H} U_T + \{W_U, W_J\}_{\beta,H}\beta_T \\
&= \left\{\frac{JW_H}{2} - W_{\gamma_*}\left(\beta + \frac{U}{2}\right), W_J\right\}_{J,H} J_T + \left\{\frac{JW_H}{2} - W_{\gamma_*}\left(\beta + \frac{U}{2}\right), W_J\right\}_{\gamma_*,H}\gamma_{*T} \\
&+ \left\{\frac{JW_H}{2} - W_{\gamma_*}\left(\beta + \frac{U}{2}\right), W_J\right\}_{U,H} U_T + \left\{\frac{JW_H}{2} - W_{\gamma_*}\left(\beta + \frac{U}{2}\right), W_J\right\}_{\beta,H}\beta_T \\
&= \left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right) J_T + \left(-\frac{J}{2}\rho_* - \left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\right)\gamma_{*T} \\
&+ \left(\frac{J}{2}\left\{\frac{JW_H}{2} - W_{\gamma_*}\left(\beta + \frac{U}{2}\right), W_H\right\}_{H,J} - \frac{1}{2}M_* W_{HJ} - \left(\beta + \frac{U}{2}\right)\left\{\frac{JW_H}{2} - W_{\gamma_*}\left(\beta + \frac{U}{2}\right), W_H\right\}_{\gamma_*,J}\right) U_T \\
&+ \left(\frac{J}{2}\left\{JW_H - UW_{\gamma_*}, W_H\right\}_{H,J} - M_* W_{HJ} - \left(\beta + \frac{U}{2}\right)\left\{JW_H - UW_{\gamma_*}, W_H\right\}_{\gamma_*,J}\right)\beta_T \\
&= \left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right) J_T + \left(-\frac{J}{2}\rho_* - \left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\right)\gamma_{*T} \\
&+ \left(\frac{J}{2}\nu_* + \frac{J}{2}\rho_*\left(\beta + \frac{U}{2}\right) - \frac{1}{2}M_* W_{HJ} + \frac{1}{2}W_H M_{*H}\left(\beta + \frac{U}{2}\right) + \left(\beta + \frac{U}{2}\right)^2\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\right) U_T \\
&+ \left(J\nu_* + \frac{J}{2}U\rho_* - M_* W_{HJ} + W_H M_H\left(\beta + \frac{U}{2}\right) + U\left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\right)\beta_T \\
&= A_{31} J_T + A_{32}\gamma_{*T} + A_{33} U_T + A_{34}\beta_T.
\end{aligned}$$

For $a_{31}, a_{32}, a_{33}, a_{34}$:

$$\begin{aligned}
W_{HJ}W_U U_X - W_{HJ}\partial_X W &= W_{HJ}W_U U_X - \{W, W_J\}_{J,H} J_X - \{W, W_J\}_{\gamma_*,H}\gamma_{*X} - \{W, W_J\}_{U,H} U_X - \{W, W_J\}_{\beta,H}\beta_X \\
&= W_H W_{JJ} J_X - (M_* W_{HJ} - W_H M_{*J})\gamma_{*X} + W_H W_{UJ} U_X - (W_\beta W_{HJ} - W_H W_{\beta J})\beta_X.
\end{aligned}$$

Taking the J-derivatives of W_U and W_β defined in equations (80) and (81) respectively, we have:

$$\begin{aligned}
W_{HJ}W_U U_X - W_{HJ}\partial_X W &= W_H W_{JJ} J_X - (M_* W_{HJ} - W_H M_{*J})\gamma_{*X} + W_H\left(\frac{1}{2}W_H + \frac{J}{2}W_{HJ} - M_{*J}\left(\beta + \frac{U}{2}\right)\right) U_X \\
&- \left(W_{HJ}(JW_H - UM_*) - W_H(W_H + JW_{HJ} - UM_{*J})\right)\beta_X \\
&= W_H W_{JJ} J_X - (M_* W_{HJ} - W_H M_{*J})\gamma_{*X} + \left(\frac{1}{2}W_H^2 + J\tau_* - W_H M_{*J}\left(\beta + \frac{U}{2}\right)\right) U_X \\
&- \left(-W_H^2 + UW_H M_{*J} - UM_* W_{HJ}\right)\beta_X \\
&= a_{31} J_X + a_{32}\gamma_{*X} + a_{33} U_X + a_{34}\beta_X.
\end{aligned}$$

Finally, in equation (92) we compute:

$$\begin{aligned}
W_{HJ}W_H\partial_T M_* &= W_H\{W_H, W_J\}_{J,\gamma_*} J_T + W_H\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\gamma_{*T} + W_H\{W_H, W_U\}_{J,\gamma_*} U_T + W_H\{W_H, W_\beta\}_{J,\gamma_*}\beta_T \\
&= -W_H\Gamma J_T + W_H\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\gamma_{*T} + W_H\left\{W_H, \frac{JW_H}{2} - W_{\gamma_*}\left(\beta + \frac{U}{2}\right)\right\}_{J,\gamma_*} U_T \\
&+ W_H\left\{W_H, JW_H - UW_{\gamma_*}\right\}_{J,\gamma_*}\beta_T \\
&= -W_H\Gamma J_T + W_H\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\gamma_{*T} + W_H\left(-\frac{1}{2}W_H M_{*H} - \left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\right) U_T \\
&+ W_H\left(-W_H M_{*H} - U\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\right)\beta_T
\end{aligned}$$

Hence for $A_{41}, A_{42}, A_{43}, A_{44}$, we have:

$$\begin{aligned}
W_{HJ}W_H\partial_T M_* - W_{HJ}M_*\partial_T W_H &= \left(-\Gamma W_H - \sigma_* M_*\right) J_T + \left(W_H\{W_H, W_{\gamma_*}\}_{J,\gamma_*} + \rho_* M_*\right) \\
&\quad + \left(-\frac{1}{2}W_H^2 M_{*H} - W_H\left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*} - M_*\left(\nu_* + \rho_*\left(\beta + \frac{U}{2}\right)\right)\right) \\
&\quad + \left(-W_H^2 M_{*H} - UW_H\{W_H, W_{\gamma_*}\}_{J,\gamma_*} - M_*(2\nu_* + U\rho_*)\right) \\
&= A_{41}J_T + A_{42}\gamma_{*T} + A_{43}U_T + A_{44}\beta_T.
\end{aligned}$$

The X-derivatives of equation (92) are:

$$\begin{aligned}
&U(W_{HJ}W_H\partial_X M_* - W_{HJ}M_*\partial_X W_H) + W_{HJ}W_H M_* U_X - W_{HJ}W_H^2 J_X \\
&= (UA_{41} - 2\tau_* W_H)J_X + UA_{42}\gamma_{*X} + (UA_{43} + 2\tau M_*)U_X + UA_{44}\beta_X
\end{aligned}$$

Appendix E Matrix elements for the Schur determinant calculation

In this appendix, we calculate the coefficients in equation (113), using the expressions for $m_{11}, m_{12}, m_{21}, m_{22}$ computed in equations (108) and (110) to (112). We derive a number of useful identities that we use freely in the subsequent calculations:

$$\begin{aligned}
\sigma_*\{W_H, W_{\gamma_*}\}_{J,\gamma_*} - \rho_*\Gamma &= \sigma_*(W_{HJ}M_{*\gamma_*} - M_{*H}M_{*J}) - (M_{*J}W_{HH} - M_{*H}W_{HJ})(M_{*H}W_{JJ} - M_{*J}W_{HJ}) \\
&= W_{HJ}\sigma_* M_{*\gamma_*} - M_{*H}M_{*J}(W_{HJ}^2 - W_{HH}W_{JJ}) \\
&\quad - \left(M_{*J}M_{*H}W_{HH}W_{JJ} - M_{*H}^2 W_{HJ}W_{JJ} - M_{*J}^2 W_{HH}W_{HJ} + M_{*H}M_{*J}W_{HJ}^2\right) \\
&= W_{HJ}\left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}\right) \\
\frac{1}{2}\sigma_* W_H M_{*H} - \nu_*\Gamma &= \frac{1}{2}W_H M_{*H}(W_{HJ}^2 - W_{HH}W_{JJ}) + \frac{1}{2}W_H W_{HH}(M_{*H}W_{JJ} - M_{*J}W_{HJ}) \\
&= \frac{1}{2}W_H W_{HJ}(M_{*H}W_{HJ} - M_{*J}W_{HH}) \\
&= -\tau_*\rho_* \\
\frac{W_H^2}{2}\sigma_* - \nu_* W_H W_{JJ} + 2\tau_*^2 &= \frac{W_H^2}{2}(W_{HJ}^2 - W_{HH}W_{JJ} + W_{HH}W_{JJ} + W_{HJ}^2) \\
&= W_H^2 W_{HJ}^2 \\
&= 4\tau_*^2 \\
2\tau_*\Gamma - W_H W_{JJ}\rho_* - W_H M_{*J}\sigma_* &= W_H(W_{HJ}(M_{*H}W_{JJ} - M_{*J}W_{HJ}) - W_{JJ}(M_{*J}W_{HH} - M_{*H}W_{HJ}) \\
&\quad - M_{*J}(W_{HJ}^2 - W_{HH}W_{JJ})) \\
&= 2W_H W_{HJ}(M_{*H}W_{JJ} - M_{*J}W_{HJ}) \\
&= 4\tau_*\Gamma \\
\sigma_* M_{*J} + \rho_* W_{JJ} &= M_{*J}(W_{HJ}^2 - W_{HH}W_{JJ}) + W_{JJ}(M_{*J}W_{HH} - M_{*H}W_{HJ}) \\
&= W_{HJ}(M_{*J}W_{HJ} - M_{*H}W_{JJ}) \\
&= -W_{HJ}\Gamma \\
\sigma_* W_H M_{*H} - 2\nu_*\Gamma &= W_H\left(M_{*H}(W_{HJ}^2 - W_{HH}W_{JJ}) + W_{HH}(M_{*H}W_{JJ} - M_{*J}W_{HJ})\right) \\
&= W_H W_{HJ}(M_{*H}W_{HJ} - M_{*J}W_{HH}) \\
&= -2\tau_*\rho_*.
\end{aligned}$$

For the coefficients $a'_{11}, b'_{11}, c'_{11}$, we have:

$$\begin{aligned}
a'_{11} &= W_H(\sigma_* A_{33} - A_{23}A_{31}) \\
&= W_H\left[\frac{J}{2}\nu_*\sigma_* + \frac{J}{2}\rho_*\sigma_*\left(\beta + \frac{U}{2}\right) - \frac{1}{2}\sigma_* M_* W_{HJ} + \frac{1}{2}\sigma_* W_H M_{*H}\left(\beta + \frac{U}{2}\right) + \sigma_*\left(\beta + \frac{U}{2}\right)^2 \{W_H, W_{\gamma_*}\}_{J,\gamma_*}\right. \\
&\quad \left. - \left(\nu_* + \rho_*\left(\beta + \frac{U}{2}\right)\right)\left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right)\right] \\
&= W_H\left[\left(\beta + \frac{U}{2}\right)^2 (\sigma_*\{W_H, W_{\gamma_*}\}_{J,\gamma_*} - \rho_*\Gamma) + \left(\beta + \frac{U}{2}\right)\left(\frac{1}{2}\sigma_* W_H M_{*H} - \nu_*\Gamma - \tau_*\rho_*\right) - \frac{1}{2}\sigma_* M_* W_{HJ} - \nu_*\tau_*\right] \\
&= W_H W_{HJ}\left(\beta + \frac{U}{2}\right)^2 (\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}) - 2\tau_*\rho_* W_H\left(\beta + \frac{U}{2}\right) - \sigma_*\tau_* M_* - \nu_*\tau_* W_H \\
b'_{11} &= iW_H(\sigma_* a_{33} - A_{23}a_{31} + 2\tau_* A_{31}) \\
&= iW_H\left(\frac{W_H^2}{2}\sigma_* + J\tau_*\sigma_* - W_H M_{*J}\sigma_*\left(\beta + \frac{U}{2}\right) - W_H W_{JJ}\left(\nu_* + \rho_*\left(\beta + \frac{U}{2}\right)\right) + 2\tau_*\left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right)\right) \\
&= iW_H\left(\frac{W_H^2}{2}\sigma_* - \nu_* W_H W_{JJ} + 2\tau_*^2 + 2J\tau_*\sigma_* + (2\tau_*\Gamma - W_H W_{JJ}\rho_* - W_H M_{*J}\sigma_*)\left(\beta + \frac{U}{2}\right)\right) \\
&= iW_H\left(4\tau_*^2 + 2J\tau_*\sigma_* + 4\tau_*\Gamma\left(\beta + \frac{U}{2}\right)\right) \\
c'_{11} &= -2\tau_* W_H a_{31} \\
&= -2\tau_* W_H^2 W_{JJ}
\end{aligned}$$

Next for $a'_{12}, b'_{12}, c'_{12}$:

$$\begin{aligned}
a'_{12} &= W_H(\sigma_* A_{32} - A_{22}A_{31}) \\
&= W_H\left[-\frac{J}{2}\rho_*\sigma_* - \sigma_*\left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*} + \rho_*\left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right)\right] \\
&= W_H\left[\left(\beta + \frac{U}{2}\right)(\rho_*\Gamma - \sigma_*\{W_H, W_{\gamma_*}\}_{J,\gamma_*}) + \tau_*\rho_*\right] \\
&= \tau_*\rho_* W_H - 2\tau_*\left(\beta + \frac{U}{2}\right)(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}) \\
b'_{12} &= -iW_H\left[\sigma_* A_{34} + \sigma_* U A_{32} - \sigma_* a_{32} - A_{24}A_{31} - UA_{22}A_{31} + A_{22}a_{31}\right] \\
&= -iW_H\left[J\nu_*\sigma_* + \sigma_* W_H M_{*H}\left(\beta + \frac{U}{2}\right) - \sigma_* M_* W_{HJ} + \frac{J}{2}U\rho_*\sigma_* + U\sigma_*\left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*}\right. \\
&\quad - \frac{J}{2}U\rho_*\sigma_* - U\sigma_*\left(\beta + \frac{U}{2}\right)\{W_H, W_{\gamma_*}\}_{J,\gamma_*} - \sigma_* W_H M_{*J} + \sigma_* W_{HJ} M_* - (2\nu_* + U\rho_*)\left(\tau_* + \frac{J\sigma_*}{2} + \Gamma\left(\beta + \frac{U}{2}\right)\right) \\
&\quad \left. + U\rho_*\left(\tau_* + \frac{J\sigma_*}{2} + \Gamma\left(\beta + \frac{U}{2}\right)\right) - \rho_* W_H W_{JJ}\right] \\
&= -iW_H\left[-2\nu_*\tau_* - W_H(\sigma_* M_{*J} + \rho_* W_{JJ}) + \left(\beta + \frac{U}{2}\right)(\sigma_* W_H M_{*H} - 2\nu_*\Gamma)\right] \\
&= 2i\tau_* W_H\left(\nu_* - \Gamma + \rho_*\left(\beta + \frac{U}{2}\right)\right) \\
c'_{12} &= W_H\left[\sigma_* a_{34} + \sigma_* U a_{32} - A_{24}a_{31} - UA_{22}a_{31}\right] \\
&= W_H\left[\sigma_*(W_H^2 + UM_* W_{HJ} - UW_H M_{*J}) + U\sigma_*(W_H M_{*J} - W_{HJ} M_*) - W_H W_{JJ}(2\nu_* + U\rho_*) + U\rho_* W_H W_{JJ}\right] \\
&= W_H\left[\sigma_* W_H^2 - 2\nu_* W_H W_{JJ}\right] \\
&= W_H^3 W_{HJ}^2 \\
&= 4\tau_*^2 W_H.
\end{aligned}$$

Now for $a'_{21}, b'_{21}, c'_{21}$:

$$\begin{aligned}
a'_{21} &= \sigma_* A_{43} - A_{23}A_{41} \\
&= \sigma_* \left(-\frac{W_H^2 M_{*H}}{2} - W_H \left(\beta + \frac{U}{2}\right) \{W_H, W_{\gamma_*}\}_{J,\gamma_*} - M_* \left(\nu_* + \rho_* \left(\beta + \frac{U}{2}\right)\right)\right) \\
&- \left(\nu_* + \rho_* \left(\beta + \frac{U}{2}\right)\right) (-\Gamma W_H - \sigma_* M_*) \\
&= W_H \left(\beta + \frac{U}{2}\right) (\rho_* \Gamma - \sigma_* \{W_H, W_{\gamma_*}\}_{J,\gamma_*}) + \nu_* \Gamma W_H - \frac{W_H^2}{2} \sigma_* M_{*H} \\
&= -2\tau \left(\beta + \frac{U}{2}\right) (\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}) + \tau_* \rho_* W_H \\
b'_{21} &= i(2\tau_* \sigma_* M_* + 2\tau_* W_H A_{23} + 2\tau_* A_{41}) \\
&= 2i\tau_* \left(\sigma_* M_* + W_H \left(\nu_* + \rho_* \left(\beta + \frac{U}{2}\right)\right) - \Gamma W_H - \sigma_* M_*\right) \\
&= 2i\tau_* W_H \left(\nu_* - \Gamma + \rho_* \left(\beta + \frac{U}{2}\right)\right) \\
c'_{21} &= 4\tau_*^2 W_H
\end{aligned}$$

Finally for $a'_{22}, b'_{22}, c'_{22}$:

$$\begin{aligned}
a'_{22} &= \sigma_* A_{42} - A_{22} A_{41} \\
&= \sigma_* \left(W_H \{W_H, W_{\gamma_*}\}_{J,\gamma_*} + \rho_* M_*\right) + \rho_* \left(-\Gamma W_H - \sigma_* M_*\right) \\
&= W_H \left(\sigma_* \{W_H, W_{\gamma_*}\}_{J,\gamma_*} - \rho_* \Gamma\right) \\
&= 2\tau_* \left(\Gamma M_{*H} + \rho_* M_{*J} + \sigma_* M_{*\gamma_*}\right) \\
b'_{22} &= -i \left(U\sigma_* A_{42} + \sigma_* A_{44} - A_{24} A_{41} - 2\tau_* W_H A_{22} - U A_{22} A_{41}\right) \\
&= -i\Big(U\sigma_* (W_H \{W_H, W_{\gamma_*}\}_{J,\gamma_*} + \rho_* M_*) + \sigma_* \left(-W_H^2 M_{*H} - U W_H \{W_H, W_{\gamma_*}\}_{J,\gamma_*} - M_*(2\nu_* + U\rho_*)\right) \\
&- (2\nu_* + U\rho_*)(-\Gamma W_H - \sigma_* M_*) + 2\tau_* \rho_* W_H + U\rho_* (-\Gamma W_H - \sigma_* M_*)\Big) \\
&= -i \left(-\sigma_* W_H^2 M_{*H} + 2\nu_* \Gamma W_H + 2\tau_* \rho_* W_H\right) \\
&= -4i\tau_* \rho_* W_H \\
c'_{22} &= 2\tau_* W_H A_{24} + 2\tau_* U W_H A_{22} \\
&= 2\tau_* W_H \left(2\nu_* + U\rho_* - U\rho_*\right) \\
&= 4\tau_* \nu_* W_H.
\end{aligned}$$

For the calculation of the quadratics in equation (114), we have:

$$\begin{aligned}
Q_{11}(\lambda, \mu) &= \left[a'_{11} + \left(\beta + \frac{U}{2}\right) a'_{21} + \left(\beta + \frac{U}{2}\right) \left(a'_{12} + \left(\beta + \frac{U}{2}\right) a'_{22}\right)\right] \lambda^2 \\
&+ \left[b'_{11} + \left(\beta + \frac{U}{2}\right) b'_{21} + \left(\beta + \frac{U}{2}\right) \left(b'_{12} + \left(\beta + \frac{U}{2}\right) b'_{22}\right)\right] \lambda\mu \\
&+ \left[c'_{11} + \left(\beta + \frac{U}{2}\right) c'_{21} + \left(\beta + \frac{U}{2}\right) \left(c'_{12} + \left(\beta + \frac{U}{2}\right) c'_{22}\right)\right] \mu^2 \\
&= -\tau_* \left[\nu_* W_H + \sigma_* M_*\right] \lambda^2 + 2i\tau_* W_H \left[2\tau_* + 2\left(\beta + \frac{U}{2}\right) \nu_* + J\sigma_*\right] \lambda\mu \\
&+ 2\tau_* W_H \left[4\tau_* \left(\beta + \frac{U}{2}\right) + 2\nu_* \left(\beta + \frac{U}{2}\right)^2 - W_H W_{JJ}\right] \mu^2 \\
&= \frac{\tau_*}{2^{15}\pi^5 \sigma_*} \left(d_2 \lambda^2 + d_1 \lambda\mu + d_0 \mu^2\right)
\end{aligned}$$

Next:

$$\begin{aligned}Q_{12}(\lambda,\mu) &= \left[a'_{12} + \left(\beta + \frac{U}{2}\right) a'_{22}\right]\lambda^2 + \left[b'_{12} + \left(\beta + \frac{U}{2}\right) b'_{22}\right]\lambda\mu + \left[c'_{12} + \left(\beta + \frac{U}{2}\right) c'_{22}\right]\mu^2 \\ &= \rho_*\tau_* W_H \lambda^2 + 2i\tau_* W_H\left[\nu_* - \Gamma - \rho_*\left(\beta + \frac{U}{2}\right)\right]\lambda\mu + 4\tau_* W_H\left[\tau_* + \nu_*\left(\beta + \frac{U}{2}\right)\right]\mu^2 \\ &= \frac{\tau_*}{2^{15}\pi^5\sigma_*}\left(b_2\lambda^2 + b_1\lambda\mu + b_0\mu^2\right).\end{aligned}$$

Using the symmetry of the matrix in equation (113), we have:

$$\begin{aligned}Q_{21}(\lambda,\mu) &= \left[a'_{21} + \left(\beta + \frac{U}{2}\right) a'_{22}\right]\lambda^2 + \left[b'_{21} + \left(\beta + \frac{U}{2}\right) b'_{22}\right]\lambda\mu + \left[c'_{21} + \left(\beta + \frac{U}{2}\right) c'_{22}\right]\mu^2 \\ &= \left[a'_{12} + \left(\beta + \frac{U}{2}\right) a'_{22}\right]\lambda^2 + \left[b'_{12} + \left(\beta + \frac{U}{2}\right) b'_{22}\right]\lambda\mu + \left[c'_{12} + \left(\beta + \frac{U}{2}\right) c'_{22}\right]\mu^2 \\ &= Q_{12}(\lambda,\mu).\end{aligned}$$

Finally:

$$\begin{aligned}Q_{22}(\lambda,\mu) &= a'_{22}\lambda^2 + b'_{22}\lambda\mu + c'_{22}\mu^2 \\ &= \frac{\tau_*}{2^{15}\pi^5\sigma_*}\left(a_2\lambda^2 + a_1\lambda\mu + a_0\mu^2\right).\end{aligned}$$

Appendix F Overview of the case $W_{JJ} \neq 0$

In this appendix, we provide an overview for finding the characteristics of the Whitham modulation equations (89) to (92) when W_{JJ} and W_{HH} are assumed to be non-vanishing. This provides justification for equation (118). In a similar manner to equation (93), we start by applying the implicit function theorem to $W_J = 0$, which yields a continuously differentiable function h such that:

$$J = h(H, U, \gamma_*, \beta).$$

Taking derivatives of $W_J = 0$ provides the relations:

$$\begin{aligned}\partial_H: &\quad h_H W_{JJ} = -W_{HJ} \\ \partial_U: &\quad h_U W_{JJ} = -W_{UJ} \\ \partial_\gamma: &\quad h_{\gamma_*} W_{JJ} = -W_{\gamma_* J} \\ \partial_\beta: &\quad h_\beta W_{JJ} = -W_{\beta J}.\end{aligned}$$

We now apply the chain rule to the T and X derivatives in the modulation equations in order to write them in terms of derivatives of the parameters. With $z = T, X$, the required derivatives are:

$$\begin{aligned}W_{JJ}\partial_z W_H &= \{W_H, W_J\}_{H,J} H_z + \{W_\gamma, W_J\}_{H,J}\gamma_{*z} + \{W_U, W_J\}_{H,J} U_z + \{W_\beta, W_J\}_{H,J}\beta_z \\ &= -\sigma_* H_z + \Gamma\gamma_{*z} - \left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right) U_z - (2\tau_* + J\sigma_* + U\Gamma)\beta_z \\ W_{JJ}\partial_z W_U &= \{W_U, W_J\}_{H,J} H_z + \{W_U, W_J\}_{\gamma_*,J}\gamma_{*z} + \{W_U, W_J\}_{U,J} U_z + \{W_U, W_J\}_{\beta,J}\beta_z \\ &= -\left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right) H_z + \left(-\frac{W_H M_{*J}}{2} + J\Gamma - \left(\beta + \frac{U}{2}\right)\{W_{\gamma_*}, W_J\}_{\gamma_*,J}\right)\gamma_{*z} \\ &+ \left(-\frac{1}{4}W_H^2 - J\tau_* - \frac{J^2}{4}\sigma_* - \frac{1}{2}M_* W_{JJ} + W_H M_{*J}\left(\beta + \frac{U}{2}\right) - J\Gamma\left(\beta + \frac{U}{2}\right) + \left(\beta + \frac{U}{2}\right)^2\{W_{\gamma_*}, W_J\}_{\gamma_*,J}\right) U_z \\ &+ \left(-\frac{1}{2}W_H^2 - 2J\tau_* - \frac{J^2}{2}\sigma_* - M_* W_{JJ} + W_H M_{*J}(\beta + U) - J\Gamma(\beta + U) + U\left(\beta + \frac{U}{2}\right)\{W_{\gamma_*}, W_J\}_{\gamma_*,J}\right)\beta_z \\ W_{JJ}\partial_z W &= W_H W_{JJ} H_z + M_* W_{JJ}\gamma_{*z} + W_U W_{JJ} U_z + W_{JJ}(J W_H - U M_*)\beta_z \\ W_{JJ}\partial_z W_{\gamma_*} &= \{W_{\gamma_*}, W_J\}_{H,J} H_z + \{W_{\gamma_*}, W_J\}_{\gamma_*,J}\gamma_{*z} + \{W_{\gamma_*}, W_J\}_{U,J} U_z + \{W_{\gamma_*}, W_J\}_{\beta,J}\beta_z \\ &= \Gamma H_z + \{W_{\gamma_*}, W_J\}_{\gamma_*,J}\gamma_{*z} + \left(-\frac{W_H M_{*J}}{2} + J\Gamma - \left(\beta + \frac{U}{2}\right)\{W_{\gamma_*}, W_J\}_{\gamma_*,J}\right) U_z \\ &+ \left(-W_H M_{*J} + J\Gamma - U\{W_{\gamma_*}, W_J\}_{\gamma_*,J}\right)\beta_z.\end{aligned}$$

As in equation (99), we write the modulation equations in the quasi-linear form:

$$A \begin{pmatrix} H_T \\ \gamma_{*T} \\ U_T \\ \beta_T \end{pmatrix} - a \begin{pmatrix} H_X \\ \gamma_{*X} \\ U_X \\ \beta_X \end{pmatrix} = 0 \tag{125}$$

with

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

$$a = \begin{pmatrix} 0 & 1 & 0 & 0 \\ UA_{21} & UA_{22} & UA_{23} - W_H W_{JJ} & UA_{24} \\ UA_{31} + a_{31} & UA_{32} + a_{32} & UA_{33} & UA_{34} + a_{34} \\ UA_{41} + 2\tau W_H & UA_{42} + W_H^2 M_{*J} & UA_{43} + a_{43} & UA_{44} + a_{44} \end{pmatrix}$$

and the coefficients

$$A_{21} = -\sigma_*$$
$$A_{22} = \Gamma$$
$$A_{23} = -\left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right)$$
$$A_{24} = -(2\tau_* + J\sigma_* + U\Gamma)$$
$$A_{31} = -\left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right)$$
$$A_{32} = -\frac{W_H M_{*J}}{2} + \frac{J}{2}\Gamma - \left(\beta + \frac{U}{2}\right)\{W_{\gamma_*}, W_J\}_{\gamma_*,J}$$
$$A_{33} = \left(-\frac{1}{4}W_H^2 - J\tau_* - \frac{J^2}{4}\sigma_* - \frac{1}{2}M_* W_{JJ} + W_H M_{*J}\left(\beta + \frac{U}{2}\right) - J\Gamma\left(\beta + \frac{U}{2}\right) + \left(\beta + \frac{U}{2}\right)^2 \{W_{\gamma_*}, W_J\}_{\gamma_*,J}\right)$$
$$A_{34} = \left(-\frac{1}{2}W_H^2 - 2J\tau_* - \frac{J^2}{2}\sigma_* - M_* W_{JJ} + W_H M_{*J}(\beta + U) - J\Gamma(\beta + U) + U\left(\beta + \frac{U}{2}\right)\{W_{\gamma_*}, W_J\}_{\gamma_*,J}\right)$$
$$A_{41} = \Gamma W_H + \sigma_* M_*$$
$$A_{42} = W_H\{W_{\gamma_*}, W_J\}_{\gamma_*,J} - \Gamma M_*$$
$$A_{43} = W_H\left(-\frac{W_H M_{*J}}{2} + \frac{J}{2}\Gamma - \left(\beta + \frac{U}{2}\right)\{W_{\gamma_*}, W_J\}_{\gamma_*,J}\right) + M_*\left(\tau_* + \frac{J}{2}\sigma_* + \Gamma\left(\beta + \frac{U}{2}\right)\right)$$
$$A_{44} = W_H(-W_H M_{*J} + J\Gamma - U\{W_{\gamma_*}, W_J\}_{\gamma_*,J}) + M_*(2\tau_* + J\sigma_* + U\Gamma)$$
$$a_{31} = -W_H W_{JJ}$$
$$a_{32} = -M_* W_{JJ}$$
$$a_{34} = -W_{JJ}(JW_H - UM_*)$$
$$a_{43} = M_* W_H W_{JJ} + W_H^2\left(\frac{W_H}{2} + \frac{JW_{HJ}}{2} - M_{*J}\left(\beta + \frac{U}{2}\right)\right)$$
$$a_{44} = W_H^2(W_H + JW_{HJ} - UM_{*J}).$$

The equation for the characteristics is

$$AX' - aT' = 0.$$

The matrices A, a satisfy equation (103), since we can compute:

$$\det(AX') = \frac{W_H W_{JJ}^2}{2^{10}\pi^4} X'^4 D \neq 0. \tag{126}$$

Moreover, we can follow the same procedure of applying the Schur determinant formula. The upper-left block matrix is:

$$P_{11} = \begin{pmatrix} 0 & -T' \\ -\sigma_*(X' - UT') & \Gamma(X' - UT') \end{pmatrix},$$

which is only singular when $X' = T' = 0$ (from equation (126)), or when $X' = UT'$. However, $X' = UT'$ leads to the determinant calculation:

$$\det(UAT' - aT') = W_H^5 W_{JJ}^2 T'^4,$$

which only vanishes when $T' = 0$, leading to the trivial solution $X' = T' = 0$. From here, the procedure of using the Schur determinant formula and manipulating the determinant calculation is the same as in the Section 3.1. The end result is given in equation (118).

References

[AS1981] M.J. ABLOWITZ AND H. SEGUR, *Solitons and the inverse scattering transform*, SIAM (1981), Philadelphia.

[Agrawal2013] G.P. AGRAWAL, *Nonlinear fiber optics*, Elsevier/Academic Press (2013), Amsterdam, Fifth edition.

[AIK1990] G.L. ALFIMOV, A.R. ITS AND N.E. KULAGIN, *Modulation instability of solutions of the nonlinear Schrödinger equation*, Teoret. Mat. Fiz. **84** (1990), no. 2, 787–793.

[Bridges2015] T.J. BRIDGES, *Breakdown of the Whitham modulation theory and the emergence of dispersion*, Stud. Appl. Math. **135** (2015), no. 3, 277–294.

[BR2019] T.J. BRIDGES AND D.J. RATLIFF, *Krein signatures and Whitham modulation theory: the sign of characteristics and the 'sign characteristic'*, Stud. Appl. Math. **142** (2019), no. 3, 314–335.

[BJ2010] J.C. BRONSKI AND M.A. JOHNSON, *The modulational instability for a generalized Korteweg-de Vries equation*, Arch. Mech. Rat. Anal. **197** (2010), no. 2, 357–400.

[BDN2011] N. BOTTMAN, B. DECONINCK AND M. NIVALA, *Elliptic solutions of the defocusing NLS equation are stable*, J. Phys. A. **44** (2011), no. 28, 1–24.

[Chen2016] F.F. CHEN, *Introduction to plasma physics and controlled fusion*, Cham: Springer International Publishing (2016). Reprint of the 1984 edition.

[DS2017] B. DECONINCK AND B.L. SEGAL, *The stability spectrum for elliptic solutions to the focusing NLS equation*, Phys. D **346** (2017), 1–19.

[DS2009] W.-P. DÜLL AND G. SCHNEIDER, *Validity of Whitham's equations for the modulation of periodic traveling waves in the NLS equation*, J. Nonlinear Sci. **19** (2009), no. 5, 453–466.

[EH2016] G.A. EL AND M.A. HOEFER, *Dispersive shock waves and modulation theory*, Phys. D **333** (2016), 11–65.

[GH2007A] T. GALLAY AND M. HĂRĂGUŞ, *Stability of small periodic waves for the nonlinear Schrödinger equation*, J. Differential Equations **243** (2007), no. 2, 544–581.

[GH2007B] T. GALLAY AND M. HĂRĂGUŞ, *Orbital stability of periodic waves for the nonlinear Schrödinger equation*, J. Dynam. Differential Equations **19** (2007), no. 4, 825–865

[GSS1987] M. GRILLAKIS, J. SHATAH AND W. STRAUSS, *Stability theory of solitary waves in the presence of symmetry, I*, J. Func. Anal. **74** (1987), no. 1, 160–197.

[GSS1990] M. GRILLAKIS, J. SHATAH AND W. STRAUSS, *Stability theory of solitary waves in the presence of symmetry, II*, J. Func. Anal. **94** (1990), no. 2, 308–348.

[Gross1961] E.P. GROSS, *Structure of a quantized vortex in boson systems*, Nuovo Cimento **20** (1961), no. 3, 454–477.

[GLCT2017] S. GUSTAFSON, S. LE COZ AND T.-P. TSAI, *Stability of periodic waves of 1D cubic nonlinear Schrödinger equations*, Appl. Math. Res. Express. AMRX, **2017** (2017), no. 2, 431–487.

[HM2003] A. HASEGAWA AND M. MATSUMOTO, *Optical solitons in fibers*, Springer, (2003), Berlin. Third edition.

[HO1972] H. HASIMOTO AND H. ONO, *Nonlinear modulation of gravity waves*, J. Phys. Soc. Japan **33** (1972), no. 3, 805–811.

[JMMP2014] C.K.R.T. JONES, R. MARANGELL, P.D. MILLER AND R.G. PLAZA, *Spectral and modulational stability of periodic wavetrains for the nonlinear Klein-Gordon equation*, Journal of Differential Equations **257** (2014), no. 12, 4632–4703.

[JZ2010] M.A. JOHNSON AND K. ZUMBRUN, *Rigorous justification of the Whitham modulation equations for the generalized Korteweg-de Vries equation*, Stud. Appl. Math. **125** (2010), no. 1, 69–89.

[Kamchatnov2000] A.M. KAMCHATNOV, *Nonlinear periodic waves and their modulations: an introductory course*, World Scientific Publishing (2000), Singapore.

[LBJM2019] K.P. LEISMAN, J.C. BRONSKI, M.A. JOHNSON AND R. MARANGELL, *Stability of Traveling wave solutions of Nonlinear Dispersive equations of NLS type*, Preprint (2019), arXiv:1910.05392.

[LTE2019] C.S. LIU, V.K. TRIPATHI AND B. ELIASSON, *High-power laser-plasma interaction*, Cambridge University Press, (2019), Cambridge.

[MOMT1976] K. MIO, T. OGINO, K. MINAMI AND S. TAKEDA, *Modified nonlinear Schrödinger equation for Alfvén waves propagating along the magnetic field in cold plasmas*, J. Phys. Soc. Japan **41** (1976), no. 1, 265–271.

[OZ2003] M. OH AND K. ZUMBRUN, *Stability of periodic solutions of conservation laws with viscosity: analysis of the evans function*, Arch. Ration. Mech. Anal. **166** (2003), no. 2, 99–166.

[OZ2006] M. OH AND K. ZUMBRUN, *Low-frequency stability analysis of periodic travelling-wave solutions of viscous conservation laws in several dimensions*, Z. Anal. Anwend. **25** (2006), no. 1, 1–21.

[Pitaevskii1961] L.P. PITAEVSKII, *Vortex lines in an imperfect Bose gas*, Sov. Phys.–JETP **13** (1961), no. 2, 451–454.

[Rowlands1974] G. ROWLANDS, *On the stability of solutions of the non-linear Schrödinger equation*, IMA J. Appl. Math. **13** (1974), no. 3, 367–377.

[Serre2005] D. SERRE, *Spectral stability of periodic solutions of viscous conservation laws: large wavelength analysis*, Communications in Partial Differential Equations **30** (2005), no. 1–2, 259–282.

[SS1999] C. SULEM AND P.-L. SULEM, *The nonlinear Schrödinger equation: self-focusing and wave collapse*, Springer-Verlag, (1999), New York.

[Whi1965A] G.B. WHITHAM, *Non-linear dispersive waves*, Proc. R. Soc. Lond. A **283** (1965), no. 1393, 238–261.

[Whi1965B] G.B. WHITHAM, *A general approach to linear and non-linear dispersive waves using a Lagrangian*, Journal of Fluid Mechanics **22** (1965), no. 2, 273–283.

[Whi1967] G.B. WHITHAM, *Variational methods and applications to water waves*, Proc. R. Soc. Lond. A **299** (1967), no. 1456, 6–25.

[Whi1970] G.B. WHITHAM, *Two-timing, variational principles and waves*, Journal of Fluid Mechanics **44** (1970), no. 2, 373–395.

[Whi1999] G.B. WHITHAM, *Linear and nonlinear waves*, John Wiley & Sons, Inc. (1999), New York. Reprint of the 1974 original.

[Zakharov1968] V.E. ZAKHAROV, *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, J. Appl. Mech. Tech. Phys. **9** (1968), no. 2, 190–194.

[Zhang2006] F. ZHANG, *The Schur complement and its applications*, Springer Science & Business Media (2006), vol. 4.