ARMA manuscript No.
(will be inserted by the editor)

The symplectic Evans matrix, and the instability of solitary waves and fronts with symmetry

Thomas J. Bridges & Gianne Derks

Abstract

Hamiltonian evolution equations which are equivariant with respect to the action of a Lie group are models for physical phenomena such as oceanographic flows, optical fibres and atmospheric flows, and such systems often have a wide variety of solitary wave or front solutions. In this paper, we present a new symplectic framework for analyzing the spectral problem associated with the linearization about such solitary waves and fronts. At the heart of the analysis is a multi-symplectic formulation of Hamiltonian partial differential equations where a distinct symplectic structure is assigned for the time and space directions, with a third symplectic structure – with two-form denoted by Ω – associated with a coordinate frame moving at the speed of the wave. This leads to a geometric decomposition and symplectification of the Evans function formulation for the linearization about solitary waves and fronts.

We introduce the concept of the *symplectic Evans matrix*, a matrix consisting of restricted $\Omega-$symplectic forms. By applying Hodge duality to the exterior algebra formulation of the Evans function, we find that the zeros of the Evans function correspond to zeros of the determinant of the symplectic Evans matrix. Based on this formulation, we prove several new properties of the Evans function. Restricting the spectral parameter λ to the real axis, we obtain rigorous results on the derivatives of the Evans function near the origin, based solely on the abstract geometry of the equations, and results for the large $|\lambda|$ behaviour which use primarily the symplectic structure, but also extend to the non-symplectic case.

The Lie group symmetry affects the Evans function by generating zero eigenvalues of large multiplicity in the so-called systems at infinity. We present a new geometric theory which describes precisely how these zero eigenvalues behave under perturbation.

By combining all these results, a new rigorous sufficient condition for instability of solitary waves and fronts is obtained.

The theory applies to a large class of solitary waves and fronts including waves which are bi-asymptotic to a nonconstant manifold of states as $|x|$ tends to infinity. To illustrate the theory, it is applied to three examples: a Boussinesq model from oceanography, a class of nonlinear Schrödinger equations from optics and a nonlinear Klein-Gordon equation from atmospheric dynamics.

Table of Contents

1. Introduction

Solitary waves and fronts are a class of solutions of translation-invariant partial-differential equations. Such partial differential equations feature as models for a large range of physical phenomena in oceanography, atmospheric dynamics, optical fibre transmission, nerve conduction, acoustical and gas dynamic flows. Predominantly, it is the stability or instability of a solitary wave or front which determines its relevance for the physical phenomena.

Many models with solitary wave and front solutions are conservative translation-invariant evolution equations with a Hamiltonian structure, and the calculus of variations, critical point theory and symplectic structure have played an important role in the analysis of the stability and instability of their solitary-wave solutions. For example, the characterization of solitary waves as critical points of the energy (or Hamiltonian) constrained to level sets of the momentum (or momentum and other constants of motion) leads to a powerful framework for proving nonlinear Lyapunov stability – when the second variation, evaluated at the constrained critical point, has a finite number of negative eigenvalues (e.g. BENJAMIN [6], BONA [11], HOLM ET AL [27], GRILLAKIS ET AL [25,26], MADDOCKS & SACHS [35] and references therein). Indeed, one could argue that the study of the stability theory of solitary waves began in this way motivated by the paradigm of solitary waves: the KdV solitary wave [6].

However, for many Hamiltonian evolution equations, particularly coupled systems of PDEs, even though the characterization of a solitary-wave or front solution as a constrained critical point is well-defined, the second variation is strongly indefinite and the relation between critical point type and stability is lost. In this case, an important first step is to study the linear stability and instability, that is, analyse the spectral problem associated with the linearization about the solitary-wave or front solution.

A dynamical systems approach for the analysis of spectral problems associated with the linearization about a solitary wave or front was first introduced by EVANS [21] in the context of the stability analysis of nerve impulses in mathematical biology. The Evans function framework was substantially generalized by ALEXANDER, GARDNER & JONES [2] to apply to a large class of parabolic PDEs. The cornerstone of the Evans theory is the *Evans function*, $D(\lambda)$, a complex analytic function of the spectral parameter $\lambda \in \mathbb{C}$. Under suitable hypotheses, the Evans function has the property that, if $\lambda_0 \in \mathbb{C}$ has positive real part and $D(\lambda_0) = 0$, then λ_0 is an unstable eigenvalue associated with the linearization about a solitary wave. There is now a wide range of analyses of the stability of solitary waves using the Evans function framework (see e.g. GARDNER & ZUMBRUN [23] and KAPITULA & SANDSTEDE [30] for recent results and references).

One way to prove the existence of unstable eigenvalues is to study the sign of $D(\lambda)$ for λ real when λ is near zero and when λ is large. When the initial-value problem for the PDE is well-posed one can expect that there exists a positive real number α such that when λ is real and $\lambda > \alpha$, $D(\lambda)$ will be of one sign; that is, there would not exist unstable eigenvalues with arbitrarily large growth rate. Assume $D(\lambda) > 0$ for $\lambda > \alpha$; then a negative sign of the slope of the Evans function for λ near zero can be used to predict the existence of unstable eigenvalues along the real λ axis. Indeed, EVANS [21] proposed a formula for $D'(0)$, and used this idea to predict instabilities in a biological model (EVANS & FEROE [22]). This idea has been substantially generalized by ALEXANDER & JONES [3] by introducing an *orientation index*, and this index has been given a topological interpretation by NII [38]. In a Hamiltonian setting there is much more geometry available and therefore one can expect to get more information about the derivatives of $D(\lambda)$.

The connection between the Evans function framework and the stability analysis of solitary wave solutions of Hamiltonian evolution equations was first studied by PEGO & WEINSTEIN [42]. They obtained the result that $D(\lambda)$ satisfies

$$D(0) = D'(0) = 0\,, \quad \operatorname{sign} D''(0) = \operatorname{sign} \frac{\mathrm{d}I}{\mathrm{d}c}\,, \tag{1.1}$$

and that $D(\lambda) \to 1$ as $\lambda \to +\infty$ along the real axis. In equation (1.1), I is the value of the momentum level set and c is the speed of the solitary wave.

The importance of the result (1.1) is twofold: it uses a property of the constrained variational principle – solitary waves characterized as critical points of the energy restricted to level sets of the momentum – without any hypotheses on the second variation, and, by the intermediate-value theorem, $\mathrm{d}I/\mathrm{d}c < 0$ is a sufficient condition for the existence of at least one unstable eigenvalue. However, in the theory of PEGO AND WEINSTEIN, the result applied to only three particular Hamiltonian PDEs, and the system of ODEs associated with the spectral problem had no special structure, requiring explicit calculations in parts of the proof of (1.1). This limited application of (1.1) to the particular PDEs studied and to cases where the solitary wave was known explicitly.

The primary difficulty with an abstract Evans function framework for Hamiltonian evolution equations is that the classical Hamiltonian formulation provides a symplectic structure for time evolution, but much of the analysis of the Evans function is associated with a dynamical system in the x-variable. In BRIDGES & DERKS [16] an abstract formulation of the Evans function for Hamiltonian PDEs was proposed based on a multi-symplectic formulation of the PDE, where distinct symplectic operators are assigned for time and space. With two symplectic structures it is possible to give an abstract geometric characterization of the ODEs associated with the Evans function framework. To be precise, a Hamiltonian system on a multi-symplectic structure will be written in the canonical form

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z)\,, \quad Z \in \mathbb{R}^{2n}\,, \quad x \in \mathbb{R}\,, \quad t \geq 0\,, \tag{1.2}$$

where $\mathbf{M}$ and $\mathbf{K}$ are skew-symmetric constant matrices, $S : \mathbb{R}^{2n} \to \mathbb{R}$ is some smooth function and ∇ is a gradient with respect to the standard inner product on $\mathbb{R}^{2n}$.

To demonstrate the multi-symplectification of a Hamiltonian PDE, we consider the nonlinear Schrödinger equation:

$$\mathrm{i}\psi_t = \psi_{xx} + W'(|\psi|^2)\psi = 0\,, \tag{1.3}$$

where $\psi(x,t)$ is a complex-valued function of x and t, and $W(\cdot)$ is some smooth real-valued function. This system can be written as a classical Hamiltonian system on an infinite-dimensional phase space. For example, let $\psi = q_1 + iq_2$, then (1.3) can be reformulated as

$$q_t = \mathbf{J}\frac{\delta H}{\delta q}\,, \quad \text{where} \quad \mathbf{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\,,$$

and

$$H(q) = \frac{1}{2}\int_{x_1}^{x_2} \left[\left(\frac{\partial q_1}{\partial x}\right)^2 + \left(\frac{\partial q_1}{\partial x}\right)^2 - W(q_1^2 + q_2^2) \right] \mathrm{d}x\,. \tag{1.4}$$

However, the phase space is infinite-dimensional and the spatial symplecticity is not explicit in this formulation, but it is implicit in the Hamiltonian function. There are several ways to formulate this PDE as a multi-symplectic system on a *finite-dimensional phase space* (see §11 for details). Effectively, one takes the Legendre transform of the Hamiltonian function (1.4). For example, with

$$p_1 + \mathrm{i}p_2 = \psi_x = (q_1 + iq_2)_x\,,$$

the form (1.2) is recovered by taking

$$\mathbf{M} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{K} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\,,$$

with

$$Z = (q,p)^T \in \mathbb{R}^4 \quad \text{and} \quad S(Z) = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}W(q_1^2 + q_2^2)\,.$$

In the multi-symplectic framework, the shape $\widehat{Z}$ of the solitary wave solution of (1.2), travelling at speed c, $Z(x,t) = \widehat{Z}(x-ct)$, is a homoclinic or heteroclinic orbit of the ODE

$$(\mathbf{K} - c\mathbf{M})\widehat{Z}_x = \nabla V(\widehat{Z})\,, \tag{1.5}$$

where V is S plus additional terms associated with symmetry (see §2 for precise definition). The linearization of (1.2) about this solitary wave reduces to a linear ODE of the form

$$U_x = \mathbf{A}(x,\lambda;\mathbf{p})\,U\,, \quad U \in \mathbb{C}^{2n}\,, \quad \lambda \in \mathbb{C}, \tag{1.6}$$

where $\lambda \in \mathbb{C}$ is the spectral parameter, $\mathbf{p}$ represents parameters associated with the family of solitary waves (see §§2-3 for precise definition), and

$$(\mathbf{K} - c\mathbf{M})\mathbf{A}(x,\lambda;\mathbf{p}) := D^2V(\widehat{Z}(x;\mathbf{p})) - \lambda\,\mathbf{M}\,. \tag{1.7}$$

The solitary wave is an orbit of the Hamiltonian ODE (1.5) with symplectic operator $(\mathbf{K}-c\mathbf{M})$ and the coefficient matrix $\mathbf{A}(x,\lambda;\mathbf{p})$ of the spectral ODE (1.6) has a multi-symplectic decomposition: $\mathbf{M}$ and $\mathbf{K}$ are the symplectic operators associated with the t and x directions respectively and $D^2V(\widehat{Z})$ is a symmetric x-dependent matrix. The above decomposition of the defining equation for the solitary wave and the spectral problem lead to an abstract generalization of (1.1) in [16]. For a large class of Hamiltonian PDEs and a wide range of solitary wave and front solutions, including waves which are biasymptotic to nonconstant manifolds as $x \to \pm\infty$, the second derivative in (1.1) takes the general form

$$\operatorname{sign} D''(0) = \varPi \operatorname{sign}\left(\frac{\mathrm{d}I}{\mathrm{d}c} - \mathcal{B}(c)\right)\,, \tag{1.8}$$

where $\varPi = \pm 1$ is a geometric sign associated with the shape of the wave, and $\mathcal{B}(c)$ is associated with the properties of the nonconstant manifold of states at infinity (for example, $\mathcal{B}(c) = 0$ for classical solitary waves). This result lead to an abstract geometric condition for instability of solitary waves or fronts. The main restriction in the theory of [16] was that the symplectic constant coefficient systems at infinity associated with (1.6) were required to have a one-dimensional stable subspace. This restriction was not essential but it simplified much of the analysis. In this paper, we generalize the theory of [16] to the case where the stable subspace is arbitrary but finite. This generalization requires the *symplectic Evans function* to be generalized to the *symplectic Evans matrix*. In addition, several other new results are proved: Hodge duality is the key to transforming the exterior-algebra definition of the Evans function introduced in [2] to the Evans matrix; the entries of the Evans matrix correspond to restricted $\varOmega$-symplectic forms; new results are obtained on the derivatives of the Evans matrix near the origin; results on the large $|\lambda|$ behaviour of the Evans function are proved, which extend also to the non-symplectic case; and a new geometric perturbation theory for the zero eigenvalues in the systems at infinity associated with symmetry is developed. These results all combine to prove that a generalization of the condition (1.8) holds. The theory is demonstrated by proving new instability results for a range of examples.

The term instability is to be interpreted as *spectral instability.* Our results give sufficient conditions for the existence of eigenvalues of the spectral problem associated with the linearization about a solitary wave or front. It is reasonable to conjecture that these results will imply the existence of unstable solutions of the Cauchy initial-value problem for the nonlinear PDE in a suitable function space, but such results are outside the scope of this paper.

In general we consider the spectral problem for solitary wave and front solutions for any Hamiltonian evolutionary equation that can be formulated as a *Hamiltonian system on a multi-symplectic structure* as in (1.2). Precise hypotheses on the system (1.2) will be given in the beginning of §2.

Remarkably, every aspect of the theory in the sequel can be proved based only on the abstract formulation (1.2) and it does not rely on the properties of a particular equation.

There are three symplectic structures associated with (1.2): the skew-symmetric matrices $\mathbf{M}$ and $\mathbf{K}$ are associated with two forms ω and κ, with a third symplectic structure associated with a coordinate frame moving at the speed c of the wave, defined by $\Omega = \kappa - c\omega$. Precise definitions of these two-forms are given in §2.

The concept of a multi-symplectic structure, in the sense used here, was introduced in Bridges [13–15], where it was shown to be a natural dynamical systems framework for analysing multi-dimensional patterns. There is an interesting connection between this formulation of Hamiltonian PDEs and the theory based on the Cartan form in field theory. Enlarge the manifold to the *extended phase space* $\mathbb{R}^{2n} \times \mathbb{R}^s$ (where s is the dimension of space-time). Then, for the case considered in this paper where $s = 2$, define the following *meta-symplectic form*

$$\Sigma_{\mathrm{meta}} = \omega \wedge \mathbf{d}x - \kappa \wedge \mathbf{d}t - \mathbf{d}S \wedge \mathbf{d}x \wedge \mathbf{d}t \in \textstyle\bigwedge^3(\mathbb{R}^{2n} \times \mathbb{R}^2)\,. \tag{1.9}$$

The connection with the Cartan form requires starting with a Lagrangian formulation of the problem which after multiple Legendre transform leads to a similar higher-order form – but without distinguishing the role of the two forms. In the case of Hamiltonian PDEs associated with first-order field theories, this connection can be made precise (cf. Marsden & Shkoller [36]). In field theory the exterior derivative of the Cartan form is sometimes called a multi-symplectic form (cf. Binz et al. [9], page 330). For an elementary and historical account of the role of the Cartan form in field theory see Kijowski & Tulczyjew ([32],§20). The Cartan framework provides a higher-order symplectic setting for Hamiltonian PDEs. However, there does not appear to be any theoretical advantage to embedding the family of two forms into the Cartan form setting. Indeed, we will show here that it is the collection of two-forms (ω, κ) on the phase space $\mathbb{R}^{2n}$ – and the canonical form (1.2) – which form the backbone of an abstract framework which is the basis of the theory presented here.

Many models of physical phenomena, particularly in atmospheric dynamics and optics, are equivariant with respect to a Lie group symmetry, in addition to a spatial translation invariance. In fact, almost all the models we have looked at have some additional symmetry. To be precise we assume that the system (1.2) is equivariant with respect to a subgroup of the Euclidean group acting on $\mathbb{R}^{2n}$, denoted by $\mathcal{G}$. Hence the system (1.2) is equivariant with respect to the group $\mathcal{G}$, *as well as spatial translations.*

The proposed framework for the stability analysis of solitary waves and fronts can be summarized as follows. The existence of a solitary wave or front is assumed. The form of this wave can be quite general; for example, we allow for the shape of the wave to be biasymptotic to a multi-dimensional invariant manifold at infinity, $\mathcal{M}_\infty$. The manifold $\mathcal{M}_\infty$ is a submanifold of $\mathbb{R}^{2n}$, and is defined by the orbit of a q-dimensional Abelian subgroup of $\mathcal{G}$. The full details of this characterization are in §2.

The spectral problem associated with the linearization about a wave reduces to a family of complex-valued ODEs on the real line of the form (1.6). An abstract multi-symplectic formulation of the Evans function is then developed for this system of ODEs. Central to the Evans function theory are the systems at infinity, defined by

$$\mathbf{A}^{\pm}(\lambda; \mathbf{p}) = \lim_{x\to\pm\infty} \mathbf{A}(x, \lambda; \mathbf{p})\,.$$

Associated with this parameter dependent matrix are the subspaces

$$E^s_{\pm}(\lambda; \mathbf{p}) = \{\, \xi \in \mathbb{C}^{2n} \ : \ \lim_{x\to+\infty} e^{\mathbf{A}^{\pm}(\lambda;\mathbf{p})x}\xi = 0 \,\}\,, \qquad \lambda \in \mathbb{C} \tag{1.10}$$

$$E^u_{\pm}(\lambda; \mathbf{p}) = \{\, \xi \in \mathbb{C}^{2n} \ : \ \lim_{x\to-\infty} e^{\mathbf{A}^{\pm}(\lambda;\mathbf{p})x}\xi = 0 \,\}\,, \qquad \lambda \in \mathbb{C} \tag{1.11}$$

and $E^c_{\pm}(\lambda; \mathbf{p})$, which is defined to be a complement of $E^s_{\pm}(\lambda; \mathbf{p}) \oplus E^u_{\pm}(\lambda; \mathbf{p})$ in $\mathbb{C}^{2n}$. In general these subspaces will differ at $\pm\infty$. However, in §3 we will prove that, when the state at infinity is an invariant manifold of relative equilibria associated with the group $\mathcal{G}$, the matrices $\mathbf{A}^+(\lambda;\mathbf{p})$ and $\mathbf{A}^-(\lambda;\mathbf{p})$ are not equal in general but have the same spectrum.

For definiteness, the following properties on the dimension of the systems at infinity are taken: for fixed values of the parameters, $\mathbf{p}$,

$$\dim E^s_\pm(0;\mathbf{p}) = \dim E^u_\pm(0;\mathbf{p}) = 1 \quad \text{thus} \quad \dim E^c_\pm(0;\mathbf{p}) = 2(n-1)\,,$$

and, the hypothesis when $\lambda \neq 0$ is that there is some $1 \le p \le n$ such that

$$\min\{\,\dim E^s_\pm(\lambda;\mathbf{p}), \dim E^u_\pm(\lambda;\mathbf{p})\,\} = p\,,$$

for all $\lambda \in \Lambda$, where Λ is subset of $\mathbb{C}_+$, the complex half-plane with positive real part, which will be specified in §3. The Ω-symplectic structure forces the dimensions of $E^s_\pm(0;\mathbf{p})$ and $E^u_\pm(0;\mathbf{p})$ to be equal and the dimension of $E^c_\pm(0;\mathbf{p})$ to be even. Furthermore, the q-dimensional Abelian group which generates $\mathcal{M}_\infty$ implies that $\dim E^c_\pm(0;\mathbf{p}) \ge 2q$. The property $\dim E^{s,u}_\pm(0;\mathbf{p}) = 1$ is not essential and many of the results, such as the construction and definition of the symplectic Evans matrix, are independent of this property. This property avoids bookkeeping and breaking the instability results into special cases, and the treatment of the case where this dimension is greater than one is best treated in particular examples and the details of this generalization will be apparent from the analysis here. The case $p=1$ is treated in detail in Bridges & Derks [16], therefore we will restrict attention to the case $p>1$. An example of the above spaces with $p=2$ is shown in Figure 12.1.

With these hypotheses on the systems at infinity, we prove in §5 that the Evans function takes the geometric form $D(\lambda;\mathbf{p}) = \det(\boldsymbol{\Omega}(\lambda;\mathbf{p}))$, where $\boldsymbol{\Omega}(\lambda;\mathbf{p})$ is the $p\times p$ *symplectic Evans matrix* and each entry of $\boldsymbol{\Omega}(\lambda;\mathbf{p})$ is an Ω-symplectic form restricted to a pair of solutions of the linearized stability problem and its adjoint. It is not difficult to prove that $D(0;\mathbf{p}) = \frac{\partial}{\partial\lambda}\big|_{\lambda=0} D(\lambda;\mathbf{p}) = 0$. However, the proof that $\frac{\partial^2}{\partial\lambda^2}\Big|_{\lambda=0} D(\lambda;\mathbf{p})$ takes a form like (1.8) requires several substantial steps. Firstly, it is essential to understand how the eigenvalues 0 of $\mathbf{A}^\pm(0;\mathbf{p})$ perturb when $\lambda\neq 0$, and for this we present a new geometric perturbation theory in §4. Secondly, the derivatives of $D(\lambda;\mathbf{p})$ at the origin are needed. It is this second issue that requires the full structure given by the multi-symplectic geometry: in §7, we present a generalization of the multi-symplectic Pego-Weinstein derivative formula ([16,42]) which provides a framework for differentiating the Evans matrix.

To give an abstract geometric instability criterion, the asymptotic properties of $D(\lambda)$ for large $|\lambda|$ are needed. In §8 we present a theory for large $|\lambda|$ asymptotics, with hypotheses that are relevant to a range of Hamiltonian PDEs. All these results combine to prove the central result, giving a general instability criterion for a large class of solitary waves and fronts, which is stated in Theorem 1.

To illustrate the theory, it is applied to three examples: a Boussinesq model from oceanography, a class of nonlinear Schrödinger equations from optics and a complex nonlinear Klein-Gordon equation from atmospheric dynamics.

The Boussinesq equations are a class of PDEs which model dispersive shallow-water waves in oceanography, as well as nonlinear strings and other physical phenomena. In §10 the theory is applied to the "good" Boussinesq model. Here our results are twofold: firstly, we prove a conjecture of Alexander & Sachs [4], that the convexity of the momentum and the Evans function are related, when evaluated on a one-parameter family of solitary waves. Secondly, we show that the class of solitary waves is in fact part of a two-parameter family. Here we use the result that the state at infinity is a manifold, and not a point, and in the two-parameter space the set of values where the wave is unstable is determined exactly.

The nonlinear Schrödinger equation (NLS), coupled NLS models and their generalizations are an important class of model equations for wave packets in electromagnetics, water waves, nonlinear optics, vortex filament dynamics and many other areas. The cubic NLS equation was first derived as a model equation for oceanographic wave packets (Benney & Newell [8]) and optics (Zakharov [45]). Since then, this model has been found to be a universal nonlinear PDE, and has been substantially generalized to include other nonlinearities and various forms of coupling (cf. [5,12,24,28,30,33,34,40,43,45–47] and references therein). A wide range of the coupled NLS models can be written in the general form

$$\begin{aligned} \mathrm{i}a_1\frac{\partial A}{\partial t} + \mathrm{i}a_2\frac{\partial A}{\partial x} + a_3\frac{\partial^2 A}{\partial x^2} &= \frac{\delta H}{\delta \overline{A}} \\ \mathrm{i}b_1\frac{\partial B}{\partial t} + \mathrm{i}b_2\frac{\partial B}{\partial x} + b_3\frac{\partial^2 B}{\partial x^2} &= \frac{\delta H}{\delta \overline{B}} \end{aligned} \tag{1.12}$$

where $a_1, a_2, a_3, b_1, b_2, b_3$ are real parameters, $A(x,t)$ and $B(x,t)$ are complex-valued functions of $(x,t) \in \mathbb{R}^2$ and $H(A, \overline{A}, B, \overline{B})$ is a real-valued function. For example, the $\chi^{(2)}$ coupled NLS model in optics is recovered by taking

$$H(A, \overline{A}, B, \overline{B}) = a_4|A|^2 + b_4|B|^2 + \mathrm{Re}(\alpha\overline{A}^2 B)\,, \tag{1.13}$$

where a_4, b_4 are real parameters and $\alpha \in \mathbb{C}$ (cf. [37,39] and references therein). In §11 it is shown that any system of the type (1.12) can be formulated as a Hamiltonian system on a multi-symplectic structure as in (1.2), with phase space $\mathbb{R}^8$.

We argue in this paper that multi-symplectification of NLS and coupled NLS models is a natural framework for analysis of the existence and linear stability of solitary wave and front solutions because it defines both a spatial symplectic structure and a temporal symplectic structure a priori – on a finite-dimensional phase space – and therefore both geometric structures are available throughout the analysis. To demonstrate how the theory of this paper applies to such models we consider in detail a multi-symplectic analysis of the stability of an oscillating front solution of the uncoupled NLS model in §11 and then sketch how the theory applied to other NLS models, as well as coupled models such as the $\chi^{(2)}$ and $\chi^{(3)}$ models in optics.

In §12 the theoretical framework is applied to a complex Klein-Gordon example. This complex nonlinear wave equation is a fundamental model equation in atmospheric dynamics [41] and for the nonlinear Kelvin-Helmholz instability [7,44]. An interesting property of this equation is that it has explicit oscillating front solutions that are biasymptotic to an invariant manifold $\mathcal{M}_\infty$ which is a periodic orbit in the phase space representation, and this periodic orbit has multiple bifurcations and stability changes and therefore provides a simple but non-trivial example of the geometric perturbation theory of §4.

2. The form of the basic solitary wave or front

In this section the basic properties of the solitary wave or front will be identified in preparation for the linear stability analysis. Here we will follow the description in §2 of [16], while also introducing a larger class of states at infinity. The starting point for the characterization of the wave is the abstract multi-symplectic formulation of the PDE,

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z)\,, \quad Z \in \mathbb{R}^{2n}\,, \quad x \in \mathbb{R}\,. \tag{2.1}$$

The basic hypotheses on the system (2.1) are that $\mathbf{M}$ and $\mathbf{K}$ are constant skew-symmetric $2n\times 2n$ matrices and $S : \mathbb{R}^{2n} \to \mathbb{R}$ is a smooth function. The gradient of S in (2.1) is with respect to an inner product on $\mathbb{R}^{2n}$, which will be denoted by $\langle\cdot,\cdot\rangle$ throughout.

Associated with $\mathbf{M}$ and $\mathbf{K}$ are the two-forms:

$$\omega(\zeta_1,\zeta_2) = \langle \mathbf{M}\zeta_1, \zeta_2\rangle\,, \quad \kappa(\zeta_1,\zeta_2) = \langle \mathbf{K}\zeta_1, \zeta_2\rangle\,, \tag{2.2}$$

and

$$\Omega(\zeta_1,\zeta_2) = \langle \mathbf{J}_c\zeta_1, \zeta_2\rangle\,, \quad \text{with} \quad \mathbf{J}_c = \mathbf{K} - c\mathbf{M}\,, \tag{2.3}$$

where ζ_1 and ζ_2 are any vectors in $\mathbb{R}^{2n}$. The two-forms ω, κ and Ω are constant and therefore closed. Neither ω or κ is required to be non-degenerate, but we take as a hypothesis that $\det \mathbf{J}_c \neq 0$ when $c \in \mathcal{C}$ (where $\mathcal{C} \subset \mathbb{R}$ is an open set) and therefore $(\Omega, \mathbb{R}^{2n})$ is a symplectic manifold. It is the natural symplectic structure associated with travelling wave states of (2.1).

2.1. The manifold at infinity

The system (2.1) is assumed to be $\mathcal{G}$-equivariant where $\mathcal{G}$ is a q-dimensional abelian subgroup of the Euclidean group, acting on $\mathbb{R}^{2n}$. The complete symmetry group of (2.1) may be larger (see §10 and §12 for examples), but it is this q-dimensional subgroup which will be used to define the *manifold at infinity.* The Lie algebra of the group $\mathcal{G}$ is

$$\mathfrak{g} = \operatorname{span}\{\xi_1, \ldots, \xi_q\},$$

and the actions of the related one-parameter subgroups of $\mathcal{G}$ are denoted by $\exp(\varepsilon\xi_j)$, $j = 1, \ldots, q$. Using multi-symplectic Noether theory (cf. [13], Appendix C), it follows that, for each such one-parameter subgroup, there is a pair of functionals (P_i, Q_i) satisfying

$$\mathbf{M}\xi_i(Z) = \nabla P_i(Z) \quad \text{and} \quad \mathbf{K}\xi_i(Z) = \nabla Q_i(Z), \quad i = 1, \ldots, q. \tag{2.4}$$

A generalization of relative equilibria, associated with the group $\mathcal{G}$, will be used to define the solution on the manifold at infinity. Classically, a relative equilibrium of a Hamiltonian system is a solution aligned with a group orbit travelling at constant speed, i.e., a solution of the form $\exp((a_1\xi_1 + \cdots + a_q\xi_q)t)\widehat{Z}(x)$ where $a_1, \ldots, a_q$ are real parameters and $\widehat{Z}(x)$ a wave shape in $\mathbb{R}^n$ (cf. [1], Chapt. 4). With this form for the relative equilibria, and if $\lim_{x\to\pm\infty}\widehat{Z}(x) = Z_0^\pm$ exists, the state at infinity is of the form $\exp((a_1\xi_1 + \cdots + a_q\xi_q)t)Z_0^\pm$.

We generalize this by allowing for both space and time variation along the group orbit (cf. [15]). The states at infinity are taken to be of the form

$$Z(x,t) = G_{\theta(x,t;\mathbf{a},\mathbf{b})}Z_0(\mathbf{a},\mathbf{b}), \tag{2.5}$$

where

$$\begin{aligned} G_{\theta(x,t;\mathbf{a},\mathbf{b})} &= \exp\{(a_1\xi_1 + \cdots + a_q\xi_q)t + (b_1\xi_1 + \cdots + b_q\xi_q)x\} \\ &= \exp\{\theta_1(x,t;\mathbf{a},\mathbf{b})\xi_1 + \cdots + \theta_q(x,t;\mathbf{a},\mathbf{b})\xi_q\}, \end{aligned} \tag{2.6}$$

with $\theta_j(x,t;\mathbf{a},\mathbf{b}) = a_jt + b_jx$, $j = 1, \ldots, q$ and

$$\mathbf{a} = (a_1, \ldots, a_q) \in \mathrm{A} \subset \mathbb{R}^q \quad \text{and} \quad \mathbf{b} = (b_1, \ldots, b_q) \in \mathrm{B} \subset \mathbb{R}^q. \tag{2.7}$$

Substitution of the expression (2.5) into (2.1) shows that $Z_0 \in \mathbb{R}^{2n}$ satisfies (cf. [16], §2)

$$\nabla S(Z_0) = \sum_{i=1}^q a_i\nabla P_i(Z_0) + \sum_{i=1}^q b_i\nabla Q_i(Z_0), \tag{2.8}$$

where $P_i(Z_0) = \mathcal{P}_i$ and $Q_i(Z_0) = \mathcal{Q}_i$ for $i = 1, \ldots, q$ and $(\mathcal{P}_1, \ldots, \mathcal{P}_q, \mathcal{Q}_1, \ldots, \mathcal{Q}_q) \in \mathbb{R}^{2q}$ are specified real parameters.

Let $Z_0(\mathbf{a},\mathbf{b}) \in \mathbb{R}^{2n}$ be any solution of (2.8), then *the invariant manifold at infinity* is

$$\mathcal{M}_\infty(\mathbf{a},\mathbf{b}) = \text{closure}\{G_\theta Z_0(\mathbf{a},\mathbf{b}) \mid G_\theta \in \mathcal{G}\}.$$

It is assumed that all solutions of (2.8) are elements of the manifold $\mathcal{M}_\infty(\mathbf{a},\mathbf{b})$. When the isotropy subgroup of Z_0 (i.e., $\{G_\theta \in \mathcal{G} \mid G_\theta Z_0 = Z_0\}$) is discrete and the subgroup $\mathcal{G}$ is compact, then $\mathcal{M}_\infty = \mathbb{T}^q$, and when the subgroup is affine, then $\mathcal{M}_\infty = \mathbb{R}^q$. Associated with the invariant manifold $\mathcal{M}_\infty(\mathbf{a},\mathbf{b})$, is a $2q$-parameter family of solutions (2.5) satisfying (2.1).

The manifolds at infinity are assumed to be nondegenerate. This means that

$$\det\begin{pmatrix} \mathcal{P}_\mathbf{a} & \mathcal{P}_\mathbf{b} \\ \mathcal{Q}_\mathbf{a} & \mathcal{Q}_\mathbf{b} \end{pmatrix} \neq 0, \tag{2.9}$$

where $\mathcal{P}_\mathbf{a}$ denotes the $q \times q$ matrix with entries $\frac{\partial}{\partial a_j}P_i(Z_0(\mathbf{a},\mathbf{b}))$, with similar notation for $\mathcal{P}_\mathbf{b}$, $\mathcal{Q}_\mathbf{a}$ and $\mathcal{Q}_\mathbf{b}$.

Remark 1. We do not consider the case where there are distinct invariant manifolds at plus and minus infinity, for example $\mathcal{M}_{+\infty}(\mathbf{a}^+,\mathbf{b}^+) \neq \mathcal{M}_{-\infty}(\mathbf{a}^-,\mathbf{b}^-)$ although this case is possible and certainly interesting.

2.2. The shape of the solitary wave or front

The complete solitary wave or front is taken to be of the following form

$$Z(x,t) = G_{\theta(x,t;\mathbf{a},\mathbf{b})}\,\mathcal{T}_{\tau(t)}\,\widetilde{Z}(x;\mathbf{a},\mathbf{b},c) \tag{2.10}$$

where $G_{\theta(x,t;\mathbf{a},\mathbf{b})}$ is the action of $\mathcal{G}$ defined in (2.6), $\mathbf{a}$, $\mathbf{b}$ are defined in (2.7), $\mathcal{T}_\tau Z(x,\cdot) = Z(x-\tau,\cdot)$ is an action of the translation group $x \mapsto x-\tau$ and $\tau(t) = ct + \tau_0$; $\widetilde{Z}(x;\cdot)$ is the *shape* of the solitary wave or front. The shape satisfies the asymptotic limits

$$\lim_{x\to-\infty} \widetilde{Z}(x;\mathbf{a},\mathbf{b},c) = Z_0^-(\mathbf{a},\mathbf{b}) \quad \text{and} \quad \lim_{x\to+\infty} \widetilde{Z}(x;\mathbf{a},\mathbf{b},c) = Z_0^+(\mathbf{a},\mathbf{b},c)\,, \tag{2.11}$$

where $Z_0^\pm$ satisfy (2.8).

For definiteness, we have taken the asymptotic state $Z_0^-(\mathbf{a},\mathbf{b})$ to be independent of c. Since $Z_0^\pm$ both lie in $\mathcal{M}_\infty(\mathbf{a},\mathbf{b})$, there is a $G_{\gamma(c)} \in \mathcal{G}$ such that

$$Z_0^+(\mathbf{a},\mathbf{b},c) = G_{\gamma(c)} Z_0^-(\mathbf{a},\mathbf{b})\,. \tag{2.12}$$

When $\mathcal{M}_\infty$ is non-trivial, the shape is in general asymptotic to distinct points on $\mathcal{M}_\infty$ as x goes to plus and minus infinity.

To simplify notation, the $(\mathbf{a},\mathbf{b},c)$ parameter dependence is represented by $\mathbf{p} = (\mathbf{a},\mathbf{b},c)$.

The shape of the wave can also be written in the form

$$\widetilde{Z}(x;\mathbf{p}) = Z_0^-(\mathbf{a},\mathbf{b}) + \widehat{Z}(x;\mathbf{p})\,,$$

with

$$\lim_{x\to-\infty} \widehat{Z}(x;\mathbf{p}) = 0 \quad \text{and} \quad \lim_{x\to+\infty} \widehat{Z}(x;\mathbf{p}) = Z_0^+(\mathbf{p}) - Z_0^-(\mathbf{a},\mathbf{b})\,.$$

Remark 2. Since $G_{\theta(x,t;\mathbf{p})}$ and $\mathcal{T}_\tau$ do not commute in general, allowing for the state at infinity to move at speed c may lead to a different family of waves. However, the analysis can be modified to treat this case.

Using the group equivariance of the governing equation (2.1), it follows that the shape $\widetilde{Z}$ satisfies the following Euler-Lagrange equation

$$\mathbf{K}\widetilde{Z}_x - c\mathbf{M}\widetilde{Z}_x = \nabla S(\widetilde{Z}) - \sum_{i=1}^q a_i \nabla P_i(\widetilde{Z}) - \sum_{i=1}^q b_i \nabla Q_i(\widetilde{Z})\,. \tag{2.13}$$

This equation can be simplified by renaming the right-hand side,

$$\mathbf{J}_c \widetilde{Z}_x = \nabla V(\widetilde{Z})\,, \quad \widetilde{Z} \in \mathbb{R}^{2n} \tag{2.14}$$

where

$$V(\widetilde{Z}) = S(\widetilde{Z}) - S(Z_0^-) - \sum_{i=1}^q a_i[P_i(\widetilde{Z}) - P_i(Z_0^-)] - \sum_{i=1}^q b_i[Q_i(\widetilde{Z}) - Q_i(Z_0^-)]\,.$$

Note that $\lim_{x\to\pm\infty} V(\widetilde{Z}) = 0$. For $x \to -\infty$ this limit is clear since $\widetilde{Z} \to Z_0^-$. As $x \to +\infty$, the limit follows from $\lim_{x\to+\infty} \widetilde{Z} = Z_0^+ = G_{\gamma(c)} Z_0^-$ and the fact that the functions S, P_i and Q_i are $\mathcal{G}$-invariant.

The characterization of the shape of the wave as given by (2.14), is a heteroclinic orbit, bi-asymptotic to $\mathcal{M}_\infty$, of the spatial Hamiltonian system $(\mathbb{R}^{2n}, \Omega, V)$. A second viewpoint is to characterize the shape as a critical point of the energy functional restricted to level sets of the momentum. To see this, rewrite (2.14) as

$$\nabla V(\widetilde{Z}) - \mathbf{K}\widetilde{Z}_x = -c\mathbf{M}\widetilde{Z}_x\,,$$

and define the energy respectively momentum functionals (assuming they exist)

$$\mathcal{H}(\widetilde{Z}) = \frac{1}{2}\int_{-\infty}^{+\infty} [\,\kappa(\widetilde{Z},\widetilde{Z}_x) + 2V(\widetilde{Z})\,]\,\mathrm{d}x \quad \text{and} \quad \mathcal{I}(\widetilde{Z}) = \frac{1}{2}\int_{-\infty}^{+\infty} \omega(\widetilde{Z},\widetilde{Z}_x)\,\mathrm{d}x\,.$$

Then (2.13) is formally recovered by

$$\frac{\delta}{\delta\widetilde{Z}}\mathcal{H}(\widetilde{Z}) = c\frac{\delta}{\delta\widetilde{Z}}\mathcal{I}(\widetilde{Z})\,,$$

which suggests that the shape of the wave can be characterized as a critical point of $\mathcal{H}$ restricted to level sets of the momentum. However this variational principle is not always straightforward to work with; in particular, the integral of $V(\widetilde{Z})$ may be unbounded from above and below. On the other hand, if we assume that $\widetilde{Z}_x$ is smooth and exponentially decaying at $\pm\infty$, then the other integrals are convergent. Although we will not use this variational principle directly, the momentum and its dependence on c will feature prominently in the sequel.

2.3. *The momentum of the wave*

The momentum density is $I(Z) = \frac{1}{2}\omega(Z,Z_x)$ but there are three candidates for the argument of this density: Z, $\widetilde{Z}$ and $\widehat{Z}$. However, since Z, $\widetilde{Z}$ and $\widehat{Z}$ are related, these densities are related.

Proposition 1.

$$I(\widetilde{Z}) = I(\widehat{Z}) + \tfrac{1}{2}\omega(Z_0^-,Z_0^+) \quad \text{and} \quad I(Z) = \mathcal{T}_{\tau(t)}\,[\,I(\widetilde{Z}) - \tfrac{1}{2}\textstyle\sum_{j=1}^q b_j\langle\widetilde{Z},\nabla P_j(\widetilde{Z})\rangle\,]\,. \tag{2.15}$$

Proof. The first expression is obtained by direct substitution. The second expression follows by direct substitution and use of the $\mathcal{G}$-invariance:

$$I(Z) = I(G_{\theta(x,t)}\mathcal{T}_{\tau(t)}\,\widetilde{Z}) = \tfrac{1}{2}\mathcal{T}_{\tau(t)}\,\omega(\widetilde{Z},\widetilde{Z}_x) + \tfrac{1}{2}\omega(G_\theta\mathcal{T}_{\tau(t)}\,\widetilde{Z},\big(\tfrac{\partial}{\partial x}G_{\theta(x,t)}\big)\,\mathcal{T}_{\tau(t)}\,\widetilde{Z})\,.$$

But $\mathbf{M}\big(\frac{\partial}{\partial x}G_{\theta(x,t)}\big)\,\mathcal{T}_{\tau(t)}\,\widetilde{Z} = \mathbf{M}G_\theta\sum_{j=1}^q b_j\xi_j(\mathcal{T}_{\tau(t)}\,\widetilde{Z}) = G_\theta\mathcal{T}_{\tau(t)}\,\sum_{j=1}^q b_j\nabla P_j(\widetilde{Z})$ and therefore

$$\begin{aligned} I(Z) &= \mathcal{T}_{\tau(t)}\,I(\widetilde{Z}) + \tfrac{1}{2}\langle\mathbf{M}G_\theta\mathcal{T}_{\tau(t)}\,\widetilde{Z},\big(\tfrac{\partial}{\partial x}G_{\theta(x,t)}\big)\,\mathcal{T}_{\tau(t)}\,\widetilde{Z}\rangle \\ &= \mathcal{T}_{\tau(t)}\,[\,I(\widetilde{Z}) - \tfrac{1}{2}\textstyle\sum_{j=1}^q b_j\langle\widetilde{Z},\nabla P_j(\widetilde{Z})\rangle\,]\,. \end{aligned}$$

The x-integral over $\mathbb{R}$ of $I(Z)$ will not exist in general, whereas both the integral of $I(\widetilde{Z})$ and $I(\widehat{Z})$ will exist if $\widetilde{Z}_x = \widehat{Z}_x$ is exponentially decaying. However, it is the density $I(\widetilde{Z})$ which appears in the stability analysis. In other words, the momentum evaluated on the shape of the solitary wave or front (and its derivative with respect to c) is the important quantity for the stability analysis. In [16] it is shown that the momentum density can be interpreted as a one-form on the heteroclinic-orbit representation on the phase space. A consequence of this characterization is the following result, (cf. Lemma 2.1 of [16]).

Lemma 1. *Let $\widetilde{Z}(x;\mathbf{p})$ be the shape of the solitary wave and suppose $\widetilde{Z}_x$ decays exponentially as $x\to\pm\infty$; then,*

$$\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}) = \int_{-\infty}^{+\infty}\omega(\widetilde{Z}_c,\widetilde{Z}_x)\,\mathrm{d}x + \tfrac{1}{2}\omega(Z_0^+,\partial_c Z_0^+)\,. \tag{2.16}$$

2.4. Summary of hypotheses

We summarize and formalize the properties of the basic travelling wave in the following.

Hypothesis *We assume that*

1. *the PDE has a multi-symplectic formulation* (2.1) *on* $\mathbb{R}^{2n}$ *with* $\mathbf{J}_c$ *non-degenerate and* $S:\mathbb{R}^{2n}\to\mathbb{R}$ *smooth (at least twice continuously differentiable). The system* (2.1) *is* $\mathcal{G}$*-equivariant with* $\mathcal{G}$ *a* q*-dimensional Abelian subgroup of the Euclidean group on* $\mathbb{R}^{2n}$*;*
2. *there are open sets* $\mathrm{A},\mathrm{B}\in\mathbb{R}^q$, $\mathcal{C}\subset\mathbb{R}$ *such that for every* $\mathbf{a}=(a_1,\dots,a_q)\in\mathrm{A}$, $\mathbf{b}=(b_1,\dots,b_q)\in\mathrm{B}$, *and* $c\in\mathcal{C}$ *there exists a travelling wave shape* $\widetilde{Z}(x;\mathbf{p})$ *satisfying* (2.13) *with* $\mathbf{p}=(\mathbf{a},\mathbf{b},c)$*;*
3. *the shape of the travelling wave is bounded for all* $x\in\mathbb{R}$ *and the shape is asymptotic to points* $Z_0^\pm$ *on* $\mathcal{M}_\infty$ *satisfying* (2.8) *and defined by*
$$Z_0^-(\mathbf{a},\mathbf{b})=\lim_{x\to-\infty}\widetilde{Z}(x;\mathbf{p}),\quad\text{and}\quad Z_0^+(\mathbf{p})=\lim_{x\to+\infty}\widetilde{Z}(x;\mathbf{p}).$$
This implies that there exists a $\gamma(c)$ *such that* $Z_0^+(\mathbf{p})=G_{\gamma(c)}Z_0^-(\mathbf{a},\mathbf{b})$*;*
4. *the manifold at infinity* $\mathcal{M}_\infty(\mathbf{a},\mathbf{b})$ *is nondegenerate: it satisfies* (2.9)*;*
5. *the derivative of the shape of the travelling wave,* $\widetilde{Z}_x$*, is exponentially decaying, with asymptotic estimate: for some* $\beta(\mathbf{p})>0$*,*
$$\lim_{x\to\pm\infty}e^{\pm\beta(\mathbf{p})x}\widetilde{Z}_x(x;\mathbf{p})=\Psi^\pm(\mathbf{p})\quad\text{and}\quad\lim_{x\to\pm\infty}\partial_x[e^{\pm\beta(\mathbf{p})x}\widetilde{Z}_x(x;\mathbf{p})]=0.$$

3. Linearization about the solitary wave or front

To investigate the linear stability of a family of travelling solitary waves or fronts with shape $\widetilde{Z}(x;\mathbf{p})$, a solution $Z(x,t)$ of the time-dependent PDE (2.1) is expressed as

$$\begin{aligned}Z(x,t)&=G_{\theta(x,t;\mathbf{p})}T_{\tau(t)}\,[\widetilde{Z}(x;\mathbf{p})+\widetilde{U}(x,t;\mathbf{p})]\\&=G_{\theta(x,t;\mathbf{p})}\,[Z_0^-(\mathbf{a},\mathbf{b})+T_{\tau(t)}(\widehat{Z}(x;\mathbf{p})+\widetilde{U}(x,t;\mathbf{p}))],\end{aligned}$$

where $\theta(x,t;\mathbf{p})$ is a linear function of x and t as in (2.6) and $\tau(t)=ct+\tau_0$. The linearized equation for $\widetilde{U}(x,t;\mathbf{p})$ is then

$$(\mathbf{K}-c\mathbf{M})\,\widetilde{U}_x+\mathbf{M}\widetilde{U}_t=\mathbf{B}(x;\mathbf{p})\widetilde{U}$$

where $\mathbf{B}(x;\mathbf{p})$ is the symmetric matrix

$$\mathbf{B}(x;\mathbf{p})=\left[D^2S-\sum_{j=1}^{q}(a_jD^2P_j+b_jD^2Q_j)\right](\widetilde{Z}(x;\mathbf{p}))\,.$$

With a spectral ansatz for $\widetilde{U}$: writing $\widetilde{U}(x,t;\mathbf{p})=e^{\lambda t}U(x,\lambda;\mathbf{p})$, the function $U(x,\lambda;\mathbf{p})$ satisfies the ODE

$$\mathbf{J}_cU_x=[\mathbf{B}(x;\mathbf{p})-\lambda\mathbf{M}]U\,,\tag{3.1}$$

or

$$U_x=\mathbf{A}(x,\lambda;\mathbf{p})U\,,\quad U\in\mathbb{C}^{2n}\tag{3.2}$$

with

$$\mathbf{A}(x,\lambda;\mathbf{p})=\mathbf{J}_c^{-1}[\mathbf{B}(x;\mathbf{p})-\lambda\mathbf{M}]\,.\tag{3.3}$$

Throughout, we will assume that $\lambda\in\Lambda$, where Λ is an open simply-connected subset of $\mathbb{C}_+$, the right-half complex λ-plane, with $0\in\overline{\Lambda}$. A second hypothesis on Λ will be given below.

The analysis of the ODE (3.2) is the basis of the linear stability analysis. A precise definition of stability will be given in §5. The basic idea is that if, for some fixed $\mathbf{p}=(\mathbf{a},\mathbf{b},c)\in\mathrm{A}\times\mathrm{B}\times\mathcal{C}$, there exists a $\lambda\in\Lambda$ with positive real part for which (3.2) has an admissible solution, then the solitary wave or front is said to be linearly unstable.

3.1. The systems at infinity

The systems at infinity are the systems of constant coefficient ordinary differential equations (3.2) in the limit at $x \to \pm\infty$. By hypothesis, the shape of the wave $\widetilde{Z}(x;\mathbf{p})$ converges and therefore the limits

$$\mathbf{B}^-(\mathbf{a},\mathbf{b}) = \lim_{x\to-\infty} \mathbf{B}(x;\mathbf{p}) = [D^2S + \sum_{j=1}^{q}(a_jD^2P + b_jD^2Q)](Z_0^-(\mathbf{a},\mathbf{b}))\,, \tag{3.4}$$

$$\mathbf{B}^+(\mathbf{p}) = \lim_{x\to+\infty} \mathbf{B}(x;\mathbf{p}) = [D^2S + \sum_{j=1}^{q}(a_jD^2P + b_jD^2Q)](Z_0^+(\mathbf{p})) \tag{3.5}$$

exist. The fact that $\mathbf{B}^-(\mathbf{a},\mathbf{b})$ is independent of c is a reflection of our choice for $Z_0^-(\mathbf{a},\mathbf{b})$ to be independent of c. The matrices $\mathbf{B}^+(\mathbf{p})$ and $\mathbf{B}^-(\mathbf{a},\mathbf{b})$ are not equal in general, even if $c=0$; but a consequence of the relation $Z_0^+ = G_\gamma Z_0^-$ is the property (cf. [16])

$$\mathbf{B}^+(\mathbf{p}) = DG_{\gamma(c)}(Z_0^-(\mathbf{a},\mathbf{b}))^{-T}\mathbf{B}^-(\mathbf{a},\mathbf{b})DG_{\gamma(c)}(Z_0^-(\mathbf{a},\mathbf{b}))^{-1}\,,$$

or, since the action $G_{\gamma(c)}$ is either compact or an affine translation,

$$\mathbf{B}^+(\mathbf{p}) = DG_{\gamma(c)}(Z_0^-(\mathbf{a},\mathbf{b}))\mathbf{B}^-(\mathbf{a},\mathbf{b})DG_{\gamma(c)}(Z_0^-(\mathbf{a},\mathbf{b}))^T\,. \tag{3.6}$$

Define

$$\mathbf{A}^\pm(\lambda;\mathbf{p}) = \lim_{x\to\pm\infty}\mathbf{A}(x,\lambda;\mathbf{p}) = \mathbf{J}_c^{-1}(\mathbf{B}^\pm(\mathbf{p}) - \lambda\mathbf{M})\,. \tag{3.7}$$

Proposition 2. $\mathbf{A}^+(\lambda;\mathbf{p}) \neq \mathbf{A}^-(\lambda;\mathbf{p})$ *in general, but the spectrum of the two matrices is the same. Moreover, if* $\zeta^-(\lambda;\mathbf{p})$ *is an eigenvector of* $\mathbf{A}^-(\lambda;\mathbf{p})$ *with eigenvalue* μ*, then*

$$\zeta^+(\lambda;\mathbf{p}) = DG_\gamma(Z_0^-(\mathbf{p}))\zeta^-(\lambda;\mathbf{p})$$

is an eigenvector of $\mathbf{A}^+(\lambda;\mathbf{p})$ *with the same eigenvalue.*

Proof. In (3.6) it is established that $\mathbf{B}^+$ and $\mathbf{B}^-$ are similar matrices. Since both $\mathbf{M}$ and $\mathbf{J}_c$ commute with DG_γ, this implies $\mathbf{A}^+ = DG_\gamma\mathbf{A}^-DG_\gamma^T$. Hence $\mathbf{A}^-$ and $\mathbf{A}^+$ are similar matrices. To prove the relation satisfied by the eigenvectors, use the fact that $\mathbf{A}^-\zeta^- = \mu\zeta^-$ implies $\mathbf{A}^+DG_\gamma\zeta^- = \mu DG_\gamma\zeta^-$.

Hypothesis 5 in §2.4 on the exponential decay of $\tilde{Z}_x$ implies that $\pm\beta(\mathbf{p})$ are eigenvalues of $\mathbf{A}^\pm(0;\mathbf{p})$.

Proposition 3. *The vector* $\Psi^\pm(\mathbf{p})$ *is an eigenvector of* $\mathbf{A}^\pm(0;\mathbf{p})$ *with eigenvalue* $\mp\beta(\mathbf{p})$*. Hence* $\pm\beta(\mathbf{p})$ *are also eigenvalues of* $\mathbf{A}^\pm(0;\mathbf{p})$*, with* $DG_\gamma(Z_0^-(\mathbf{a},\mathbf{b})^T\Psi^+(\mathbf{p})$ *the eigenvector of* $\mathbf{A}^-(0;\mathbf{p})$ *with eigenvalue* $-\beta(\mathbf{p})$*, and* $DG_\gamma(Z_0^-(\mathbf{a},\mathbf{b}))\Psi^-(\mathbf{p})$ *is the eigenvector of* $\mathbf{A}^+(0;\mathbf{p})$ *with eigenvalue* $\beta(\mathbf{p})$*.*

Proof. By differentiating the Euler-Lagrange equation (2.13) with respect to x, we get

$$\mathbf{B}(x;\mathbf{p})\widetilde{Z}_x(x;\mathbf{p}) = \mathbf{J}_c\widetilde{Z}_{xx}(x;\mathbf{p}).$$

Multiplying this equation with $e^{\pm\beta(\mathbf{p})x}$ and rewriting it, gives

$$\mathbf{B}(x;\mathbf{p})e^{\pm\beta(\mathbf{p})x}\widetilde{Z}_x(x;\mathbf{p}) = \mathbf{J}_c\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\pm\beta(\mathbf{p})x}\widetilde{Z}_x(x;\mathbf{p})\right) \mp \beta(\mathbf{p})\mathbf{J}_ce^{\pm\beta(\mathbf{p})x}\widetilde{Z}_x(x;\mathbf{p}).$$

From hypothesis 5 in §2.4 it follows that $\lim_{x\to\pm\infty}\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\pm\beta(\mathbf{p})x}\widetilde{Z}_x(x;\mathbf{p})\right) = 0$. Hence taking the limit for $x\to\pm\infty$ in the equation above gives $\mathbf{B}^\pm\Psi^\pm = \mp\beta\mathbf{J}_c\Psi^\pm$, in other words,

$$\mathbf{A}^\pm(0;\mathbf{p})\Psi^\pm(\mathbf{p}) = \mp\beta(\mathbf{p})\Psi^\pm(\mathbf{p}).$$

Since $\mathbf{A}^\pm(0;\mathbf{p}) = \mathbf{J}_c^{-1}\mathbf{B}^\pm(\mathbf{p})$, with $\mathbf{J}_c^{-1}$ skew-symmetric and $\mathbf{B}^\pm$ symmetric, we get immediately that $\pm\beta(\mathbf{p})$ are also eigenvalues of $\mathbf{A}^\pm(0;\mathbf{p})$.

The other statements in Proposition 3 are consequences of Proposition 2.

The spectral problem for $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ can be written $\det(\mathbf{A}^{\pm}(\lambda;\mathbf{p})-\mu\mathbf{I})=0$, or

$$\Delta(\mu,\lambda;\mathbf{p})=0\,,\quad \text{with}\quad \Delta(\mu,\lambda;\mathbf{p})=\det[\mathbf{B}^{\pm}(\mathbf{p})-\mu\mathbf{J}_c-\lambda\mathbf{M}]\,. \tag{3.8}$$

A superscript $\pm$ is not needed on Δ because the spectrum of $\mathbf{A}^{+}(\lambda;\mathbf{p})$ and $\mathbf{A}^{-}(\lambda;\mathbf{p})$ are the same.

An interesting property of the spectrum of $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ is that $-\mu$ is an eigenvalue of $\mathbf{A}^{\pm}(-\lambda;\mathbf{p})$ whenever μ is an eigenvalue of $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$. This follows from the multi-symplectic reflection property:

$$\begin{aligned}\Delta(\mu,\lambda;\mathbf{p}) &= \det[\mathbf{B}^{\pm}(\mathbf{p})-\mu\mathbf{J}_c-\lambda\mathbf{M}]\\ &= \det[\mathbf{B}^{\pm}(\mathbf{p})^T-\mu\mathbf{J}_c^T-\lambda\mathbf{M}^T]=\det[\mathbf{B}^{\pm}(\mathbf{p})+\mu\mathbf{J}_c+\lambda\mathbf{M}]=\Delta(-\mu,-\lambda;\mathbf{p})\,.\end{aligned}$$

We will assume that the spectrum of $\mathbf{A}^{\pm}(0;\mathbf{p})$ has exactly one positive real simple eigenvalue, $\beta(\mathbf{p})$; one negative real simple eigenvalue, $-\beta(\mathbf{p})$, and the remainder of the spectrum is purely imaginary. When $\lambda\in\Lambda$ (i.e., $\lambda\neq 0$) we assume the spectrum of $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ consists of $p=q+1$ eigenvalues with strictly negative real part, and the remainder of the spectrum has non-negative real part. The p μ-roots of $\Delta(\mu,\lambda;\mathbf{p})$ with negative real part are denoted by $\mu_j(\lambda;\mathbf{p})$ for $j=1,\ldots,p$ and ordered so that $\Re(\mu_1(\lambda;\mathbf{p}))\le\Re(\mu_2(\lambda;\mathbf{p}))\le\ldots\le\Re(\mu_p(\lambda;\mathbf{p}))<0$ and

$$\lim_{\lambda\to 0}\mu_1(\lambda;\mathbf{p})=-\beta(\mathbf{p})\,,\quad\text{and}\quad \lim_{\lambda\to 0}\mu_j(\lambda;\mathbf{p})=0\,,\quad \text{for } j=2,\ldots,p\,. \tag{3.9}$$

The eigenvalues $\mu_2(\lambda;\mathbf{p}),\ldots,\mu_p(\lambda;\mathbf{p})$ are perturbations of the eigenvalue 0, and we will assume that all those eigenvalues 0 are associated with symmetry (we will specify this assumption later in this section). In the next section (§4), we develop a geometric perturbation theory which predicts the small λ behaviour of $\mu_2(\lambda;\mathbf{p}),\ldots,\mu_p(\lambda;\mathbf{p})$ and we show that the behaviour of these eigenvalues near $\lambda=0$ can be expressed in terms of polynomials in λ.

The second hypothesis on the set Λ is that each of the eigenvalues $\mu_1(\lambda;\mathbf{p}),\ldots,\mu_p(\lambda;\mathbf{p})$ is simple for all $\lambda\in\Lambda$. This hypothesis ensures that each eigenvalue depends analytically on λ. A further discussion of analyticity of the eigenvalues and its implications is given in §9. Note that we do not require simplicity of the other eigenvalues in the spectrum of $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ (i.e., the ones with non-negative real part).

The eigenvectors associated with the eigenvalues $\mu_i(\lambda;\mathbf{p})$ of $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ are denoted by $\zeta_i^{\pm}(\lambda;\mathbf{p})$ and they satisfy

$$[\mathbf{B}^{\pm}(\mathbf{p})-\mu_i(\lambda;\mathbf{p})\mathbf{J}_c-\lambda\mathbf{M}]\zeta_i^{\pm}(\lambda;\mathbf{p})=0\,,\quad i=1,\ldots,p\,. \tag{3.10}$$

Since μ_i, $i=1,\ldots,p$ are simple eigenvalues, the eigenvectors $\zeta_i^{\pm}(\lambda;\mathbf{p})$ can be chosen to be analytic functions of λ for all $\lambda\in\Lambda$ (cf. Kato [31]).

By the chosen ordering of the eigenvalues (cf. (3.9)) and Proposition 3, it follows that

$$\lim_{\lambda\to 0}\zeta_1^{+}(\lambda;\mathbf{p})=\Psi^{+}(\mathbf{p})\,.$$

In §4, it will be shown that if $\mathbf{B}^{+}$ has exactly q eigenvalues 0 and the manifold at infinity is non-degenerate (i.e., (2.9) is satisfied), then for $k=2,\ldots,p$ the eigenvalues $\mu_k(\lambda)$ have the following asymptotic expansion $\mu_k(\lambda)=\mu_k^1\lambda+\mathcal{O}(\lambda^2)$, for λ near 0. By hypothesis, $\Re(\mu_k^1)<0$. Furthermore, it will be proved that the eigenvectors $\zeta_k^{+}(\lambda)$ have the following asymptotic expansion

$$\zeta_k^{+}(\lambda)=\mathbb{C}\sum_{j=1}^{q}\chi_{jk}^{+}\left(\xi_k(Z_0^{+})+\left[\mu_j^1\frac{\partial Z_0^{+}}{\partial b_k}+(1-c\mu_k^1)\frac{\partial Z_0^{+}}{\partial a_k}\right]\lambda+\mathcal{O}(\lambda^2)\right),$$

where χ_{jk}^{+} are constants. Hence, it is natural to assume that in the limit as $\lambda\to 0$, the eigenvectors $\zeta_2^{+}(\lambda;\mathbf{p}),\ldots,\zeta_p^{+}(\lambda;\mathbf{p})$ are related to the symmetry. In other words we assume that the two sets of bases

$$\mathrm{span}\{\zeta_2^{+}(0;\mathbf{p}),\cdots,\zeta_p^{+}(0;\mathbf{p})\}\quad\text{and}\quad \mathrm{span}\{\xi_1(Z_0^{+}(\mathbf{p})),\cdots,\xi_q(Z_0^{+}(\mathbf{p}))\}$$

span the same vector space, where $\zeta_j^+(0;\mathbf{p}) = \lim_{\lambda\to 0}\zeta_j^+(\lambda;\mathbf{p})$, $j=2,\dots,p$. Therefore, there exists an invertible $q\times q$ matrix $\boldsymbol{\chi}^+(\mathbf{p})$ such that

$$[\zeta_2^+(0;\mathbf{p}),\cdots,\zeta_p^+(0;\mathbf{p})] = [\xi_1(Z_0^+(\mathbf{p})),\cdots,\xi_q(Z_0^+(\mathbf{p}))]\boldsymbol{\chi}^+(\mathbf{p})\,. \tag{3.11}$$

3.2. Adjoint eigenvectors and meromorphicity

The adjoint eigenvectors $v^\pm(\lambda;\mathbf{p})$ of $\mathbf{A}^\pm(\lambda;\mathbf{p})$ corresponding to an eigenvalue $\mu(\lambda;\mathbf{p})\in\mathbb{C}$ satisfy

$$\mathbf{A}^\pm(\lambda;\mathbf{p})^*\, v^\pm(\lambda;\mathbf{p}) = \overline{\mu}(\lambda;\mathbf{p})v^\pm(\lambda;\mathbf{p})\,,$$

where the superscript $*$ indicates complex-conjugate transpose. Due to the multi-symplectic decomposition of $\mathbf{A}^\pm(\lambda;\mathbf{p})$, the vectors $v^\pm(\lambda;\mathbf{p})$ have an interesting and useful relationship with the vectors $\zeta^\pm(\lambda;\mathbf{p})$. Let

$$\eta^\pm(\lambda;\mathbf{p}) = \mathbf{J}_c^{-1}\overline{v^\pm(\lambda;\mathbf{p})}\,.$$

Then these vectors satisfy

$$[\mathbf{B}^\pm(\mathbf{p}) + \mu(\lambda;\mathbf{p})\mathbf{J}_c + \lambda\mathbf{M}]\eta^\pm(\lambda;\mathbf{p}) = 0\,. \tag{3.12}$$

In other words,

$$\mathbf{A}^\pm(-\lambda;\mathbf{p})\eta^\pm(\lambda;\mathbf{p}) = -\mu(\lambda;\mathbf{p})\eta^\pm(\lambda;\mathbf{p}). \tag{3.13}$$

For the simple eigenvalues $\mu_1(\lambda;\mathbf{p}),\dots,\mu_p(\lambda;\mathbf{p})$, define $\eta_j^\pm(\lambda;\mathbf{p})$ to be the eigenvector of $\mathbf{A}^\pm(-\lambda;\mathbf{p})$ with eigenvalue $-\mu_j(\lambda;\mathbf{p})$. The natural normalisation of the eigenvectors $\zeta_i^\pm$ and $\eta_j^\pm$ is the Ω-symplectic normalisation

$$\Omega(\eta_j^\pm,\zeta_i^\pm) = \begin{cases} 1, & \text{if } i=j \\ 0, & \text{otherwise}\end{cases} \quad \text{for} \quad 1\le i,j\le p\,.$$

In a similar way to Proposition 2, we can prove that

Proposition 4. $\eta_j^+(\lambda;\mathbf{p}) = DG_{\gamma(c)}(Z_0^-(\mathbf{p}))\eta_j^-(\lambda;\mathbf{p})$.

So the normalisation for η_j^- can be written as

$$\Omega(\eta_j^-, DG_\gamma(Z_0^-)^T\zeta_i^+) = \begin{cases} 1, & \text{if } i=j \\ 0, & \text{otherwise}\end{cases} \quad \text{for} \quad 1\le i,j\le p\,. \tag{3.14}$$

The eigenvector $\eta_1^-(\lambda;\mathbf{p})$ can be extended analytically at $\lambda=0$, since $\mu_1(\lambda;\mathbf{p})\to -\beta(\mathbf{p})$ as $\lambda\to 0$ and $-\beta(\mathbf{p})$ is a simple eigenvalue. At $\lambda=0$, $\eta_1^-(0;\mathbf{p})$ is an eigenvector of $\mathbf{A}^-(0;\mathbf{p})$ with eigenvalue $-\mu_1(0,\mathbf{p}) = \beta(\mathbf{p})$. From Proposition 3 it follows that $\eta_1^-(0,\mathbf{p})$ is a scalar multiple of $\Psi^-(\mathbf{p})$, i.e., $\eta_1^-(0,\mathbf{p}) = \chi_{00}^-(\mathbf{p})\Psi^-(\mathbf{p})$. Since $\zeta_1^+(0;\mathbf{p}) = \Psi^+(\mathbf{p})$, the eigenvector normalization $\Omega(\eta_1^-, DG_\gamma(Z_0^-)^T\zeta_1^+) = 1$ gives that the scalar $\chi_{00}^-(\mathbf{p})$ associated with $\eta_1^-(0;\mathbf{p})$ is

$$\begin{aligned}\chi_{00}^-(\mathbf{p}) &= \Omega(\Psi^-(\mathbf{p}), DG_\gamma(Z_0^-(\mathbf{p}))^T\Psi^+(\mathbf{p}))^{-1}\\ &= \lim_{x\to\infty}[\,e^{2\beta(\mathbf{p})x}\Omega(\widetilde{Z}_x(-x;\mathbf{p}), DG_\gamma(Z_0^-(\mathbf{p}))^T\widetilde{Z}_x(x;\mathbf{p}))\,]^{-1}\,.\end{aligned}$$

In other words, there are non-negative integers $m_2^-,\dots,m_p^-$ such that the vectors

$$\widetilde{\eta}_i^-(\lambda;\mathbf{p}) = \lambda^{m_i^-}\eta_i^-(\lambda;\mathbf{p})\,, \quad i=2,\dots,p$$

are differentiable in a neighbourhood of 0 in $\overline{\Lambda}$ (see KATO [31]).

We have seen that

$$\eta_1^-(\lambda;\mathbf{p}) \to \Omega(\Psi^-(\mathbf{p}), DG_\gamma(Z_0^-(\mathbf{p}))^T\Psi^+(\mathbf{p}))^{-1}\Psi^-(\mathbf{p}) = \chi_{00}^-(\mathbf{p})\Psi^-(\mathbf{p}), \quad \text{for} \quad \lambda \to 0\,.$$

Also, by taking the limit for $\lambda \to 0$, we get for $j = 2, \ldots, p$

$$\mathbf{A}^-(0;\mathbf{p})\zeta_j^-(0;\mathbf{p}) = \mu_j(0;\mathbf{p})\zeta_j^-(0;\mathbf{p}) = 0 = -\mu_j(0;\mathbf{p})\eta_j^-(0;\mathbf{p}) = \mathbf{A}^-(0;\mathbf{p})\eta_j^-(0;\mathbf{p}),$$

since $\mu_j(0,\mathbf{p}) = 0$. Therefore, it is natural to require that the same space is spanned by the two sets of bases $\{\eta_2^-(0,\mathbf{p}), \ldots, \eta_p^-(0,\mathbf{p})\}$ and $\{\zeta_2^-(0,\mathbf{p}), \ldots, \zeta_p^-(0,\mathbf{p})\}$; i.e., that there exists an invertible $q \times q$ matrix $\boldsymbol{\chi}^-(\mathbf{p})$ such that

$$[\widetilde{\eta}_2^-(0;\mathbf{p})|\cdots|\widetilde{\eta}_p^-(0;\mathbf{p})] = [\xi_1(Z_0^-)|\cdots|\xi_q(Z_0^-)]\boldsymbol{\chi}^-(\mathbf{p})\,, \tag{3.15}$$

where $\widetilde{\eta}_i^-(0;\mathbf{p}) = \lim\limits_{\lambda\to 0} \lambda^{m_i^-}\eta_i^-(\lambda;\mathbf{p})$, $i = 2, \ldots, p$ as defined above.

3.3. The stable subspace of $U_x = \mathbf{A}(x,\lambda;\mathbf{p})U$

With the above properties of the asymptotic constant coefficient systems at $x = \pm\infty$, we can use standard theory to establish properties of solutions of (3.2) for finite x. Hypothesis 5 of §2.4 ensures that the derivative of the solitary wave decays exponentially fast. This implies that the shape of the wave decays exponentially to the asymptotic states $Z_0^\pm$ and the matrices $\mathbf{A}(x;\mathbf{p})$ and $\mathbf{B}(x;\mathbf{p})$ will decay exponentially to the asymptotic matrices $\mathbf{A}^\pm(\mathbf{p})$ respectively $\mathbf{B}^\pm(\mathbf{p})$.

The stable subspace of (3.2) is the space of all solutions which decay exponentially to 0 as $x \to +\infty$. Since $\mathbf{A}^+(\lambda;\mathbf{p})$ has p eigenvalues with strictly negative real part, it follows from the standard theory of asymptotic systems (cf. [18], [20]) that there exists p linearly independent solutions of (3.2),

$$\{\ U_i^+(x,\lambda;\mathbf{p}) \mid i = 1, \ldots, p\ \} \tag{3.16}$$

with the property that

$$\lim_{x\to+\infty} e^{-\mu_i(\lambda;\mathbf{p})x}U_i^+(x,\lambda;\mathbf{p}) = \zeta_i^+(\lambda;\mathbf{p})\,, \quad i = 1, \ldots, p\,. \tag{3.17}$$

The analyticity of the functions $U_i^+(x,\lambda;\mathbf{p})$ does not follow from the classical Levinson theory. However, we can appeal to a result of PEGO & WEINSTEIN ([42], Proposition 1.2) to show that each $U_i^+(x,\lambda;\mathbf{p})$ is analytic for $\lambda \in \Lambda$, and that the limit is uniform on compact subsets of Λ.

Due to the ordering of the eigenvalues, it follows that

$$U_1^+(x,0;\mathbf{p}) := \lim_{\lambda\to 0} U_1^+(x,\lambda;\mathbf{p}) = \widetilde{Z}_x(x;\mathbf{p}) := \xi_0(\widetilde{Z}(x;\mathbf{p}))\,, \tag{3.18}$$

and (3.11) implies that

$$U_{i+1}^+(x,0;\mathbf{p}) := \lim_{\lambda\to 0} U_{i+1}^+(x,\lambda;\mathbf{p}) = \sum_{j=1}^{q} \chi_{ji}^+\xi_j(\widetilde{Z}(x;\mathbf{p}))\,, \quad \text{for} \quad i = 1, \ldots, q = p-1\,. \tag{3.19}$$

3.4. The unstable subspace of $W_x = \mathbf{A}(x, -\lambda; \mathbf{p})W$

The adjoint of (3.2) will also be used in the analysis. Classically, the adjoint problem associated with (3.2) is

$$\Upsilon_x = -\mathbf{A}(x, \lambda; \mathbf{p})^* \Upsilon \,. \tag{3.20}$$

However, the multi-symplectic decomposition of $\mathbf{A}(x, \lambda; \mathbf{p})$ leads to

$$\mathbf{A}(x, \lambda; \mathbf{p})^* = [\mathbf{B}(x; \mathbf{p}) - \overline{\lambda}\mathbf{M}^T](\mathbf{J}_c^{-1})^T = -[\mathbf{B}(x; \mathbf{p}) + \overline{\lambda}\mathbf{M}]\mathbf{J}_c^{-1} \,.$$

Hence, with

$$\Upsilon(x, \lambda; \mathbf{p}) = \mathbf{J}_c \overline{W(x, \lambda; \mathbf{p})} \,, \tag{3.21}$$

it follows that $W(x, \lambda; \mathbf{p})$ satisfies

$$\mathbf{J}_c W_x = [\mathbf{B}(x; \mathbf{p}) + \lambda \mathbf{M}]W \,, \tag{3.22}$$

or

$$W_x = \mathbf{A}(x, -\lambda; \mathbf{p})W \,, \quad W \in \mathbb{C}^{2n} \,. \tag{3.23}$$

A consequence of this form is that the two equations (3.2) and (3.23) are identical when $\lambda = 0$.

The constant coefficient systems at $x = \pm\infty$ associated with (3.23) are

$$W_x = \mathbf{A}^{\pm}(-\lambda; \mathbf{p})W \,, \quad W \in \mathbb{C}^{2n} \,, \tag{3.24}$$

where

$$\mathbf{A}^{\pm}(-\lambda; \mathbf{p}) = \lim_{x \to \pm\infty} \mathbf{A}(x, -\lambda; \mathbf{p}) \,. \tag{3.25}$$

Using the standard asymptotic theory, we can define solutions of the dual equation (3.23). Just as before, we can define p independent solutions $W_i^-(x, \lambda; \mathbf{p})$, $i = 1, \ldots, p$, $\lambda \neq 0$, for the dual system. Since $-\mu_j(\lambda; \mathbf{p})$, $j = 1, \ldots, p$ are the adjoint eigenvalues (see (3.13)), these solutions will span the p-dimensional *unstable* subspace of (3.23). These solutions satisfy the asymptotic estimates

$$\lim_{x \to -\infty} e^{\mu_i(\lambda; \mathbf{p})x} W_i^-(x, \lambda; \mathbf{p}) = \eta_i^-(\lambda; \mathbf{p}) \,, \quad i = 1, \ldots, p \,. \tag{3.26}$$

By Proposition 1.2 of [42], each of the functions $W_i^-(x, \lambda; \mathbf{p})$ is analytic for $\lambda \in \Lambda$, and the limit is uniform on compact subsets of Λ.

Due to the ordering of the eigenvalues

$$W_1^-(x, 0; \mathbf{p}) := \lim_{\lambda \to 0} W_1^-(x, \lambda; \mathbf{p}) = \chi_{00}^-(\mathbf{p}) \widetilde{Z}_x(x; \mathbf{p}) = \chi_{00}^-(\mathbf{p}) \xi_0(\widetilde{Z}(x; \mathbf{p})) \,.$$

Since the dual vectors $\eta_i^-(\lambda; \mathbf{p})$ are meromorphic functions of λ, the poles can be eliminated by defining functions $\widetilde{W}_i^-(x, \lambda; \mathbf{p})$ which are asymptotic to the analytic functions $\widetilde{\eta}_i^-(\lambda; \mathbf{p})$,

$$\widetilde{W}_i^-(x, \lambda; \mathbf{p}) = \lambda^{m_i^-} W_i^-(x, \lambda; \mathbf{p}) \,, \quad i = 1, \ldots, p \,. \tag{3.27}$$

For $i = 1, \ldots, q = p - 1$ the solutions $\widetilde{W}_{i+1}^-(x, \lambda; \mathbf{p})$ at $\lambda = 0$ are also known because of equation (3.15):

$$\widetilde{W}_{i+1}^-(x, 0; \mathbf{p}) := \lim_{\lambda \to 0} \lambda^{m_{i+1}^-} W_{i+1}^-(x, \lambda; \mathbf{p}) = \sum_{j=1}^{q} \chi_{ji}^-(\mathbf{p}) \xi_j(\widetilde{Z}(x; \mathbf{p})) \,. \tag{3.28}$$

3.5. Summary of hypotheses

We summarize and formalize the hypotheses for the analysis of the linear stability problem (3.2)-(3.3).

Hypothesis *We assume that*

1. *the spectral parameter* $\lambda \in \Lambda$ *where* Λ *is an open simply-connected subset of* $\mathbb{C}_+$ *, the right-half complex* λ*-plane, and* $0 \in \overline{\Lambda}$*;*
2. *for all* $\lambda \in \Lambda$*, the spectrum of* $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ *(defined in* (3.7)*) consists of* p *simple eigenvalues with strictly negative real part, where* $p = q+1$*, and* $2n-p$ *eigenvalues with non-negative real part;*
3. *the spectrum of the linearized system at infinity (i.e. the matrix* $\mathbf{A}^{\pm}(0;\mathbf{p})$*) is of the form* $\{\beta(\mathbf{p}), -\beta(\mathbf{p})\} \cup \Sigma_0(\mathbf{p})$ *with* $\Sigma_0(\mathbf{p}) \subset i\mathbb{R}$ *and* $\beta(\mathbf{p})$ *real and positive for each* $\mathbf{p} \in \mathrm{A} \times \mathrm{B} \times \mathcal{C}$*. The eigenvalues* $\pm\beta(\mathbf{p})$ *are simple;*
4. *the following vector spaces are isomorphic:*
$$\mathrm{span}\{\zeta_2^+(0;\mathbf{p}),\cdots,\zeta_p^+(0;\mathbf{p})\} \quad \textit{and} \quad \mathrm{span}\{\xi_1(Z_0^+(\mathbf{p})),\cdots,\xi_q(Z_0^+(\mathbf{p}))\}\,,$$
and
$$\mathrm{span}\{\eta_2^-(0;\mathbf{p}),\cdots,\eta_p^-(0;\mathbf{p})\} \quad \textit{and} \quad \mathrm{span}\{\xi_1(Z_0^-(\mathbf{p})),\cdots,\xi_q(Z_0^-(\mathbf{p}))\}\,.$$
This implies that there exist invertible $q\times q$ *matrices* $\boldsymbol{\chi}^{\pm}$ *such that*
$$[\zeta_2^+(0;\mathbf{p}),\cdots,\zeta_p^+(0;\mathbf{p})] = [\xi_1(Z_0^+(\mathbf{p})),\cdots,\xi_q(Z_0^+(\mathbf{p}))]\,\boldsymbol{\chi}^+(\mathbf{p})$$
and
$$[\xi_1(Z_0^-(\mathbf{p})),\cdots,\xi_q(Z_0^-(\mathbf{p}))]\,\boldsymbol{\chi}^-(\mathbf{p}) = [\widetilde{\eta}_2^-(0;\mathbf{p})|\cdots|\widetilde{\eta}_p^-(0;\mathbf{p})]\,.$$

Remark 3. In §4 it is shown that property 4 above is satisfied if $B^{\pm}(\mathbf{p})$ has exactly q eigenvalues 0, and the manifold at infinity is non-degenerate.

4. Perturbation of the eigenvalues 0 of the matrix $\mathbf{A}^{\pm}(0;\mathbf{p})$

When $\lambda = 0$, the spectrum of $\mathbf{A}^{\pm}(0;\mathbf{p})$ has, by hypothesis, one pair of hyperbolic eigenvalues with the remainder of the spectrum on the imaginary axis and predominantly at the origin. An important factor in the stability analysis is how the eigenvalues 0 of $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ perturb when $\lambda \neq 0$. In this section we show that the perturbation of these eigenvalues 0 is determined by the parameter structure of the family of invariant manifolds at infinity, $\mathcal{M}_\infty(\mathbf{p})$. This geometric perturbation theory involves an interplay between the temporal and spatial symplectic structures.

To simply the exposition, we first give the details for the case where the symmetry group $\mathcal{G}$ is one dimensional (hence generically the manifold at infinity $\mathcal{M}_\infty(\mathbf{p})$ is one dimensional), and then state the modification necessary for the higher-dimensional case. Before considering the perturbation of the eigenvalues 0 of $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ for $|\lambda|$ small, we recall the parameter structure associated with the manifold at infinity $\mathcal{M}_\infty(\mathbf{p})$.

Consider the original PDE in the form

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z)\,, \quad Z \in \mathbb{R}^{2n}\,. \tag{4.1}$$

Let $\mathcal{G}$ be a one parameter subgroup of the Euclidean group on $\mathbb{R}^{2n}$ with action G_θ and infinitesimal generator $\xi(Z) = \frac{\mathrm{d}}{\mathrm{d}\theta}G_\theta Z|_{\theta=0}$. Associated with this one-parameter group is the pair of functions $P(Z)$ and $Q(Z)$ which satisfy

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}P(Z+\epsilon W)\,|_{\epsilon=0} = \omega(\xi(Z),W) \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}\epsilon}Q(Z+\epsilon W)\,|_{\epsilon=0} = \kappa(\xi(Z),W)\,, \tag{4.2}$$

where W is any vector in $\mathbb{R}^{2n}$, and ω and κ are the pre-symplectic two-forms associated with time and space respectively. These equalities define P and Q (up to an arbitrary constant) in the sense that $\xi(Z)$ is either $\mathbf{G}Z$ for some skew-symmetric matrix $\mathbf{G}$ in which case the functions P and Q are explicitly determined quadratic functions, or $\xi(Z) = V \in \mathbb{R}^{2n}$ and P and Q are explicitly determined linear functions. Level sets of the functions P and Q define surfaces in $\mathbb{R}^{2n}$ but they are not invariant in any sense: indeed, they combine to form a conservation law: $\partial_t P + \partial_x Q = 0$, and this follows from (4.2) (see [13] Appendix C).

The state at infinity is part of a *two-parameter* family of manifolds, $\mathcal{M}_\infty(a,b)$, of relative equilibria and each member of this family is of the form

$$Z(x,t) = G_{\theta(x,t)} Z_0 \quad \text{with} \quad \theta(x,t) = at + bx + \theta_0 \tag{4.3}$$

and $(Z_0, a, b) \in \mathbb{R}^{2n} \times \mathbb{R} \times \mathbb{R}$ which satisfy

$$\nabla S(Z_0) = a\nabla P(Z_0) + b\nabla Q(Z_0)\,, \quad P(Z_0) = \mathcal{P} \quad \text{and} \quad Q(Z_0) = \mathcal{Q}\,, \tag{4.4}$$

with $(\mathcal{P}, \mathcal{Q}) \in \mathbb{R}^2$ specified. A member of the two-parameter family is called non-degenerate when

$$\det \begin{bmatrix} \mathcal{P}_a & \mathcal{P}_b \\ \mathcal{Q}_a & \mathcal{Q}_b \end{bmatrix} \neq 0\,, \tag{4.5}$$

where $\mathcal{P}_a = \frac{\partial \mathcal{P}}{\partial a}$, etcetera. As indicated by the free parameter θ_0 in the definition of $\theta(x,t)$, equation (4.4) defines a one dimensional manifold if $\mathrm{Fix}(Z_0)$ is discrete. Hence in this case, the manifold $\mathcal{M}_\infty(a,b)$ is a one-dimensional manifold embedded in a two parameter family.

Let $Z(x,t) = G_{\theta(x,t)}\mathcal{T}_{ct+t_0}[Z_0 + \widehat{U}(x,t)]$; substitute into (4.1) and linearize about the invariant manifold at infinity, relative to a frame moving at speed c (i.e. linearize about $G_{\theta(x,t)}\mathcal{T}_{ct+t_0}Z_0$)

$$\mathbf{M}\widehat{U}_t + \mathbf{J}_c\widehat{U}_x = \mathbf{B}(a,b)\widehat{U}\,, \tag{4.6}$$

where $\mathbf{B}(a,b)$ is the symmetric matrix

$$\mathbf{B}(a,b) = D^2S(Z_0(a,b)) - aD^2P(Z_0(a,b)) - bD^2Q(Z_0(a,b))\,. \tag{4.7}$$

Here, we suppress the $\pm$ superscript as Z_0 can be any state in the manifold at infinity $\mathcal{M}_\infty$. Let $\widehat{U}(x,t) = \mathrm{e}^{\mu x + \lambda t}U$, with $U \in \mathbb{C}^{2n}$, and define the following spectral problem , where (μ, λ) are considered as a spectral pair,

$$\mathbf{B}(a,b)U = \mu\mathbf{J}_cU + \lambda\mathbf{M}U\,, \quad U \in \mathbb{C}^{2n}\,, \tag{4.8}$$

with spectral function

$$\Delta(\mu,\lambda;a,b,c) = \det[\mathbf{B}(a,b) - \mu\mathbf{J}_c - \lambda\mathbf{M}]\,. \tag{4.9}$$

This spectral function does not depend on the choice of $Z_0 \in \mathcal{M}_\infty$.

Lemma 2. *Assume $q = 1$ and suppose there exists a two-parameter family of nondegenerate manifolds at infinity, $\{\mathcal{M}_\infty(a,b) \mid a \in \mathrm{A},\, b \in \mathrm{B}\}$ with relative equilibria of the form (4.3), (4.4) and suppose $\mathbf{B}(a,b)$ has exactly one eigenvalue 0 with eigenvector $\xi(Z_0(a,b))$. Then, as $|\lambda| + |\mu| \to 0$, the spectral function satisfies*

$$\Delta(\mu,\lambda;a,b) = \frac{\Pi}{\|\xi(Z_0)\|^2}\left[\mathcal{Q}_b\mu^2 + (\mathcal{Q}_a + \mathcal{P}_b)\mu(\lambda - c\mu) + \mathcal{P}_a(\lambda - c\mu)^2\right] + o((|\lambda| + |\mu|)^2) \tag{4.10}$$

where Π is the product of the non-zero eigenvalues of the matrix $\mathbf{B}(a,b)$. If moreover,

$$\mathcal{Q}_b - c(\mathcal{Q}_a + \mathcal{P}_b) + c^2\mathcal{P}_a \neq 0\,, \tag{4.11}$$

there are two distinct eigenvalues of order λ as $|\lambda|$ is small.

Moreover, the eigenvectors $U \in \mathbb{C}^{2n}$ *associated with these eigenvalues have the following form as* $|\lambda| \to 0$

$$U(\lambda; a, b) = \mathbb{C}\left(\xi(Z_0) + \mu(\lambda)\frac{\partial Z_0}{\partial b} + (\lambda - c\mu(\lambda))\frac{\partial Z_0}{\partial a} + \mathcal{O}(|\lambda|^2)\right), \tag{4.12}$$

where $\mathbb{C}$ *represents an arbitrary complex constant and* $\mu(\lambda)$ *are the solutions of* (4.10) *which vanish in the limit as* $\lambda \to 0$.

Remark 4. With no limitations on the value of c, the hypothesis (4.11) is necessary. However, in the examples where c is restricted by the existence of the solitary wave, this hypothesis is often automatically satisfied.

Proof. The proof of (4.12) is based primarily on the following three results,

$$\mathbf{B}\xi(Z_0) = 0, \quad \mathbf{B}(\partial_a Z_0) = \nabla P(Z_0) \quad \text{and} \quad \mathbf{B}(\partial_b Z_0) = \nabla Q(Z_0). \tag{4.13}$$

These expressions can be verified using the symmetry and parameter structure of (4.4). The expression (4.12) is then verified by substitution into (4.8).

To prove the leading order behaviour of the spectral function, we expand everything in terms of the eigenfunctions of $\mathbf{B}(a,b)$. Let $\nu_1, \ldots, \nu_{2n}$ be the $2n$ eigenvalues of the symmetric matrix $\mathbf{B}$ with $\nu_1 = 0$ and let Π be the product of the $2n-1$ non-zero eigenvalues. Let $u_1, \ldots, u_{2n}$ be the $2n$ orthonormalized eigenvectors associated with $\nu_1, \ldots, \nu_{2n}$ with $u_1 = \xi(Z_0)/\|\xi(Z_0)\|$. Then, with

$$\mathbf{T} = [u_1|u_2|\cdots|u_{2n}] \in g\ell(2n, \mathbb{R}) \cap O(2n),$$

the spectral function can be transformed to

$$\Delta(\mu, \lambda; a, b) = \det[\mathbf{T}^T(\mathbf{B}(a,b) - \lambda\mathbf{M} - \mu\mathbf{J}_c)\mathbf{T}] = \det[\mathbf{T}^T\mathbf{B}(a,b)\mathbf{T} - \lambda\mathbf{T}^T\mathbf{M}\mathbf{T} - \mu\mathbf{T}^T\mathbf{J}_c\mathbf{T}], \tag{4.14}$$

with

$$\mathbf{T}^T\mathbf{B}\mathbf{T} = \begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & \widetilde{\mathbf{D}} \end{bmatrix}, \quad \mathbf{T}^T\mathbf{M}\mathbf{T} = \begin{bmatrix} 0 & -\mathbf{m}^T \\ \mathbf{m} & \widetilde{\mathbf{M}} \end{bmatrix}, \quad \mathbf{T}^T\mathbf{J}_c\mathbf{T} = \begin{bmatrix} 0 & -\mathbf{n}^T \\ \mathbf{n} & \widetilde{\mathbf{J}_c} \end{bmatrix}, \tag{4.15}$$

where $\widetilde{\mathbf{D}} = \mathrm{diag}[\nu_2, \ldots, \nu_{2n}]$,

$$\mathbf{m} = \frac{1}{\|\xi(Z_0)\|}\begin{pmatrix} \langle u_2, \nabla P(Z_0)\rangle \\ \vdots \\ \langle u_{2n}, \nabla P(Z_0)\rangle \end{pmatrix}, \quad \mathbf{n} = \frac{1}{\|\xi(Z_0)\|}\begin{pmatrix} \langle u_2, \nabla Q(Z_0) - c\nabla P(Z_0)\rangle \\ \vdots \\ \langle u_{2n}, \nabla Q(Z_0) - c\nabla P(Z_0)\rangle \end{pmatrix}. \tag{4.16}$$

Then

$$\begin{aligned} \Delta(\mu, \lambda; a, b) &= \det[\mathbf{T}^T\mathbf{B}(a,b)\mathbf{T} - \lambda\mathbf{T}^T\mathbf{M}\mathbf{T} - \mu\mathbf{T}^T\mathbf{J}_c\mathbf{T}] \\ &= \det\begin{bmatrix} 0 & \lambda\mathbf{m}^T + \mu\mathbf{n}^T \\ -\lambda\mathbf{m} - \mu\mathbf{n} & \widetilde{\mathbf{D}} - \lambda\widetilde{\mathbf{M}} - \mu\widetilde{\mathbf{J}_c} \end{bmatrix} \\ &= \det[\widetilde{\mathbf{D}} - \lambda\widetilde{\mathbf{M}} - \mu\widetilde{\mathbf{J}_c}]\left[(\lambda\mathbf{m}^T + \mu\mathbf{n}^T)(\widetilde{\mathbf{D}} - \lambda\widetilde{\mathbf{M}} - \mu\widetilde{\mathbf{J}_c})^{-1}(\lambda\mathbf{m} + \mu\mathbf{n})\right] \\ &= \Pi\left[\mathbf{m}^T\widetilde{\mathbf{D}}^{-1}\mathbf{m}\lambda^2 + (\mathbf{m}^T\widetilde{\mathbf{D}}^{-1}\mathbf{n} + \mathbf{n}^T\widetilde{\mathbf{D}}^{-1}\mathbf{m})\mu\lambda + \mathbf{n}^T\widetilde{\mathbf{D}}^{-1}\mathbf{n}\mu^2\right] + \cdots \end{aligned} \tag{4.17}$$

since $\Pi = \det(\widetilde{D})$, and the dots represent terms of degree higher than two in λ and μ. The proof will be completed by showing that

$$\mathbf{m}^T\widetilde{\mathbf{D}}^{-1}\mathbf{m} = \|\xi(Z_0)\|^{-2}\frac{\partial P}{\partial a}, \quad \mathbf{n}^T\widetilde{\mathbf{D}}^{-1}\mathbf{n} = \|\xi(Z_0)\|^{-2}\left(\frac{\partial Q}{\partial b} - c\left(\frac{\partial Q}{\partial a} + \frac{\partial P}{\partial b}\right) + c^2\frac{\partial P}{\partial a}\right), \tag{4.18}$$

and

$$\mathbf{m}^T\widetilde{\mathbf{D}}^{-1}\mathbf{n}+\mathbf{n}^T\widetilde{\mathbf{D}}^{-1}\mathbf{m} = \|\xi(Z_0)\|^{-2}\left(\frac{\partial P}{\partial b}+\frac{\partial Q}{\partial a}-2c\frac{\partial P}{\partial a}\right) . \tag{4.19}$$

To prove these identities, expand $\nabla P(Z_0)$, $\nabla Q(Z_0)$, $\partial_a Z_0$ and $\partial_b Z_0$ in terms of the eigenfunctions of $\mathbf{B}$,

$$\nabla P(Z_0) = \sum_{j=2}^{2n} r_j^{(1)} u_j \,, \quad \nabla Q(Z_0) = \sum_{j=2}^{2n} r_j^{(2)} u_j \,, \tag{4.20}$$

and

$$\partial_a Z_0 = \mathbb{R}u_1 + \sum_{j=2}^{2n} s_j^{(1)} u_j \,, \quad \partial_b Z_0 = \mathbb{R}u_1 + \sum_{j=2}^{2n} s_j^{(2)} u_j \,. \tag{4.21}$$

It follows from (4.13) and these expressions that

$$\nu_j s_j^{(1)} = r_j^{(1)} \quad \text{and} \quad \nu_j s_j^{(2)} = r_j^{(2)} \,, \quad j=2,\ldots,2n \,. \tag{4.22}$$

In terms of the coefficients in the eigenfunction expansion,

$$\mathbf{m} = \|\xi(Z_0)\|^{-1}\begin{pmatrix} r_2^{(1)} \\ \vdots \\ r_{2n}^{(1)} \end{pmatrix} \quad \text{and} \quad \mathbf{n} = \|\xi(Z_0)\|^{-1}\begin{pmatrix} r_2^{(2)} - c r_2^{(1)} \\ \vdots \\ r_{2n}^{(2)} - c r_{2n}^{(1)} \end{pmatrix} . \tag{4.23}$$

Therefore

$$\mathbf{m}^T\widetilde{\mathbf{D}}^{-1}\mathbf{m} = \|\xi(Z_0)\|^{-2}\sum_{j=1}^{2n}\left(r_j^{(1)}\right)^2/\nu_j \,, \tag{4.24}$$

with similar expressions for other bilinear forms in (4.18) and (4.19).

On the other hand,

$$\frac{\partial P}{\partial a} = \langle \nabla P(Z_0), \partial_a Z_0\rangle = \left\langle \sum_{j=2}^{2n} r_j^{(1)} u_j, \mathbb{R}u_1 + \sum_{j=2}^{2n}\nu_j^{-1} r_j^{(1)} u_j \right\rangle = \sum_{j=2}^{2n}(r_j^{(1)})^2/\nu_j \,. \tag{4.25}$$

Similarly,

$$\frac{\partial Q}{\partial b} = \sum_{j=2}^{2n}(r_j^{(2)})^2/\nu_j \quad \text{and} \quad \frac{\partial Q}{\partial a} = \frac{\partial P}{\partial b} = \sum_{j=2}^{2n}(r_j^{(1)} r_j^{(2)})/\nu_j \,. \tag{4.26}$$

Comparing these equations with the expansions of (4.18) and (4.19) and noting that $\widetilde{\mathbf{D}}^{-1} = \mathrm{diag}(\nu_2^{-1},\ldots,\nu_{2n}^{-1})$ completes the proof.

A consequence of this result is that the μ-roots of the spectral function (4.9) satisfy a homogeneous quadratic in μ and λ, for $|\lambda|+|\mu|$ sufficiently small, of the form

$$A\mu^2 + 2B\mu\lambda + C\lambda^2 + \cdots = 0 \,, \tag{4.27}$$

with

$$A = c^2\mathcal{P}_a - c(\mathcal{Q}_a+\mathcal{P}_b)+\mathcal{Q}_b \,, \quad B = \frac{1}{2}(\mathcal{Q}_a+\mathcal{P}_b) - c\mathcal{P}_a \quad \text{and} \quad C = \mathcal{P}_a \,. \tag{4.28}$$

Another characterization of these coefficients is

$$A = \begin{pmatrix} -c \\ 1 \end{pmatrix}^T \begin{bmatrix} \mathcal{P}_a & \mathcal{P}_b \\ \mathcal{Q}_a & \mathcal{Q}_b \end{bmatrix} \begin{pmatrix} -c \\ 1 \end{pmatrix} , \quad C = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \begin{bmatrix} \mathcal{P}_a & \mathcal{P}_b \\ \mathcal{Q}_a & \mathcal{Q}_b \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$B = \tfrac{1}{2}\begin{pmatrix} -c \\ 1 \end{pmatrix}^T \begin{bmatrix} \mathcal{P}_a & \mathcal{P}_b \\ \mathcal{Q}_a & \mathcal{Q}_b \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \tfrac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \begin{bmatrix} \mathcal{P}_a & \mathcal{P}_b \\ \mathcal{Q}_a & \mathcal{Q}_b \end{bmatrix} \begin{pmatrix} -c \\ 1 \end{pmatrix} .$$

The discriminant of the homogeneous quadratic (4.27) is independent of c and non-zero since

$$B^2 - AC = -\det \begin{bmatrix} \mathcal{P}_a & \mathcal{P}_b \\ \mathcal{Q}_a & \mathcal{Q}_b \end{bmatrix} , \tag{4.29}$$

and the determinant is non-zero by hypothesis. With $A \neq 0$ (cf. (4.11)) the homogeneous quadratic has two distinct μ-roots for $\lambda \neq 0$, and $|\lambda|$ sufficiently small, of the form

$$\mu = \frac{1}{A}(-B \pm \sqrt{B^2 - AC})\lambda + \cdots . \tag{4.30}$$

For λ real, positive and small, the qualitative position of the μ-roots of the spectral function is shown in Figure 4.1.

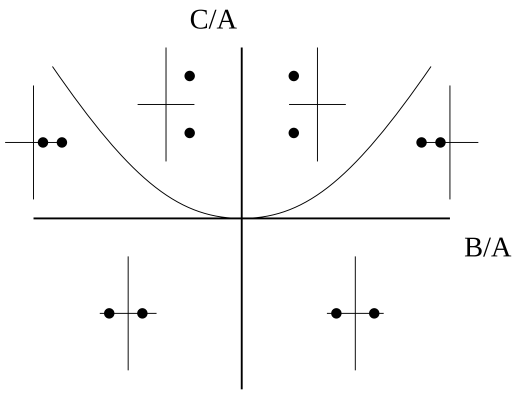

Fig. 4.1. Qualitative position of the μ-roots of (4.27) for λ real, positive and small, as a function of B/A and C/A.

In other words, the parameter structure of the manifold $\mathcal{M}_\infty$ determines precisely how the eigenvalues 0 of $\mathbf{A}^{\pm}(0;a,b)$ perturb when $|\lambda|$ is non-zero and small. Furthermore, if $q = 1$, then Hypotheses 2 and 4 in §3.5 imply that only one eigenvalue with negative real part bifurcates from 0. This suggests that we should expect C/A to be negative in the examples.

The sign of the discriminant (4.29) also plays a role in the linear stability of the state at infinity. In particular, if

$$\det \begin{bmatrix} \mathcal{P}_a & \mathcal{P}_b \\ \mathcal{Q}_a & \mathcal{Q}_b \end{bmatrix} > 0 \tag{4.31}$$

the invariant manifold at infinity is linearly unstable, and therefore may generate a branch of unstable continuous spectrum in the linearization about a solitary wave or front. To see this let $\mu = \mathrm{i}k$ with $k \in \mathbb{R}$; then for $|\lambda| + |k|$ sufficiently small, solving (4.27) and using that $C \neq 0$ leads to

$$\lambda = \mathrm{i}\frac{B}{C}k \pm \frac{k}{C}\left[\det \begin{pmatrix} \mathcal{P}_a & \mathcal{P}_b \\ \mathcal{Q}_a & \mathcal{Q}_b \end{pmatrix}\right]^{1/2} + \cdots \tag{4.32}$$

showing that positivity of the determinant results in an unstable eigenvalue (i.e. $\mathrm{Re}(\lambda) > 0$) with k small. $C \neq 0$ follows from the fact that $C = \mathcal{P}_a$, the 2×2 matrix in (4.31) is symmetric, and its determinant is positive.

This theory can be extended in a straightforward way to the general case where the Lie group $\mathcal{G}$ is q-dimensional (hence generically $\mathcal{M}_\infty$ is q-dimensional as well). As before, we assume that the manifolds at infinity $\mathcal{M}_\infty(\mathbf{a},\mathbf{b})$ are non-degenerate manifolds in a $2q$ parameter family.

Lemma 3. *Suppose there exists a $2q$-parameter family of nondegenerate manifolds at infinity $\mathcal{M}_\infty(\mathbf{a},\mathbf{b})$ and suppose $\mathbf{B}(\mathbf{a},\mathbf{b})$ has exactly q eigenvalues 0. Then, as $|\lambda|+|\mu|\to 0$, the spectral function satisfies*

$$\Delta(\mu,\lambda;\mathbf{a},\mathbf{b},c) = \widehat{\Pi}\,\det[\,\mathcal{Q}_{\mathbf{b}}\mu^2 + (\mathcal{Q}_{\mathbf{a}}+\mathcal{P}_{\mathbf{b}})\mu(\lambda-c\mu) + \mathcal{P}_{\mathbf{a}}(\lambda-c\mu)^2\,] + o((|\lambda|+|\mu|)^{2q})\,,$$

where $\widehat{\Pi} = \Pi\left(\prod_{i=1}^q \|\xi_i(Z_0)\|^2\right)^{-1}$ and Π is again the product of the non-zero eigenvalues of $\mathbf{B}$, $\mathcal{Q}_{\mathbf{a}}$ denotes the $q\times q$ matrix with entries $\frac{\partial \mathcal{Q}_i}{\partial a_j}$, with similar notation for $\mathcal{Q}_{\mathbf{b}}$, $\mathcal{P}_{\mathbf{a}}$ and $\mathcal{P}_{\mathbf{b}}$.

With a suitable generalization of (4.11), we can expect that as $\lambda\to 0$ the q eigenvalues $\mu_{i+1}(\lambda;\mathbf{p})$, $i=1,\ldots,q$, will perturb as $\mu_{i+1}(\lambda;\mathbf{p}) = \mu_{i+1}^{(1)}(\mathbf{p})\lambda + \mathcal{O}(\lambda^2)$, With each eigenvector perturbing as

$$U_{i+1}(\lambda;\mathbf{p}) = \mathbb{C}\sum_{j=1}^{q}\chi_{ij}(\mathbf{p})\left(\xi_j(Z_0) + [\,\mu_{i+1}^{(1)}\frac{\partial Z_0}{\partial b_j} + (1-c\mu_{i+1}^{(1)})\frac{\partial Z_0}{\partial a_j}\,]\,\lambda + \mathcal{O}(|\lambda|^2)\right)\,, \tag{4.33}$$

where $\chi_{ij}(\mathbf{p})$ are functions of $\mathbf{p}$ only, and determined by the matrices $\mathcal{Q}_{\mathbf{a}}$, $\mathcal{Q}_{\mathbf{b}}$, $\mathcal{P}_{\mathbf{a}}$ and $\mathcal{P}_{\mathbf{b}}$.

5. Hodge duality and the symplectic Evans matrix

Classically, a basic solitary wave or front is said to be *linearly unstable* or *spectrally unstable* if, for some $\lambda\in\mathbb{C}$ with positive real part, the linearized stability equation has a solution which is bounded for all $x\in\mathbb{R}$. Geometrically, the basic state is said to be linearly unstable if, for some $\lambda\in\mathbb{C}$ with $\Re(\lambda)>0$, the plane spanned by the solutions which are bounded as $x\to+\infty$ has a nontrivial intersection with the plane spanned by solutions which are bounded as $x\to-\infty$. This geometric interpretation was introduced by Evans [21] for the case where the space of solutions which are bounded as $x\to+\infty$ is one-dimensional. The generalization – to a class of parabolic PDEs where this dimension of the space of solutions as $x\to+\infty$ is arbitrary – was given by Alexander, Gardner & Jones [2]. Pego & Weinstein [42] first applied the Evans function theory to a class of Hamiltonian evolution equations.

In this section we extend the Evans function theory in several directions for Hamiltonian evolution equations. In Bridges & Derks [16] a new symplectic characterization of the Evans function for Hamiltonian evolution equations was given, for the case where the space of bounded solutions as $x\to+\infty$ (or $x\to-\infty$) is one-dimensional. Here we extend this theory to the case where this space has dimension p with p any natural number less than the dimension of the phase space. The result is a symplectification of the Evans function generalization introduced in [2]. Using Hodge duality, we find that the Evans function corresponds to the determinant of a $p\times p$ matrix of restricted Ω-symplectic forms. This $p\times p$ matrix we call *the symplectic Evans matrix.*

There is another generalization of interest: in the classical Evans function theory the asymptotic systems at infinity are required to be strictly hyperbolic. We are interested in systems with symmetry, and continuous symmetry will force the asymptotic matrices $\mathbf{A}^\pm(0;\mathbf{p})$ (and sometimes $\mathbf{A}^\pm(\lambda;\mathbf{p})$) to have eigenvalues 0.

Therefore we modify the class of admissible unstable eigenfunctions: if for some $\lambda\in\mathbb{C}$ with $\Re(\lambda)>0$, the basic state is said to be linearly unstable if the space of all *exponentially* decaying solutions as $x\to+\infty$ has a nontrivial intersection with the space of solutions which grow at most algebraically as $x\to-\infty$. If there are no eigenvalues of $\mathbf{A}^\pm(\lambda;\mathbf{p})$ on the imaginary axis (which includes all the examples presented in §§10-12), this criterion is equivalent to the usual criterion for linear instability: the eigenfunction decays exponentially as $x\to\pm\infty$. This condition also eliminates trivial instabilities due to exponential flow along a non-compact group orbit (see §6 of [16] for an example of such phenomena).

The starting point for the analysis is the linearized system

$$U_x = \mathbf{A}(x,\lambda;\mathbf{p})U\,, \quad U \in \mathbb{C}^{2n}\,, \quad x \in \mathbb{R}, \quad \lambda \in \Lambda\,, \tag{5.1}$$

with $\mathbf{A}(x,\lambda;\mathbf{p})$ as defined in (3.3). Central to the application of the Evans function theory is the behaviour of the asymptotic systems associated with (5.1):

$$\mathbf{A}^{\pm}(\lambda;\mathbf{p}) = \lim_{x\to\pm\infty} \mathbf{A}(x,\lambda;\mathbf{p})\,. \tag{5.2}$$

By hypothesis 2 in §3.5, for all $\lambda \in \Lambda$, the spectrum of $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ has p eigenvalues with strictly negative real part, denoted by

$$\{\mu_j(\lambda;\mathbf{p}) \mid j = 1,\dots,p\}\,, \tag{5.3}$$

and satisfying (3.9). The remainder of the spectrum has non-negative real part.

The space of solutions of (5.1) which decay exponentially as $x \to +\infty$ is spanned by

$$\{\ U_i^+(x,\lambda;\mathbf{p}) \mid i = 1,\dots,p\ \}\,, \tag{5.4}$$

and these solutions satisfy the asymptotic estimates (3.17), and are analytic for $\lambda \in \Lambda$. Define

$$Y^+(x,\lambda;\mathbf{p}) = U_1^+(x,\lambda;\mathbf{p}) \wedge \cdots \wedge U_p^+(x,\lambda;\mathbf{p}) \in {\textstyle\bigwedge}^p(\mathbb{C}^{2n})\,. \tag{5.5}$$

Then, $Y^+(x,\lambda;\mathbf{p})$ satisfies the asymptotic estimate

$$\lim_{x\to+\infty} \mathrm{e}^{-\sigma_+(\lambda;\mathbf{p})x} Y^+(x,\lambda;\mathbf{p}) = \zeta_1^+(\lambda;\mathbf{p}) \wedge \cdots \wedge \zeta_p^+(\lambda;\mathbf{p})\,, \tag{5.6}$$

where

$$\sigma_+(\lambda;\mathbf{p}) = \mu_1(\lambda;\mathbf{p}) + \cdots + \mu_p(\lambda;\mathbf{p})\,,$$

and $\{\zeta_1^+(\mathbf{p}),\dots,\zeta_p^+(\mathbf{p})\}$ span the stable subspace $E_+^s(\lambda;\mathbf{p})$ of $\mathbf{A}^+(\lambda;\mathbf{p})$.

The space of solutions of (5.1) which grow at most algebraically as $x \to -\infty$ is spanned by

$$\{\ U_i^-(x,\lambda;\mathbf{p}) \mid i = p+1,\dots,2n\ \}\,. \tag{5.7}$$

Define

$$Y^-(x,\lambda;\mathbf{p}) = U_{p+1}^-(x,\lambda;\mathbf{p}) \wedge \cdots \wedge U_{2n}^-(x,\lambda;\mathbf{p}) \in {\textstyle\bigwedge}^{2n-p}(\mathbb{C}^{2n})\,. \tag{5.8}$$

Then, $Y^-(x,\lambda;\mathbf{p})$ satisfies the asymptotic estimate

$$\lim_{x\to-\infty} \mathrm{e}^{-\sigma_-(\lambda;\mathbf{p})x} Y^-(x,\lambda;\mathbf{p}) = \zeta_{p+1}^-(\lambda;\mathbf{p}) \wedge \cdots \wedge \zeta_{2n}^-(\lambda;\mathbf{p})\,, \tag{5.9}$$

where $\{\zeta_{p+1}^-(\mathbf{p}),\dots,\zeta_{2n}^-(\mathbf{p})\}$ span the unstable and centre subspaces $E_-^u(\lambda;\mathbf{p}) + E_-^c(\lambda;\mathbf{p})$ of $\mathbf{A}^-(\lambda;\mathbf{p})$ and $\sigma_-(\lambda;\mathbf{p}) = \mathrm{Tr}(\mathbf{A}^{\pm}(\lambda;\mathbf{p})) - \sigma_+(\lambda;\mathbf{p})$. Note that $\mathrm{Tr}(\mathbf{A}^+(\lambda;\mathbf{p})) = \mathrm{Tr}(\mathbf{A}^-(\lambda;\mathbf{p}))$, since $\mathbf{A}^+$ and $\mathbf{A}^-$ have the same spectrum.

At least one of the eigenvalues of $\mathbf{A}^{\pm}(\lambda;\mathbf{p})$ has positive real part – since one of these is $\beta(\mathbf{p})$ when $\lambda = 0$ and the real part of this eigenvalue can not change sign for $\lambda \in \Lambda$. Therefore, the sum $\sigma_-(\lambda;\mathbf{p})$ has strictly positive real part: the decay of $Y^-(x,\lambda;\mathbf{p})$ to 0 as $x \to -\infty$ is exponential.

If $Y^+(x,\lambda;\mathbf{p})$ and $Y^-(x,\lambda;\mathbf{p})$ have a nontrivial intersection for some $(\lambda,\mathbf{p})$ with $\mathrm{Re}(\lambda) > 0$; that is, if $Y^-(x,\mathbf{p}) \wedge Y^+(x,\mathbf{p}) = 0$, then a linear combination of $U_1^+(x,\lambda;\mathbf{p}),\dots,U_p^+(x,\lambda;\mathbf{p})$ is a candidate for an unstable eigenfunction. This geometric construction is the basis for a generalization of the Evans function theory introduced in [2]; the Evans function takes the form,

$$\mathcal{D}(\lambda;\mathbf{p}) = \mathrm{e}^{-\int_0^x \mathrm{Tr}(\mathbf{A}(s,\lambda;\mathbf{p},\mathbf{b},c))\,\mathrm{d}s}\, Y^-(x,\lambda;\mathbf{p}) \wedge Y^+(x,\lambda;\mathbf{p})\,, \quad \lambda \in \Lambda\,. \tag{5.10}$$

This definition is a slight modification of that in [2]: here, instead of requiring exponential decay of the set $\{U^-_{p+1}(x,\mathbf{p}),\ldots,U^-_{2n}(x,\mathbf{p})\}$ as $x\to-\infty$, we require only that the set grows at most algebraically as $x\to-\infty$. However, as noted above, $Y^-(x,\lambda;\mathbf{p})$ will decay exponentially as $x\to-\infty$ in any case. (See [23,30] for further generalizations of the definition of the Evans function in (5.10).)

The multi-symplectic decomposition of $\mathbf{A}(x,\lambda;\mathbf{p})$ leads to an immediate simplification of the expression (5.10). Since $\mathbf{B}(x;\mathbf{p})$ is symmetric and $\mathbf{J}_c$ and $\mathbf{M}$ are skew-symmetric, it follows that

$$\mathrm{Tr}(\mathbf{A}(x,\lambda;\mathbf{p}))=\mathrm{Tr}(\mathbf{J}_c^{-1}\mathbf{B}(x;\mathbf{p}))-\lambda\mathrm{Tr}(\mathbf{J}_c^{-1}\mathbf{M})=-\lambda\tau(c)\,,$$

where

$$\tau(c)=\mathrm{Tr}(\mathbf{J}_c^{-1}\mathbf{M})=-\mathrm{Tr}(\mathbf{J}_c^{-1}\tfrac{\mathrm{d}}{\mathrm{d}c}\mathbf{J}_c)=-\frac{\frac{\mathrm{d}}{\mathrm{d}c}\det\mathbf{J}_c}{\det\mathbf{J}_c}\,,$$

or $\frac{\mathrm{d}}{\mathrm{d}c}\det\mathbf{J}_c+\tau(c)\det\mathbf{J}_c=0$. Hence

$$\mathcal{D}(\lambda;\mathbf{p})=\mathrm{e}^{\lambda\tau(c)x}\,Y^-(x,\lambda;\mathbf{p})\wedge Y^+(x,\lambda;\mathbf{p})\,. \tag{5.11}$$

The function $\mathcal{D}(\lambda;\mathbf{p})\in\bigwedge^{2n}(\mathbb{C}^{2n})$ and can therefore be characterized as a complex function times a volume form. Choose a canonical volume form on $\mathbb{C}^{2n}$:

$$\mathcal{V}=\mathbf{e}_1\wedge\cdots\wedge\mathbf{e}_{2n}\,,$$

where $\{\mathbf{e}_1,\ldots,\mathbf{e}_{2n}\}$ is an orthonormal basis for $\mathbb{C}^{2n}$. Then

$$\mathcal{D}(\lambda;\mathbf{p})=D(\lambda;\mathbf{p})\mathcal{V}\,,\quad D\,:\,\Lambda\times\mathrm{A}\times\mathrm{B}\times\mathcal{C}\to\mathbb{C}. \tag{5.12}$$

When $p=1$, it is proved by Bridges & Derks [16], that $D(\lambda;\mathbf{p})$ can be characterized as a restriction of the Ω-symplectic form. When $p>1$ we prove that $D(\lambda;\mathbf{p})$ is the determinant of a $p\times p$ matrix of restricted Ω-symplectic forms.

In preparation for the definition of this matrix, we introduce a family of dual solutions to (5.1), which are related to solutions of the adjoint problem. Let

$$\{\ W_j^-(x,\lambda;\mathbf{p})\ |\ j=1,\ldots,p\ \} \tag{5.13}$$

be solutions of the equation

$$W_x=\mathbf{A}(x,-\lambda;\mathbf{p})W\,,\quad W\in\mathbb{C}^{2n}\,, \tag{5.14}$$

with the asymptotic property (3.26). These functions are related to solutions of the adjoint of (5.1) as indicated in §3.

The set (5.13) forms a spanning set for the solutions of (5.14) which decay exponentially as $x\to-\infty$. This follows since

$$\lim_{x\to-\infty}\mathbf{A}(x,-\lambda;\mathbf{p})=\mathbf{A}^-(-\lambda;\mathbf{p})\,,$$

and the spectrum of $\mathbf{A}^-(-\lambda;\mathbf{p})$ is minus the spectrum of $\mathbf{A}^\pm(\lambda;\mathbf{p})$ (cf. §3). Hence the spectrum of $\mathbf{A}^-(-\lambda;\mathbf{p})$ has exactly p eigenvalues with strictly positive real part.

Definition 1. (**The symplectic Evans matrix**). For any $\mathbf{p}\in\mathrm{A}\times\mathrm{B}\times\mathcal{C}$ and $\lambda\in\Lambda$, the symplectic Evans matrix is defined by

$$\boldsymbol{\Omega}(\lambda;\mathbf{p})=\begin{bmatrix}\Omega(W_1^-(x,\lambda;\mathbf{p}),U_1^+(x,\lambda;\mathbf{p}))&\cdots&\Omega(W_1^-(x,\lambda;\mathbf{p}),U_p^+(x,\lambda;\mathbf{p}))\\ \vdots&\ddots&\vdots\\ \Omega(W_p^-(x,\lambda;\mathbf{p}),U_1^+(x,\lambda;\mathbf{p}))&\cdots&\Omega(W_p^-(x,\lambda;\mathbf{p}),U_p^+(x,\lambda;\mathbf{p}))\end{bmatrix} \tag{5.15}$$

where $U_i^+(x,\lambda;\mathbf{p})$ and $W_j^-(x,\lambda;\mathbf{p})$, $1\le i,j\le p$, are as defined above.

This matrix is called the *symplectic Evans matrix* because each entry is a restricted Ω-symplectic form. It is not a symplectic matrix in the sense that there exists a skew-symmetric matrix $\mathbf{J}$ such that $\boldsymbol{\Omega}^*\mathbf{J}\boldsymbol{\Omega} = \mathbf{J}$.

The fact that each entry of the symplectic Evans matrix is independent of x is a consequence of the following property.

Proposition 5. *Let $W_j^-(x,\lambda;\mathbf{p})$ be any solution of (5.14) and let $U_i^+(x,\lambda;\mathbf{p})$ be any solution of (5.1). Then $\Omega(W_j^-(x,\lambda;\mathbf{p}),U_i^+(x,\lambda;\mathbf{p}))$ is independent of x.*

Proof. First note that (5.1) and (5.14) can be written in the form

$$\mathbf{J}_c U_x = [\mathbf{B}(x;\mathbf{p}) - \lambda\mathbf{M}]U \quad \text{and} \quad \mathbf{J}_c W_x = [\mathbf{B}(x;\mathbf{p}) + \lambda\mathbf{M}]W\,. \tag{5.16}$$

Hence

$$\Omega(\frac{\mathrm{d}}{\mathrm{d}x}W_j^-, U_i^+) = \langle[\mathbf{B}(x;\mathbf{p}) + \lambda\mathbf{M}]W_j^-, U_i^+\rangle\,,$$

$$\Omega(W_j^-, \frac{\mathrm{d}}{\mathrm{d}x}U_i^+) = -\langle W_j^-, [\mathbf{B}(x;\mathbf{p}) - \lambda\mathbf{M}]U_i^+\rangle = -\langle[\mathbf{B}(x;\mathbf{p}) + \lambda\mathbf{M}]W_j^-, U_i^+\rangle\,,$$

and so

$$\frac{\mathrm{d}}{\mathrm{d}x}\Omega(W_j^-, U_i^+) = \Omega(\frac{\mathrm{d}}{\mathrm{d}x}W_j^-, U_i^+) + \Omega(W_j^-, \frac{\mathrm{d}}{\mathrm{d}x}U_i^+) = 0\,,$$

since $\mathbf{B}(x;\mathbf{p})$ is symmetric and $\mathbf{M}$ is skew-symmetric.

It remains to relate the symplectic Evans matrix to the complex function $D(\lambda;\mathbf{p})$ in (5.12). An important part of this relation is the connection between the set of adjoint solutions (5.13) which decay exponentially as $x \to -\infty$ and the set of solutions (5.7) which grow at most algebraically as $x \to -\infty$. We will show that this connection is a consequence of Hodge duality.

Lemma 4. *Let $\mathbf{p} \in \mathrm{A} \times \mathrm{B} \times \mathcal{C}$ be fixed. For any $\lambda \in \Lambda$, $\mathcal{D}(\lambda;\mathbf{p}) = D(\lambda;\mathbf{p})\mathcal{V}$ with*

$$D(\lambda;\mathbf{p}) = \det[\boldsymbol{\Omega}(\lambda;\mathbf{p})]\,. \tag{5.17}$$

Proof. The forms $Y^-(x,\lambda;\mathbf{p})$ and $Y^+(x,\lambda;\mathbf{p})$ are in $\bigwedge^{2n-p}(\mathbb{C}^{2n})$ and $\bigwedge^p(\mathbb{C}^{2n})$ respectively and the vector spaces $\bigwedge^{2n-p}(\mathbb{C}^{2n})$ and $\bigwedge^p(\mathbb{C}^{2n})$ are isomorphic, and the isomorphism is provided by the Hodge star operator. In [17] it is proved that $\star Y^-(x,\lambda;\mathbf{p}) \in \bigwedge^p(\mathbb{C}^{2n})$ satisfies

$$(\star Y^-(x,\lambda;\mathbf{p}))_x = \left[\mathrm{Tr}(\mathbf{A}(x,\overline{\lambda};\mathbf{p}))\mathbf{I}_p - (\mathbf{A}^{(p)}(x,\lambda;\mathbf{p}))^*\right](\star Y^-(x,\lambda;\mathbf{p}))\,, \tag{5.18}$$

where $\mathbf{I}_p$ is the identity on $\bigwedge^p(\mathbb{C}^{2n})$ and $(\mathbf{A}^{(p)}(\mathbf{p}))^*$ is the adjoint of the derivation $\mathbf{A}^{(p)}(\mathbf{p})$ (cf. [17]). Define

$$\mathcal{Y}^-(x,\lambda;\mathbf{p}) = \mathrm{e}^{\lambda\tau(c)x}Y^-(x,\lambda;\mathbf{p})\,. \tag{5.19}$$

Then $\star\mathcal{Y}^-(x,\lambda;\mathbf{p}) = \mathrm{e}^{\overline{\lambda}\tau(c)x}(\star Y^-(x,\lambda;\mathbf{p}))$ is a decomposable element in $\bigwedge^p(\mathbb{C}^{2n})$ and it satisfies the differential equation

$$(\star\mathcal{Y}^-(x,\lambda;\mathbf{p}))_x = [-(\mathbf{A}^{(p)}(x,\lambda;\mathbf{p}))^*](\star\mathcal{Y}^-(x,\lambda;\mathbf{p}))\,. \tag{5.20}$$

On each of the vector spaces $\bigwedge^p(\mathbb{R}^{2n})$, $k = 1,\cdots,2n$, there is a natural induced inner product. Let $u_1,\cdots,u_p$ and $v_1,\cdots,v_p$ be any vectors in $\mathbb{R}^{2n}$, and let

$$U = u_1 \wedge \cdots \wedge u_p \in \textstyle\bigwedge^p(\mathbb{R}^{2n}) \quad \text{and} \quad V = v_1 \wedge \cdots \wedge v_p \in \textstyle\bigwedge^p(\mathbb{R}^{2n})\,.$$

Then the inner product of U and V is defined by

$$[\![U,V]\!]_p = \det \begin{bmatrix} \langle u_1, v_1\rangle & \cdots & \langle u_1, v_p\rangle \\ \vdots & \ddots & \vdots \\ \langle u_p, v_1\rangle & \cdots & \langle u_p, v_p\rangle \end{bmatrix} . \tag{5.21}$$

The definition extends to any pair of elements of $\bigwedge^p(\mathbb{R}^{2n})$ (i.e. not necessarily decomposable) by linearity, and satisfies all the conditions of an inner product (cf. DARLING [19], §1.6). The inner products $\langle\cdot,\cdot\rangle$ and $[\![\cdot,\cdot]\!]_p$ extend in a natural way to inner products on $\mathbb{C}^{2n}$ and $\bigwedge^p(\mathbb{C}^{2n})$ respectively.

Using the complex inner product $[\![\cdot,\cdot]\!]_p$ on $\bigwedge^p(\mathbb{C}^{2n})$, it follows from the Hodge theory and the definition of $\mathcal{D}(\lambda;\mathbf{p})$ that

$$\mathcal{D}(\lambda;\mathbf{p}) = \mathcal{Y}^-(x,\lambda;\mathbf{p}) \wedge Y^+(x,\lambda;\mathbf{p}) = [\![\star\mathcal{Y}^-(x,\lambda;\mathbf{p}), Y^+(x,\lambda;\mathbf{p})]\!]_p \, \mathcal{V} . \tag{5.22}$$

Since $\star\mathcal{Y}^-(x,\lambda;\mathbf{p})$ is decomposable (cf. [17]), there exist p functions $\delta_1(x,\lambda;\mathbf{p}),\ldots,\delta_p(x,\lambda;\mathbf{p})$ such that

$$\star\mathcal{Y}^-(x,\lambda;\mathbf{p}) = \overline{\delta_1(x,\lambda;\mathbf{p})} \wedge \cdots \wedge \overline{\delta_p(x,\lambda;\mathbf{p})} , \tag{5.23}$$

where the overbar indicates complex conjugation. Then, using the definition of the inner product $[\![\cdot,\cdot]\!]_p$,

$$[\![\star\mathcal{Y}^-(x,\lambda;\mathbf{p}), Y^+(x,\lambda;\mathbf{p})]\!]_p = \det \begin{bmatrix} \langle \delta_1(x,\lambda;\mathbf{p}), U_1^+(x,\lambda;\mathbf{p})\rangle & \cdots & \langle \delta_1(x,\lambda;\mathbf{p}), U_p^+(x,\lambda;\mathbf{p})\rangle \\ \vdots & \ddots & \vdots \\ \langle \delta_p(x,\lambda;\mathbf{p}), U_1^+(x,\lambda;\mathbf{p})\rangle & \cdots & \langle \delta_p(x,\lambda;\mathbf{p}), U_p^+(x,\lambda;\mathbf{p})\rangle \end{bmatrix} .$$

It remains to connect the functions $\delta_j(x,\lambda;\mathbf{p})$ with the functions $W_j^-(x,\lambda;\mathbf{p})$. Each of the functions $\delta_j(x,\lambda;\mathbf{p})$ satisfies

$$\frac{d}{dx}\overline{\delta_j} = -\mathbf{A}(x,\lambda;\mathbf{p})^* \overline{\delta_j} ;$$

that is, each $\overline{\delta_j(x,\lambda;\mathbf{p})}$ is a solution of the adjoint equation and we have shown that each solution of the adjoint equation has the form $\mathbf{J}_c\overline{W(x,\lambda;\mathbf{p})}$ with $W(x,\lambda;\mathbf{p})$ satisfying $W_x = \mathbf{A}(x,-\lambda;\mathbf{p})W$. With the further requirement that the set $\{\delta_1(x,\lambda;\mathbf{p}),\ldots,\delta_p(x,\lambda;\mathbf{p})\}$ span the space of solutions of the adjoint equation which are exponentially decaying as $x\to-\infty$ (see [17]), it follows that the two sets

$$\{\delta_1(x,\lambda;\mathbf{p}),\ldots,\delta_p(x,\lambda;\mathbf{p})\} \quad \text{and} \quad \{\mathbf{J}_c W_1^-(x,\lambda;\mathbf{p}),\ldots,\mathbf{J}_c W_p^-(x,\lambda;\mathbf{p})\}$$

span the same space. Hence there exists an invertible $p\times p$ matrix $\mathbf{E}(\lambda;\mathbf{p})$ satisfying

$$[\delta_1(x,\lambda;\mathbf{p})|\cdots|\delta_p(x,\lambda;\mathbf{p})] = \mathbf{J}_c[W_1^-(x,\lambda;\mathbf{p})|\cdots|W_p^-(x,\lambda;\mathbf{p})]\mathbf{E}(\lambda;\mathbf{p}) . \tag{5.24}$$

Therefore, combining (5.22) and (5.24) we find

$$D(\lambda;\mathbf{p}) = [\![\star\mathcal{Y}^-(x,\lambda;\mathbf{p}), Y^+(x,\lambda;\mathbf{p})]\!]_p = \det[\mathbf{E}(\lambda;\mathbf{p})]\det[\boldsymbol{\Omega}(\lambda;\mathbf{p})] , \tag{5.25}$$

where $\boldsymbol{\Omega}(\lambda;\mathbf{p})$ is the symplectic Evans matrix. Since $\mathbf{E}(\lambda;\mathbf{p})$ is invertible and analytic for all $\lambda\in\Lambda$, the functions $\{U_{p+1}^-(\mathbf{p}),\ldots,U_{2n}^-(\mathbf{p})\}$ can be scaled (and permuted if necessary) in such a way that the determinant of $\mathbf{E}(\lambda;\mathbf{p})$ is unity.

This characterization of the Evans function and its properties combine to prove the following.

Corollary 1. *For any $\mathbf{p}=(\mathbf{a},\mathbf{b},c)\in \mathrm{A}\times\mathrm{B}\times\mathcal{C}$, if there exists a value of $\lambda\in\Lambda$ such that the determinant of the symplectic Evans matrix vanishes, the basic solitary wave or front is linearly spectrally unstable.*

5.1. Differentiating the Evans matrix

In order to analyse the Evans function for λ near 0, we will need the derivative of symplectic Evans matrix. The following result is a generalization of the Pego-Weinstein derivative formula ([42]) to the multi-dimensional case and a generalization of the multi-symplectic derivative formula of [16] to the case $p>1$. As in [16], we show that the derivative with respect to λ of the restricted $\varOmega$-symplectic form leads to an expression in terms of the pre-symplectic structure ω, associated with the time-dependence of the PDE. To recognize the fact that we will also be differentiating with respect to λ, differentiation with respect to x will be considered as partial differentiation.

Proposition 6. *Let $U_i^+(x,\lambda;\mathbf{p})$ and $W_j^-(x,\lambda;\mathbf{p})$ be as above. Then, for $1\le i,j\le p$,*

$$\omega(W_j^-(x,\lambda;\mathbf{p}),U_i^+(x,\lambda;\mathbf{p})) = \partial_x\left[\varOmega(\partial_\lambda W_j^-(x,\lambda;\mathbf{p}),U_i^+(x,\lambda;\mathbf{p}))\right] \tag{5.26}$$

and

$$\omega(W_j^-(x,\lambda;\mathbf{p}),U_i^+(x,\lambda;\mathbf{p})) = -\partial_x\left[\varOmega(W_j^-(x,\lambda;\mathbf{p}),\partial_\lambda U_i^+(x,\lambda;\mathbf{p}))\right]. \tag{5.27}$$

Proof. First note that $\partial_\lambda U_i^+(x,\lambda;\mathbf{p})$ and $\partial_\lambda W_j^-(x,\lambda;\mathbf{p})$ satisfy

$$\mathbf{J}_c(\partial_\lambda U_i^+)_x = [\mathbf{B}(x;\mathbf{p})-\lambda\mathbf{M}](\partial_\lambda U_i^+) - \mathbf{M}U_i^+ \tag{5.28}$$

and

$$\mathbf{J}_c(\partial_\lambda W_j^-)_x = [\mathbf{B}(x;\mathbf{p})+\lambda\mathbf{M}](\partial_\lambda W_j^-) + \mathbf{M}W_j^-\,. \tag{5.29}$$

The proof is now a direct calculation:

$$\partial_x\varOmega(\partial_\lambda W_j^-,U_i^+) = \langle\mathbf{J}_c(\partial_\lambda W_j^-)_x,U_i^+\rangle - \langle\partial_\lambda W_j^-,\mathbf{J}_c(U_i^+)_x\rangle = \langle\mathbf{M}W_j^-,U_i^+\rangle = \omega(W_j^-,U_i^+)\,,$$

for the first expression, with a similar calculation for the second one.

Lemma 5. *For any fixed $\mathbf{p}=(\mathbf{a},\mathbf{b},c)\in\mathrm{A}\times\mathrm{B}\times\mathcal{C}$ and all $\lambda\in\Lambda$ and any $R,S>0$ we have*

$$\frac{\partial}{\partial\lambda}\boldsymbol{\varOmega}(\lambda;\mathbf{p}) = \int_{-S}^{R}\begin{bmatrix}\omega(W_1^-,U_1^+) & \cdots & \omega(W_1^-,U_p^+)\\ \vdots & \ddots & \vdots\\ \omega(W_p^-,U_1^+) & \cdots & \omega(W_p^-,U_p^+)\end{bmatrix}\mathrm{d}x + BC(R,S)\,, \tag{5.30}$$

where $BC(R,S)=BC_1(R,S)+BC_2(R,S)$ and

$$BC_1(R,S) = \begin{bmatrix}\varOmega(\partial_\lambda W_1^-(-S,\lambda;\mathbf{p}),U_1^+(-S,\lambda;\mathbf{p})) & \cdots & \varOmega(\partial_\lambda W_1^-(-S,\lambda;\mathbf{p}),U_p^+(-S,\lambda;\mathbf{p}))\\ \vdots & \ddots & \vdots\\ \varOmega(\partial_\lambda W_p^-(-S,\lambda;\mathbf{p}),U_1^+(-S,\lambda;\mathbf{p})) & \cdots & \varOmega(\partial_\lambda W_p^-(-S,\lambda;\mathbf{p}),U_p^+(-S,\lambda;\mathbf{p}))\end{bmatrix},$$

$$BC_2(R,S) = \begin{bmatrix}\varOmega(W_1^-(R,\lambda;\mathbf{p}),\partial_\lambda U_1^+(R,\lambda;\mathbf{p})) & \cdots & \varOmega(W_1^-(R,\lambda;\mathbf{p}),\partial_\lambda U_p^+(R,\lambda;\mathbf{p}))\\ \vdots & \ddots & \vdots\\ \varOmega(W_p^-(R,\lambda;\mathbf{p}),\partial_\lambda U_1^+(R,\lambda;\mathbf{p})) & \cdots & \varOmega(W_p^-(R,\lambda;\mathbf{p}),\partial_\lambda U_p^+(R,\lambda;\mathbf{p}))\end{bmatrix}.$$

Proof. Every entry of $\boldsymbol{\Omega}(\lambda;\mathbf{p})$ is of the form $\Omega(W_j^-(x,\lambda;\mathbf{p}),U_i^+(x,\lambda;\mathbf{p}))$ and

$$\begin{aligned}\frac{\partial}{\partial\lambda}\left[\Omega(W_j^-(x,\lambda;\mathbf{p}),U_i^+(x,\lambda;\mathbf{p}))\right] &= \Omega(\partial_\lambda W_j^-(x,\lambda;\mathbf{p}),U_i^+(x,\lambda;\mathbf{p}))\\ &\quad + \Omega(W_j^-(x,\lambda;\mathbf{p}),\partial_\lambda U_i^+(x,\lambda;\mathbf{p}))\,.\end{aligned}\tag{5.31}$$

To evaluate the right-hand side of this expression, integrate the two identities in Proposition 6 over x,

$$\int_{-S}^{0}\omega(W_j^-,U_i^+)\,\mathrm{d}x = \Omega(\partial_\lambda W_j^-,U_i^+)\big|_{x=0} - \Omega(\partial_\lambda W_j^-,U_i^+)\big|_{x=-S}\,,$$

$$\int_{0}^{R}\omega(W_j^-,U_i^+)\,\mathrm{d}x = -\,\Omega(W_j^-,\partial_\lambda U_i^+)\big|_{x=R} + \Omega(W_j^-,\partial_\lambda U_i^+)\big|_{x=0}\,.$$

Now, since $\Omega(W_j^-,U_i^+)$ and $\partial_\lambda\Omega(W_j^-,U_i^+)$ are independent of x,

$$\partial_\lambda\Omega(W_j^-(x,\lambda;\mathbf{p}),U_i^+(x,\lambda;\mathbf{p})) = \partial_\lambda\Omega(W_j^-(x,\lambda;\mathbf{p}),U_i^+(x,\lambda;\mathbf{p}))\big|_{x=0}\,.$$

Hence,

$$\partial_\lambda\Omega(W_j^-,U_i^+) = \int_{-S}^{R}\omega(W_j^-,U_i^+)\,\mathrm{d}x + \Omega(\partial_\lambda W_j^-,U_i^+)\big|_{x=-S} + \Omega(W_j^-,\partial_\lambda U_i^+)\big|_{x=R}\,,$$

and this expression holds for each $1\le i,j\le p$.

The second derivative of $\boldsymbol{\Omega}(\lambda;\mathbf{p})$ is obtained by formal differentiation of (5.30). In the sequel, the idea will be to take the limit as R and S go to infinity.

When $p=1$, the symplectic Evans matrix reduces to a scalar and the derivative formula, when $D(\lambda_0;\mathbf{p})=0$, to

$$\left.\frac{\partial}{\partial\lambda}\right|_{\lambda=\lambda_0} D(\lambda;\mathbf{p}) = \int_{-\infty}^{+\infty}\omega(W_1^-(x,\lambda_0;\mathbf{p}),U_1^+(x,\lambda_0;\mathbf{p}))\,\mathrm{d}x\,,$$

which is the multi-symplectic characterization of the Pego-Weinstein derivative formula (cf. [16]).

6. A geometric condition for linearized instability

To deduce a geometric criterion for linear instability we will study the slope of the Evans function for λ real and near 0. An explicit formula is deduced for this slope. Then by the intermediate value theorem, if there exists real positive value of λ, denoted λ_∞ with $\lambda_\infty\in\Lambda\cap\mathbb{R}$, with the property that the sign of $D(\lambda_\infty;\mathbf{p})$ is opposite the sign of the slope of $D(\lambda;\mathbf{p})$ near 0, we can conclude the existence of an unstable eigenvalue $\lambda_0\in\Lambda\cap\mathbb{R}$ with $0<\lambda_0<\lambda_\infty$.

The advantage of such a geometric criterion for instability is that properties of the basic state which are known and readily available can be used to predict the existence of an unstable eigenvalue of the spectral problem, thereby avoiding direct analysis of the spectral problem, which is in general intractable.

Theorem 1. *For some fixed* $\mathbf{p}=(\mathbf{a},\mathbf{b},c)\in\mathrm{A}\times\mathrm{B}\times\mathcal{C}$, *let* $\lambda_\infty\in\Lambda\cap\mathbb{R}$ *be a positive value of* λ *and let* $d_\infty(\mathbf{p})=D(\lambda_\infty;\mathbf{p})$. *Recall that (see section 3)*

$$\begin{aligned}\chi_{00}^-(\mathbf{p}) &= \lim_{x\to\infty}\left[\,e^{2\beta(\mathbf{p})x}\,\Omega(\widetilde{Z}_x(-x;\mathbf{p}),DG_\gamma(Z_0^-(\mathbf{p}))^T\widetilde{Z}_x(x;\mathbf{p}))\,\right]^{-1}\\ &= \left[\,\Omega(\Psi^-(\mathbf{p}),DG_\gamma(Z_0^-(\mathbf{p}))^T\Psi^+(\mathbf{p}))\,\right]^{-1}.\end{aligned}$$

If

$$d_\infty(\mathbf{p})\,\chi_{00}^-(\mathbf{p})\,\left[\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \tfrac{1}{2}\,\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p}))\right] < 0,$$

then the solitary wave or front $G_{\theta(x,t;\mathbf{p})}\widetilde{Z}(x-ct;\mathbf{p})$ *is linearly spectrally unstable.*

Note that, as in the $p=1$ case [16], the sign of $\left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \frac{1}{2}\,\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p}))\right]$ plays an important role in determining the stability of a wave. Methods for determining d_∞ are considered in §8. The proof of Theorem 1 requires establishing the sign of the slope of the symplectic Evans function for λ real and near 0. In the next section we will show that the symplectic Evans function has the following Taylor expansion for λ near 0,

$$D(\lambda,\mathbf{p}) = \lambda^2\,\chi_{00}^-(\mathbf{p})\ \det(\partial\mathcal{D}(\mathbf{p}))\ \left[\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \tfrac{1}{2}\,\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p}))\right] + \mathcal{O}(\lambda^3), \tag{6.1}$$

where $\partial\mathcal{D}\in g\ell(q,\mathbb{C})$ is the matrix with entries

$$(\partial\mathcal{D})_{ij} = \sum_{k=1}^{q}\Big(\chi_{ki}^+\,\Omega(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\,\widetilde{\eta}_{j+1}^-,\xi_k(Z_0^-)) + \chi_{kj}^-\,\Omega(\xi_k(Z_0^+),\,\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\,\zeta_{i+1}^+)\Big) \tag{6.2}$$

(see section 3 for the definitions of the symbols used in this expression). By using the normalisation of η_i^-, and the geometric perturbation theory of §4 we can characterize this matrix precisely: we obtain the remarkable result that $\det(\partial\mathcal{D})=1$.

Lemma 6. *If $\mathbf{B}^\pm(\mathbf{p})$ has exactly q eigenvalues 0, then the matrix $\partial\mathcal{D}$ is equal to the identity matrix.*

Proof. From equation (4.33) for the eigenvectors of the linearised system and the definition of the matrix $\boldsymbol{\chi}^+(\mathbf{p})$, it follows that

$$\zeta_{i+1}^+(\lambda;\mathbf{p}) = \zeta_{i+1}^{+,0}(\mathbf{p}) + \lambda\zeta_{i+1}^{+,1}(\mathbf{p}) + \mathcal{O}(|\lambda|^2)\,,$$

where

$$\zeta_{i+1}^{+,0}(\mathbf{p}) = \sum_{j=1}^{q}\chi_{ij}^+(\mathbf{p})\xi_j(Z_0^+(\mathbf{p}))\quad\text{and}\quad \zeta_{i+1}^{+,1}(\mathbf{p}) = \sum_{j=1}^{q}\chi_{ij}^+(\mathbf{p})\ \left[\mu_{i+1}^{(1)}(\mathbf{p})\frac{\partial Z_0^+}{\partial b_j} + (1-c\mu_{i+1}^{(1)}(\mathbf{p}))\frac{\partial Z_0^+}{\partial a_j}\ \right]$$

with $\mu_{i+1}^{(1)}(\mathbf{p})$ the first-order term in the series expansion of $\mu_{i+1}(\lambda;\mathbf{p})$.

By definition, $\eta_{i+1}(\lambda;\mathbf{p})$ is a solution of the adjoint system (3.13) with $\mu=\mu_{i+1}(\lambda;\mathbf{p})$, hence equation (4.33) implies that for λ near 0

$$\widetilde{\eta}_{i+1}^-(\lambda;\mathbf{p}) = \eta_{i+1}^{-,0}(\mathbf{p}) + \lambda\eta_{i+1}^{-,1}(\mathbf{p}) + \mathcal{O}(|\lambda|^2)\,,$$

where

$$\eta_{i+1}^{-,0}(\mathbf{p}) = \sum_{j=1}^{q}\chi_{ij}^-(\mathbf{p})\xi_j(Z_0^-(\mathbf{p}))\quad\text{and}\quad \eta_{i+1}^{-,1}(\mathbf{p}) = -\sum_{j=1}^{q}\chi_{ij}^-(\mathbf{p})\ \left[\mu_{i+1}^{(1)}(\mathbf{p})\frac{\partial Z_0^-}{\partial b_j} + (1-c\mu_{i+1}^{(1)}(\mathbf{p}))\frac{\partial Z_0^-}{\partial a_j}\ \right].$$

Using that $\xi_k(Z_0^+) = DG_\gamma(Z_0^-)\xi_k(Z_0^-)$, we get that elements of $\partial\mathcal{D}$ can be written as

$$(\partial\mathcal{D})_{ij} = \Omega(\eta_{j+1}^{-,1},DG_\gamma(Z_0^-)^T\zeta_{i+1}^{+,0}) + \Omega(DG_\gamma(Z_0^-)\eta_{j+1}^{-,0},\zeta_{i+1}^{+,1}).$$

Now recall that η_j^- is normalised such that $\Omega(\eta_j^-,DG_\gamma(Z_0^-)^T\zeta_i^+) = \delta_{ij}$. Using the expansions as given above, this implies that

$$\delta_{ij} = \Omega(\eta_j^{-,1},DG_\gamma(Z_0^-)^T\zeta_i^{+,0}) + \Omega(\eta_j^{-,0},DG_\gamma(Z_0^-)^T\zeta_i^{+,1}).$$

since $\Omega(\xi_k(Z_0),\xi_l(Z_0))=0$, for all $k,l=1,\ldots,q$. Using that

$$\Omega(DG_\gamma(Z_0^-)\eta_{j+1}^{-,0},\zeta_{i+1}^{+,1}) = \Omega(\eta_{j+1}^{-,0},DG_\gamma(Z_0^-)^T\zeta_{i+1}^{+,1}),$$

we can conclude that $(\partial\mathcal{D})_{ij}=\delta_{ij}$.

7. The symplectic Evans matrix for λ near zero

In this section we study entries in the symplectic Evans matrix for λ real, positive and near zero, and prove the properties of the slope of the determinant of this matrix needed for the proof of Theorem 1. The next result establishes the expansion (6.1).

Theorem 2. *Suppose* $\mathbf{p} = (\mathbf{a}, \mathbf{b}, c) \in \mathrm{A} \times \mathrm{B} \times \mathcal{C}$ *are fixed. Let* λ *be real and positive. Then*

$$\det \boldsymbol{\Omega}(0;\mathbf{p}) = 0\,, \qquad \frac{\partial}{\partial\lambda}\bigg|_{\lambda=0} \det \boldsymbol{\Omega}(\lambda;\mathbf{p}) = 0\,,$$

and

$$\frac{\partial^2}{\partial\lambda^2}\bigg|_{\lambda=0} \det \boldsymbol{\Omega}(\lambda;\mathbf{p}) = 2\chi_{00}^-(\mathbf{p}) \det(\partial\mathcal{D}(\mathbf{p})) \left[\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \tfrac{1}{2}\,\omega(Z_0^+(\mathbf{p}), \partial_c Z_0^+(\mathbf{p}))\right].$$

The proof of this theorem will follow from a sequence of lemmas. The idea will be to expand each entry in the symplectic Evans matrix in a Taylor series in λ, and then use the interplay between the Ω-symplectic structure and the ω-symplectic structure to characterize the coefficients in the Taylor expansions.

In the limit as $\lambda \to 0$, the adjoint functions $W_j^\pm(x,\lambda;\mathbf{p})$ have poles and this has lead in §3.4 to the introduction of the functions $\widetilde{W}_j^\pm(x,\lambda;\mathbf{p})$ which are analytic functions of λ when $\lambda = 0$. Thus for the limit as $\lambda \to 0$ it is useful to reformulate $\boldsymbol{\Omega}(\lambda;\mathbf{p})$ in terms of these functions. Let

$$\mathbf{N}(\lambda) = \mathrm{diag}[1, \lambda, \dots, \lambda] \in g\ell(p, \mathbb{C})\,,$$

and define $\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p}) = \mathbf{N}(\lambda)\,\boldsymbol{\Omega}(\lambda;\mathbf{p})$. Then

$$\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p}) = \begin{pmatrix} \Omega(W_1^-(x,\lambda;\mathbf{p}), U_1^+(x,\lambda;\mathbf{p})) & \cdots & \Omega(W_1^-(x,\lambda;\mathbf{p}), U_p^+(x,\lambda;\mathbf{p})) \\ \Omega(\widetilde{W}_2^-(x,\lambda;\mathbf{p}), U_1^+(x,\lambda;\mathbf{p})) & \cdots & \Omega(\widetilde{W}_2^-(x,\lambda;\mathbf{p}), U_p^+(x,\lambda;\mathbf{p})) \\ \vdots & \ddots & \vdots \\ \Omega(\widetilde{W}_p^-(x,\lambda;\mathbf{p}), U_1^+(x,\lambda;\mathbf{p})) & \cdots & \Omega(\widetilde{W}_p^-(x,\lambda;\mathbf{p}), U_p^+(x,\lambda;\mathbf{p})) \end{pmatrix}, \tag{7.1}$$

and

$$\det \boldsymbol{\Omega}(\lambda;\mathbf{p}) = \lambda^{-(p-1)} \det \widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p})\,.$$

We will prove Theorem 2 by finding the first few terms of the series expansion of each entry of $\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p})$, for λ small. The functions U_i^+ and $\widetilde{W}_j^-$ are generators of the symmetry group $\mathcal{G}$ or the translation group, in the limit for $\lambda = 0$ (cf. §3 and §4). Explicitly, we have

$$\begin{gathered} U_1^+(x,0;\mathbf{p}) = \xi_0(\widetilde{Z}(x;\mathbf{p})) \quad \text{and} \quad \widetilde{W}_1^-(x,0;\mathbf{p}) = W_1^-(x,0;\mathbf{p}) = \chi_{00}^-(\mathbf{p})\xi_0(\widetilde{Z}(x;\mathbf{p})) \\ U_{i+1}^+(x,0;\mathbf{p}) = \textstyle\sum_{j=1}^q \chi_{ji}^+(\mathbf{p})\xi_j(\widetilde{Z}(x;\mathbf{p})) \quad \text{and} \quad \widetilde{W}_{i+1}^-(x,0;\mathbf{p}) = \sum_{j=1}^q \chi_{ji}^-(\mathbf{p})\xi_j(\widetilde{Z}(x;\mathbf{p})), \end{gathered} \tag{7.2}$$

for $i = 1, \dots, q = p-1$.

Lemma 7. *For any fixed* $\mathbf{p} = (\mathbf{a}, \mathbf{b}, c) \in \mathrm{A} \times \mathrm{B} \times \mathcal{C}$,

$$\widetilde{\boldsymbol{\Omega}}_{ij}(0;\mathbf{p}) = \Omega(\widetilde{W}_i^-(x,0;\mathbf{p}), U_j^+(x,0;\mathbf{p}))\,, \quad \textit{for} \quad 1 \le i,j \le p\,.$$

Proof. By the definition of $\widetilde{\boldsymbol{\Omega}}_{ij}$ and the expressions for $U_i^+(x,0;\mathbf{p})$ and $\widetilde{W}_j^-(x,0;\mathbf{p})$ we get for $1 \le i,j \le q$

$$\begin{aligned} \widetilde{\boldsymbol{\Omega}}_{11} &= \Omega(W_1^-, U_1^+) = \Omega(\chi_{00}^- \widetilde{Z}_x, \widetilde{Z}_x) = \chi_{00}^- \Omega(\widetilde{Z}_x, \widetilde{Z}_x) = 0 \\ \widetilde{\boldsymbol{\Omega}}_{1(i+1)} &= \Omega(\widetilde{W}_1^-, U_{i+1}^+) = \chi_{00}^- \lim_{x\to-\infty} \textstyle\sum_{k=1}^q \chi_{ki}^+ \Omega(\widetilde{Z}_x, \xi_k(\widetilde{Z})) = 0 \\ \widetilde{\boldsymbol{\Omega}}_{(j+1)1} &= \Omega(\widetilde{W}_{j+1}^-, U_1^+) = \lim_{x\to-\infty} \textstyle\sum_{k=1}^q \chi_{kj}^- \Omega(\xi_k(\widetilde{Z}), \widetilde{Z}_x) = 0 \end{aligned}$$

since $\lim_{x\to-\infty} \widetilde{Z}_x = 0$ and $\lim_{x\to-\infty} \xi_k(\widetilde{Z}) = \xi_k(Z_0^-)$, for $k = 1, \dots, q$.

Furthermore,

$$\mathbf{J}_c \xi_l(Z_0^-) = (\mathbf{K} - c\mathbf{M})\xi_l(Z_0^-) = \nabla Q_l(Z_0^-) - c\nabla P_l(Z_0^-)\,, \quad \text{for } l = 1, \dots, q\,,$$

and

$$\langle \xi_k(Z_0^-), \nabla Q_l(Z_0^-)\rangle = 0 = \langle \xi_k(Z_0^-), \nabla P_l(Z_0^-)\rangle \quad \text{for } 1 \le k, l \le q = p-1\,,$$

because of the invariance of P_l and Q_l under the action of $\mathcal{G}$. Therefore $\Omega(\xi_l(Z_0^-), \xi_k(Z_0^-)) = 0$ and

$$\widetilde{\boldsymbol{\Omega}}_{(i+1)(j+1)} = \Omega(\widetilde{W}_{i+1}^-, U_{j+1}^+) = \sum_{k=1}^{q}\sum_{l=1}^{q} \chi_{kj}^+ \chi_{li}^- \Omega(\xi_l(Z_0^-), \xi_k(Z_0^-)) = 0,$$

for $1 \le i, j \le q = p-1$.

To derive the next term in the series expansion of $\widetilde{\boldsymbol{\Omega}}_{ij}$ about $\lambda = 0$, we use the derivative formula developed in §5. Lemma 5 shows that we need the behaviour of $U_i^+(x, 0; \mathbf{p})$ and $\frac{\mathrm{d}}{\mathrm{d}\lambda}\widetilde{W}_j^-(x, 0; \mathbf{p})$ for $x \to -\infty$ and $\widetilde{W}_j^-(x, 0; \mathbf{p})$ and $\frac{\mathrm{d}}{\mathrm{d}\lambda}U_i^+(x, 0; \mathbf{p})$ for $x \to \infty$.

Lemma 8. *Let* $1 \le i \le p$. *For* $x \to \infty$ *we have*

$$\frac{\partial}{\partial\lambda}U_i^+(x,\lambda;\mathbf{p}) = \mathrm{e}^{\mu_i(\lambda;\mathbf{p})x}\left(x\frac{\partial}{\partial\lambda}\mu_i(\lambda;\mathbf{p})\zeta_i^+(\lambda;\mathbf{p}) + \frac{\partial}{\partial\lambda}\zeta_i^+(\lambda;\mathbf{p})\right) + o(\mathrm{e}^{\mu_i(\lambda;\mathbf{p})x}(1+x)), \tag{7.3}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial\lambda^2}U_i^+(x,\lambda;\mathbf{p}) = \mathrm{e}^{\mu_i(\lambda;\mathbf{p})x}\Big(&x^2\left(\tfrac{\partial}{\partial\lambda}\mu_i(\lambda;\mathbf{p})\right)^2\zeta_i^+(\lambda;\mathbf{p}) + \tfrac{\partial^2}{\partial\lambda^2}\zeta_i^+(\lambda;\mathbf{p})\\ &+ x\left[\tfrac{\partial^2}{\partial\lambda^2}\mu_i(\lambda;\mathbf{p})\,\zeta_i^+(\lambda;\mathbf{p}) + 2\tfrac{\partial}{\partial\lambda}\mu_i(\lambda;\mathbf{p})\tfrac{\partial}{\partial\lambda}\zeta_i^+(\lambda;\mathbf{p})\right]\Big)\\ &+ o(\mathrm{e}^{\mu_i(\lambda;\mathbf{p})x}(1+x+x^2)),\end{aligned}$$

uniform for λ *in a compact set. Also, for* $x \to -\infty$ *we have*

$$\frac{\partial}{\partial\lambda}\widetilde{W}_i^-(x,\lambda;\mathbf{p}) = \left(-x\frac{\partial}{\partial\lambda}\mu_i(\lambda;\mathbf{p})\widetilde{\eta}_i^-(\lambda;\mathbf{p}) + \frac{\partial}{\partial\lambda}\widetilde{\eta}_i^-(\lambda;\mathbf{p})\right)\mathrm{e}^{-\mu_i(\lambda;\mathbf{p})x} + o(\mathrm{e}^{-\mu_i(\lambda;\mathbf{p})x}(1+x)) \tag{7.4}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial\lambda^2}\widetilde{W}_i^-(x,\lambda;\mathbf{p}) = \mathrm{e}^{-\mu_i(\lambda;\mathbf{p})x}\Big(&-x^2\left(\tfrac{\partial}{\partial\lambda}\mu_i(\lambda;\mathbf{p})\right)^2\widetilde{\eta}_i^-(\lambda;\mathbf{p}) + \tfrac{\partial^2}{\partial\lambda^2}\widetilde{\eta}_i^-(\lambda;\mathbf{p})\\ &- x\left[\tfrac{\partial^2}{\partial\lambda^2}\mu_i(\lambda;\mathbf{p})\,\widetilde{\eta}_i^-(\lambda;\mathbf{p}) + 2\tfrac{\partial}{\partial\lambda}\mu_i(\lambda;\mathbf{p})\tfrac{\partial}{\partial\lambda}\widetilde{\eta}_i^-(\lambda;\mathbf{p})\right]\Big)\\ &+ o(\mathrm{e}^{-\mu_i(\lambda;\mathbf{p})x}(1+x+x^2)),\end{aligned}$$

uniform for λ *in a compact set.*

Proof. By definition, for $1 \le i \le p$

$$U_i^+(x,\lambda;\mathbf{p})\mathrm{e}^{-\mu_i(\lambda;\mathbf{p})x} = \zeta_i^+(\lambda;\mathbf{p}) + o(1), \quad \text{for} \quad x \to \infty,$$

uniform for λ in a compact set. Hence we also have

$$\frac{\partial}{\partial\lambda}\left[U_i^+(x,\lambda;\mathbf{p})\mathrm{e}^{-\mu_i(\lambda;\mathbf{p})x}\right] = \frac{\partial}{\partial\lambda}\zeta_i^+(\lambda;\mathbf{p}) + o(1), \quad \text{for} \quad x \to \infty.$$

A straightforward calculation shows that

$$\frac{\partial}{\partial\lambda}\left[U_i^+(x,\lambda;\mathbf{p})\mathrm{e}^{-\mu_i(\lambda;\mathbf{p})x}\right] = \left(\frac{\partial}{\partial\lambda}U_i^+(x,\lambda;\mathbf{p}) - x\left(\frac{\partial}{\partial\lambda}\mu_i(\lambda;\mathbf{p})\right)U_i^+(x,\lambda;\mathbf{p})\right)\mathrm{e}^{-\mu_i(\lambda;\mathbf{p})x}.$$

Combining these facts we get the first expression in the lemma. The other expressions are verified in a similar way.

Lemma 9. *For any fixed* $\mathbf{p} = (\mathbf{a}, \mathbf{b}, c) \in \mathrm{A} \times \mathrm{B} \times \mathcal{C}$,

$$\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \boldsymbol{\Omega}_{11}(\lambda;\mathbf{p}) = \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{\boldsymbol{\Omega}}_{1j}(\lambda;\mathbf{p}) = \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{\boldsymbol{\Omega}}_{i1}(\lambda;\mathbf{p}) = 0, \quad \text{for } i,j \geq 2\,.$$

For $1 \leq i,j \leq q$, *define*

$$(\partial\mathcal{D})_{ij}(\mathbf{p}) = \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{\boldsymbol{\Omega}}_{i+1,j+1}(\lambda;\mathbf{p}) = \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \Omega(\widetilde{W}^-_{i+1}(x,\lambda;\mathbf{p}), U^+_{j+1}(x,\lambda;\mathbf{p}))\,;$$

then,

$$(\partial\mathcal{D})_{ij}(\mathbf{p}) = \sum_{k=1}^{q}\left[\chi^+_{kj}(\mathbf{p})\Omega(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\widetilde{\eta}^-_{i+1}(\lambda;\mathbf{p}), \xi_k(Z^-_0(\mathbf{p}))) + \chi^-_{ki}(\mathbf{p})\Omega(\xi_k(Z^+_0(\mathbf{p})), \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\zeta^+_{j+1}(\lambda;\mathbf{p}))\right].$$

Proof. First note that the derivative formula in Lemma 5 also holds if we replace $\boldsymbol{\Omega}(\lambda;\mathbf{p})$ by $\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p})$ (by replacing each W_i^- with $\widetilde{W}_i^-$). Each term in the integrand on the right-hand side of (5.30) is of the form $\omega(\widetilde{W}_i^-, U_j^+)$. We will show that each of these terms vanish when $\lambda = 0$. Define

$$\chi^+_{00} = 1 \quad \text{and} \quad \chi^\pm_{0i} = \chi^\pm_{j0} = 0\,, \quad \text{for } 1 \leq i,j \leq q\,.$$

(Recall that χ^-_{00} is defined in §3.) Then (note that $q = p-1$)

$$\widetilde{W}^-_{i+1}(x,0;\mathbf{p}) = \sum_{k=0}^{q} \chi^-_{ki}(\mathbf{p})\xi_k(\widetilde{Z}(x;\mathbf{p})) \quad \text{for } i = 0,\ldots,q\,,$$

and

$$U^+_{j+1}(x,0;\mathbf{p}) = \sum_{l=0}^{q} \chi^+_{lj}(\mathbf{p})\xi_l(\widetilde{Z}(x;\mathbf{p})) \quad \text{for } j = 0,\ldots,q\,.$$

Hence for $0 \leq i,j \leq q$

$$\begin{aligned}\omega(\widetilde{W}^-_{i+1}(x,\lambda;\mathbf{p}), U^+_{j+1}(x,\lambda;\mathbf{p}))\Big|_{\lambda=0} &= \sum_{k=0}^{q}\sum_{l=0}^{q} \chi^-_{ki}(\mathbf{p})\chi^+_{lj}(\mathbf{p})\omega(\xi_k(\widetilde{Z}(x;\mathbf{p})), \xi_l(\widetilde{Z}(x;\mathbf{p})))\\ &= \sum_{k=1}^{q}\sum_{l=1}^{q} \chi^-_{ki}(\mathbf{p})\chi^+_{lj}(\mathbf{p})\langle\, \xi_l(\widetilde{Z}(x;\mathbf{p})), \nabla P_k(\widetilde{Z}(x;\mathbf{p}))\rangle\\ &\quad + \chi^-_{00}(\mathbf{p})\omega(\widetilde{Z}_x(x;\mathbf{p}), \widetilde{Z}_x(x;\mathbf{p}))\,.\end{aligned}$$

But, $\omega(\widetilde{Z}_x, \widetilde{Z}_x) = 0$ by skew-symmetry and for each k,l, $\langle \xi_l(\widetilde{Z}), \nabla P_k(\widetilde{Z})\rangle$ vanishes because the group $\mathcal{G}$ is Abelian and so the functionals P_k are invariant under the action of $\mathcal{G}$. Therefore since the integrand vanishes for each fixed x, $\int_{-\infty}^{\infty} \omega(\widetilde{W}^-_{i+1}, U^+_{j+1})\Big|_{\lambda=0} \mathrm{d}x = 0$.

Next we consider the boundary terms $BC(R,S)$ in (5.30) in the limit as $R,S \to \infty$. For $0 \leq j \leq q$ and $\lambda = 0$, $\frac{\partial}{\partial\lambda}\widetilde{W}^-_{j+1}(x,\lambda;\mathbf{p})$ is not exponentially growing for $x \to -\infty$ and $U^+_1(x,\lambda;\mathbf{p}) = \widetilde{Z}_x(x;\mathbf{p})$ is exponentially decaying for $x \to -\infty$. Hence at $\lambda = 0$,

$$\lim_{x\to-\infty} \Omega\left(\frac{\partial}{\partial\lambda}\widetilde{W}^-_{j+1}(x,\lambda;\mathbf{p}), U^+_1(x,\lambda;\mathbf{p})\right) = 0\,.$$

Also, $\lim_{x\to\infty} \Omega(W^-_{j+1}(x,\lambda;\mathbf{p}), \frac{\partial}{\partial\lambda}U^+_1(x,\lambda;\mathbf{p})) = 0$, since $\frac{\partial}{\partial\lambda}U^+_1(x,\lambda;\mathbf{p})$ is exponentially decaying and $\widetilde{W}^-_{j+1}(x,\lambda;\mathbf{p}) = \sum_{k=0}^{q}\chi^-_{kj}(\mathbf{p})\xi_k(\widetilde{Z}(x;\mathbf{p}))$ is bounded for $x \to \infty$ when $\lambda = 0$.

In a similar way we can show that the boundary terms for the case $1 \leq i \leq p$ and $j = 1$ vanish. This completes the proof of the first three expressions in the lemma.

To prove the expression for the $q \times q$ matrix $\partial\mathcal{D}$, we first use Lemma 8 to show that, for $x \to -\infty$ and $1 \le i, j \le q$,

$$\begin{aligned} \Omega(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{W}^-_{i+1}(x,\lambda;\mathbf{p}), U^+_{j+1}(x,0;\mathbf{p})) &= -x\, \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \mu_{i+1}(\lambda;\mathbf{p}) \Omega(\widetilde{W}^-_{i+1}(x,0;\mathbf{p}), U^+_{j+1}(x,0;\mathbf{p})) \\ &\quad + \Omega(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{\eta}^-_{i+1}(\lambda;\mathbf{p}), \sum_{k=1}^q \chi^+_{kj}(\mathbf{p})\xi_k(Z^-_0(\mathbf{p}))) + o(1) \\ &= -x\, \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \mu_{i+1}(\lambda;\mathbf{p}) \boldsymbol{\Omega}_{(i+1)(j+1)}(0;\mathbf{p}) \\ &\quad + \sum_{k=1}^q \chi^+_{kj}(\mathbf{p}) \Omega(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{\eta}^-_{i+1}(\lambda;\mathbf{p}), \xi_k(Z^-_0(\mathbf{p}))) + o(1). \end{aligned}$$

Here we have used that $\mu_{i+1}(0;\mathbf{p}) = 0$. In Lemma 7 it is shown that $\boldsymbol{\Omega}_{(i+1)(j+1)}(0;\mathbf{p}) = 0$, hence

$$\lim_{x\to-\infty} \Omega(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\widetilde{W}^-_{i+1}(x,\lambda;\mathbf{p}), U^+_{j+1}(x,0;\mathbf{p})) = \textstyle\sum_{k=1}^q \chi^+_{kj}(\mathbf{p})\Omega(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{\eta}^-_{i+1}(\lambda;\mathbf{p}), \xi_k(Z^-_0(\mathbf{p}))).$$

The limit as $x \to +\infty$ can be derived in a similar way.

Apart from the asymptotic approximations in Lemma 8, we can also find explicit expression for the functions $\frac{\partial}{\partial\lambda}U^+_i(x,\lambda;\mathbf{p})$ and $\frac{\partial}{\partial\lambda}\widetilde{W}^-_j(x,\lambda;\mathbf{p})$ at $\lambda = 0$.

Lemma 10. *There exist constants $c^\pm_1(\mathbf{p})$ such that*

$$\begin{aligned} \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} U^+_1(x,\lambda;\mathbf{p}) &= -\tfrac{\partial}{\partial c}\widetilde{Z}(x;\mathbf{p}) + c^+_1(\mathbf{p})\widetilde{Z}_x(x;\mathbf{p}) \\ \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{W}^-_1(x,\lambda;\mathbf{p}) &= \chi^-_{00}(\mathbf{p})\tfrac{\partial}{\partial c}\widetilde{Z}(x;\mathbf{p}) + c^-_1(\mathbf{p})\widetilde{Z}_x(x;\mathbf{p}) \end{aligned} \tag{7.5}$$

For $1 \le i \le q$ there exist functions $y^\pm_{i+1}(x;\mathbf{p}) \in \ker[\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})]$ such that

$$\begin{aligned} \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} U^+_{i+1}(x,\lambda;\mathbf{p}) &= \sum_{k=1}^q \chi^+_{ki}(\mathbf{p})\frac{\partial}{\partial a_k}\widetilde{Z}(x;\mathbf{p}) + y^+_{i+1}(x;\mathbf{p}) \\ \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{W}^-_{i+1}(x,\lambda;\mathbf{p}) &= -\sum_{k=1}^q \chi^-_{ki}(\mathbf{p})\frac{\partial}{\partial a_k}\widetilde{Z}(x;\mathbf{p}) + y^-_{i+1}(x;\mathbf{p}). \end{aligned}$$

Proof. By substituting $\lambda = 0$ into (5.28) we get for $1 \le i \le p$

$$[\,\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})\,]\,\big(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} U^+_i(x,\lambda;\mathbf{p})\big) = -\mathbf{M}U^+_i(x,0;\mathbf{p}).$$

For $i = 1$ this gives

$$[\,\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})\,]\,\big(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} U^+_1(x,\lambda;\mathbf{p})\big) = -\mathbf{M}\widetilde{Z}_x(x;\mathbf{p}) \tag{7.6}$$

and for $i \ge 1$ we get

$$[\,\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})\,]\,\big(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} U^+_{i+1}(x,\lambda;\mathbf{p})\big) = -\mathbf{M}\sum_{k=1}^q \chi^+_{ki}(\mathbf{p})\xi_k(\widetilde{Z}(x;\mathbf{p})) = -\sum_{k=1}^q \chi^+_{ki}(\mathbf{p})\nabla P_k(\widetilde{Z}(x;\mathbf{p})).$$

Similarly, from (5.16) and (5.29) it follows $[\,\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})\,]\left(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{W}^-_1(x,\lambda;\mathbf{p})\right) = \chi^-_{00}(\mathbf{p})\mathbf{M}\widetilde{Z}_x(x;\mathbf{p})$ and for $i \ge 1$ $[\,\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})\,]\left(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{W}^-_{i+1}(x,\lambda;\mathbf{p})\right) = \sum_{k=1}^q \chi^-_{ki}(\mathbf{p})\nabla P_k(\widetilde{Z}(x;\mathbf{p}))$.

Furthermore, the existence of a multi-parameter family of travelling waves implies that we can differentiate the Euler-Lagrange equation with respect to c and $\mathbf{p}$. This gives

$$[\,\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})\,]\tfrac{\partial}{\partial c}\widetilde{Z}(x;\mathbf{p}) = \mathbf{M}\widetilde{Z}_x(x;\mathbf{p}) \tag{7.7}$$

and for $k \geq 1$

$$[\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})]\tfrac{\partial}{\partial a_k}\widetilde{Z}(x;\mathbf{p}) = -\nabla P_k(\widetilde{Z}(x;\mathbf{p}))\,. \tag{7.8}$$

Combining (7.7) with the expression (7.6) for $\frac{\partial}{\partial\lambda}\big|_{\lambda=0} U_1^+(x,\lambda;\mathbf{p})$, we get

$$[\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})]\left(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} U_1^+(x,\lambda;\mathbf{p}) + \tfrac{\partial}{\partial c}\widetilde{Z}(x;\mathbf{p})\right) = 0\,.$$

Hence $\frac{\partial}{\partial\lambda}\big|_{\lambda=0} U_1^+(x,\lambda;\mathbf{p}) = -\frac{\partial}{\partial c}\widetilde{Z}(x;\mathbf{p}) + y_1^+(x;\mathbf{p})$, where $y_1^+(x;\mathbf{p})$ is some element from the kernel of $[\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})]$. Both $\frac{\partial}{\partial\lambda}\big|_{\lambda=0} U_1^+(x,\lambda;\mathbf{p})$ and $\frac{\partial}{\partial c}\widetilde{Z}(x;\mathbf{p})$ are exponentially decaying for $x\to\infty$. By Hypothesis 3 of §3.5, the only exponentially decaying solution in the kernel of $[\mathbf{J}_c D_x - \mathbf{B}(x;\mathbf{p})]$ is $\widetilde{Z}_x$, hence there exists a constant $c_1^+(\mathbf{p})$ such that $y_1^+(x;\mathbf{p}) = c_1^+(\mathbf{p})\widetilde{Z}_x(x;\mathbf{p})$. In a similar way we can derive the other expressions in the lemma.

The remaining step, before we can start proving Theorem 2 is an analysis of $\frac{\partial^2}{\partial\lambda^2}\Big|_{\lambda=0} \boldsymbol{\Omega}_{11}(\lambda;\mathbf{p})$.

Lemma 11. *For any* $\mathbf{p}\in \mathrm{A}\times\mathrm{B}\times\mathcal{C}$,

$$\begin{aligned}\tfrac{\partial^2}{\partial\lambda^2}\Big|_{\lambda=0} \widetilde{\boldsymbol{\Omega}}_{11}(\lambda;\mathbf{p}) &= \tfrac{\partial^2}{\partial\lambda^2}\Big|_{\lambda=0} \Omega(\widetilde{W}_1^-(x,\lambda;\mathbf{p}), U_1^+(x,\lambda;\mathbf{p}))\\ &= 2\chi_{00}^-(\mathbf{p})\,\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \chi_{00}^-(\mathbf{p})\,\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p}))\end{aligned}$$

Proof. From (7.5) it follows

$$\begin{aligned}&\int_{-\infty}^{\infty}\omega\left(\widetilde{W}_1^-(x,0;\mathbf{p}),\, \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} U_1^+(x,\lambda;\mathbf{p})\right)\,\mathrm{d}x\\ &\qquad= \int_{-\infty}^{\infty}\omega\left(\chi_{00}^-(\mathbf{p})\,\widetilde{Z}_x(x;\mathbf{p}), -\tfrac{\partial}{\partial c}\widetilde{Z}(x;\mathbf{p}) + c_1^+\widetilde{Z}_x(x;\mathbf{p})\right)\,\mathrm{d}x\\ &\qquad= \chi_{00}^-(\mathbf{p})\int_{-\infty}^{\infty}\omega(\widetilde{Z}_c(x;\mathbf{p})\,,\,\widetilde{Z}_x(x;\mathbf{p}))\,\mathrm{d}x\\ &\qquad= \chi_{00}^-(\mathbf{p})\,\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \chi_{00}^-(\mathbf{p})\,\tfrac{1}{2}\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p})),\end{aligned}$$

using Lemma 1. In a similar way we can show that

$$\int_{-\infty}^{\infty}\omega\left(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\widetilde{W}_1^-(x,\lambda;\mathbf{p})\,,\,U_1^+(x,0;\mathbf{p})\right)\,\mathrm{d}x = \chi_{00}^-(\mathbf{p})\,\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \chi_{00}^-(\mathbf{p})\,\tfrac{1}{2}\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p})),$$

Finally we analyse the boundary terms. Using (7.5), we get

$$\begin{aligned}\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\,\Omega\left(\tfrac{\partial}{\partial\lambda}\widetilde{W}_1^-(x,\lambda;\mathbf{p}), U_1^+(x,\lambda;\mathbf{p})\right) &= \Omega\left(\tfrac{\partial^2}{\partial\lambda^2}\Big|_{\lambda=0}\widetilde{W}_1^-(x,\lambda;\mathbf{p})\,,\,\widetilde{Z}_x(x;\mathbf{p})\right)\\ &\quad+ \Omega(\chi_{00}^-(\mathbf{p})\tfrac{\partial}{\partial c}\widetilde{Z}(x;\mathbf{p}) + c_1^-\widetilde{Z}_x(x;\mathbf{p}), -\tfrac{\partial}{\partial c}\widetilde{Z}(x;\mathbf{p}) + c_1^+\widetilde{Z}_x(x;\mathbf{p}))\,.\end{aligned}$$

So for $x\to-\infty$ this expression vanishes. In a similar way we can show that the other boundary terms disappear.

Now we have all information needed to prove Theorem 2.

Proof of Theorem 2 Lemmas 7 and 9 imply that

$$\widetilde{\boldsymbol{\Omega}}_{ij}(0;\mathbf{p}) = \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\widetilde{\boldsymbol{\Omega}}_{ij}(\lambda;\mathbf{p}) = 0, \quad \text{for } i=1 \text{ or } j=1 \quad .$$

Hence we can conclude that for $\lambda\to 0$

$$\widetilde{\boldsymbol{\Omega}}_{ij}(\lambda;\mathbf{p}) = \mathcal{O}(\lambda^2), \quad \text{for } i=1 \text{ or } j=1.$$

Moreover, with Lemma 11 we have

$$\widetilde{\boldsymbol{\Omega}}_{11}(\lambda;\mathbf{p}) = \lambda^2\chi_{00}^-(\mathbf{p})\,\left[\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \tfrac{1}{2}\,\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p}))\right] + \mathcal{O}(\lambda^3), \quad \text{as } \lambda\to 0\,.$$

Furthermore, Lemma 7 and Lemma 9 show that $\widetilde{\boldsymbol{\Omega}}_{ij}(0;\mathbf{p}) = 0$ and $\frac{\partial}{\partial\lambda}\big|_{\lambda=0} \widetilde{\boldsymbol{\Omega}}_{ij}(\lambda;\mathbf{p}) = (\partial\mathcal{D})_{(i-1)(j-1)}(\mathbf{p})$, so altogether

$$\widetilde{\boldsymbol{\Omega}}_{ij}(\lambda;\mathbf{p}) = \lambda(\partial\mathcal{D})_{(i-1)(j-1)}(\mathbf{p}) + \mathcal{O}(\lambda^2), \quad \text{for } \lambda \to 0, \text{ if } i,j \geq 2.$$

So we are in the following situation:

$$\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p}) = \left(\begin{array}{c|c} \lambda^2\chi^-_{00}(\mathbf{p})\left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \frac{1}{2}\,\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p}))\right] + \mathcal{O}(\lambda^3) & \mathcal{O}(\lambda^2) \\ \hline \mathcal{O}(\lambda^2) & \lambda\partial\mathcal{D}(\mathbf{p}) + \mathcal{O}(\lambda^2) \end{array}\right)$$

$$= \lambda^p \left(\begin{array}{c|c} \lambda\chi^-_{00}(\mathbf{p})\left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \frac{1}{2}\,\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p}))\right] + \mathcal{O}(\lambda^2) & \mathcal{O}(\lambda) \\ \hline \mathcal{O}(\lambda) & \partial\mathcal{D}(\mathbf{p}) + \mathcal{O}(\lambda) \end{array}\right).$$

From the last part of this section (Lemma 12), it follows that the determinant of such a matrix is

$$\det(\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p})) = \lambda^{1+p}\,\chi^-_{00}(\mathbf{p})\,\det(\partial\mathcal{D}(\mathbf{p}))\,\left[\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x;\mathbf{p})) - \tfrac{1}{2}\,\omega(Z_0^+(\mathbf{p}),\partial_c Z_0^+(\mathbf{p}))\right] + \mathcal{O}(\lambda^{2+p}).$$

Using that by definition $\det(\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p})) = \lambda^{p-1}\det(\boldsymbol{\Omega}(\lambda;\mathbf{p}))$, we get the expressions for the derivatives of the determinant of the symplectic Evans matrix as stated. □

Now we provide a proof of the basic property needed to complete the last step in the proof of Theorem 2.

Lemma 12. *Let* $\mathrm{B} \in g\ell(m,\mathbb{C})$ *be a matrix depending analytically on* $\lambda \in \Lambda \subset \mathbb{C}$, *where* Λ *is an open simply-connected set with the origin as a limit point, with the form*

$$\mathrm{B}(\lambda) = \left(\begin{array}{c|c} \mathrm{B}^1(\lambda) & \mathrm{B}^2(\lambda) \\ \hline \mathrm{B}^3(\lambda) & \mathrm{B}^4(\lambda) \end{array}\right), \quad \textit{with} \quad \mathrm{B}^1(\lambda) \in g\ell(m_1,\mathbb{C}), \quad 0 < m_1 < m\,.$$

Suppose $\mathrm{B}^j(0) = 0$ *for* $j = 1,2,3$. *Then*

$$\det(\mathrm{B}(\lambda)) = \det(\mathrm{B}^1(\lambda))\det(\mathrm{B}^4(\lambda)) + \mathcal{O}(\lambda^{m_1+1}) \quad \text{as } \lambda \to 0\,.$$

Proof. To prove this statement, we will use induction on m. If $m = 2$, then $m_1 = 1$, hence all $\mathrm{B}^i(\lambda)$ are complex numbers and

$$\det(\mathrm{B}(\lambda)) = \mathrm{B}^1(\lambda)\mathrm{B}^4(\lambda) - \mathrm{B}^2(\lambda)\mathrm{B}^3(\lambda) = \det(\mathrm{B}^1(\lambda))\det(\mathrm{B}^4(\lambda)) + \mathcal{O}(\lambda^{1+1})\,.$$

Assume that the statement is true for some $m \geq 2$. Now consider some matrix $\mathrm{B} \in g\ell(m+1,\mathbb{C})$ which satisfies the above conditions. We use $[\mathrm{B}(\lambda)]_{1j}$ to denote the matrix $\mathrm{B}(\lambda)$ in $g\ell(m,\mathbb{C})$ with the first row and the jth column removed, and we use $\mathrm{B}_{ij}(\lambda)$ to denote the element of $\mathrm{B}(\lambda)$ in the ith row and jth column.

By definition we have

$$\det(\mathrm{B}(\lambda)) = \sum_{j=1}^{m}\mathrm{B}_{1j}(\lambda)\det([\mathrm{B}]_{1j}(\lambda)) = \sum_{j=1}^{m_1}\mathrm{B}^1_{1j}(\lambda)\det([\mathrm{B}]_{1j}(\lambda)) + \sum_{j=m_1+1}^{m}\mathrm{B}_{1j}(\lambda)\det([\mathrm{B}]_{1j}(\lambda)).$$

For any $m_1 + 1 \leq j \leq m$, the matrix $([\mathrm{B}]_{1j}(\lambda))$ satisfies the induction assumption. Hence for $m_1 + 1 \leq j \leq m$, we have

$$\det([\mathrm{B}]_{1j})(\lambda) = \det([\mathrm{B}^1]_{1m_1}(\lambda))\det(\tilde{\mathrm{B}}^4(\lambda)) + \mathcal{O}(\lambda^{(m_1-1)+1})\,,$$

where $\tilde{\mathrm{B}}^4(\lambda)$ is the matrix $\mathrm{B}^4(\lambda)$ minus the (m_1-j)th column, plus the m_1th column of $\mathrm{B}^3(\lambda)$.

Since $(\mathrm{B}^1)_{ij} = \mathcal{O}(\lambda)$ and $(\tilde{B}^4)_{i1}(\lambda) = \mathcal{O}(\lambda)$, we see that for $m_1 + 1 \le j \le m$ we have

$$\det([\mathrm{B}]_{1j}(\lambda)) = \mathcal{O}(\lambda^{(m_1-1)+1}) = \mathcal{O}(\lambda^{m_1}) .$$

Using that $\mathrm{B}_{1j}(\lambda) = \mathrm{B}^2_{1(j-m_1)}(\lambda) = \mathcal{O}(\lambda)$, we get

$$\sum_{j=m_1+1}^{m} \mathrm{B}_{1j}(\lambda)\det([\mathrm{B}]_{1j}(\lambda)) = \mathcal{O}(\lambda^{m_1+1}) .$$

Furthermore, if $m_1 = 1$, then

$$\sum_{j=1}^{m_1} \mathrm{B}^1_{1j}(\lambda)\det([\mathrm{B}]_{1j}(\lambda)) = \det(\mathrm{B}^1(\lambda))\det(\mathrm{B}^4(\lambda)) .$$

And if $m_1 > 1$, then for any $1 \le j \le m$ the matrix $([\mathrm{B}]_{1j}(\lambda))$ satisfies the induction hypothesis. Hence for $1 \le j \le m_1$, we have

$$\det([\mathrm{B}]_{1j}(\lambda)) = \det([\mathrm{B}^1]_{1j}(\lambda))\det(\mathrm{B}^4(\lambda)) + \mathcal{O}(\lambda^{(m_1-1)+1}) .$$

So altogether we get

$$\begin{aligned}\det(\mathrm{B}(\lambda)) &= \textstyle\sum_{j=1}^{m_1} \mathrm{B}^1_{1j}(\lambda)\det([\mathrm{B}^1]_{1j}(\lambda))\det(\mathrm{B}^4(\lambda)) + \mathcal{O}(\lambda^{m_1+1})\\ &= \det(\mathrm{B}^1(\lambda))\det(\mathrm{B}^4(\lambda)) + \mathcal{O}(\lambda^{m_1+1}) .\end{aligned}$$

which is the desired statement, completing the proof.

8. Asymptotics of the Evans matrix for large $|\lambda|$

If the time-dependent PDE is well-posed as an evolution equation, we would expect the Evans function to be non-zero when $\Re(\lambda)$ is large. In other words, for a well-posed PDE, there are no unstable eigenvalues with arbitrarily large growth rates.

The Evans function is real when λ is real. Therefore for λ real, positive and large, we can expect the Evans function for a well-posed PDE to be nonzero and of one sign. There are several results in the literature which prove this property, with varying hypotheses and degree of generality. For a class of parabolic PDEs, asymptotic results on the Evans function for large $|\lambda|$ are proved in Alexander et al. ([2], Proposition 2.2), and this result is substantially generalized in Gardner & Zumbrun ([23],§2.2). For Hamiltonian PDEs – with the spectral ODEs having the property $p = 1$ – asymptotic results are proved in Pego & Weinstein ([42], Proposition 1.17 and Corollary 1.18). For a large class of parameter-dependent parabolic and nonlinear Schrödinger type equations, Kapitula & Sandstede [30] prove results on the large $|\lambda|$ behavior of the Evans functions, including uniformity with respect to parameters. However, none of the above results applies to Hamiltonian systems with $p > 1$ which are not necessarily of the "quasi-parabolic" or nonlinear Schrödinger type.

The most natural generalization of the above results we need is a symplectification and generalizations of Proposition 1.17 of [42]. In this secttion, we simplify their proof and abstract it to the case $p = 2$. Further generalization is possible, but this is the relevant case for the examples in §§10-12. The parameter dependence (other than λ) is not important in this section and is therefore suppressed to simplify notation.

Restrict λ to be real and positive. To analyse the Evans matrix for large λ, we introduce a small parameter $\varepsilon = \lambda^{-\alpha}$ with $\alpha > 0$. The value of α will depend on the problem we are looking at; in the examples in this paper, $\alpha = \frac{1}{2}$.

The strategy is to first deduce the asymptotics for large λ for the functions $U(x,\lambda)$ and $W(x,\lambda)$ which satisfy the linear spectral ODE and its adjoint. Introduce a new variable $s=\varepsilon^{-\gamma}x$ and a selfadjoint transformation $T(\varepsilon)$, which rescales the entries in the vector-valued functions $U(x,\lambda)$ and $W(x,\lambda)$. If we define

$$U^{\rm new}(s)=T(\varepsilon)U(\varepsilon^{\gamma}s) \quad \text{and} \quad W^{\rm new}(s)=T(\varepsilon)W(\varepsilon^{\gamma}s)\,,$$

the linear problem (3.1) and the adjoint problem (3.22) are transformed to

$$\lambda T(\varepsilon)\,\left[\mathbf{J}_c^{\rm new}(\varepsilon)\tfrac{\rm d}{{\rm d}s}-\mathbf{B}^{\rm new}(\varepsilon^{\gamma}s,\varepsilon)+\mathbf{M}^{\rm new}(\varepsilon)\right]U^{\rm new}(s)=0, \tag{8.1}$$

$$\lambda T(\varepsilon)\,\left[\mathbf{J}_c^{\rm new}(\varepsilon)\tfrac{\rm d}{{\rm d}s}-\mathbf{B}^{\rm new}(\varepsilon^{\gamma}s,\varepsilon)-\mathbf{M}^{\rm new}(\varepsilon)\right]W^{\rm new}(s)=0, \tag{8.2}$$

where

$$\mathbf{J}_c^{\rm new}(\varepsilon)=\varepsilon^{-\gamma+1/\alpha}T(\varepsilon)^{-1}\mathbf{J}_cT(\varepsilon)^{-1}, \quad \mathbf{M}^{\rm new}(\varepsilon)=T(\varepsilon)^{-1}\mathbf{M}T(\varepsilon)^{-1}$$

and

$$\mathbf{B}^{\rm new}(x,\varepsilon)=\varepsilon^{1/\alpha}T(\varepsilon)^{-1}\mathbf{B}(x)T(\varepsilon)^{-1}.$$

The parameter γ is chosen such that the largest eigenvalues of the scaled linear systems at infinity are of order 1 (in our examples we find that $\gamma=1$), the transformation $T(\varepsilon)$ is required to be such that $\mathbf{J}_c^{\rm new}(0)=\lim\limits_{\varepsilon\to 0}\mathbf{J}_c^{\rm new}(\varepsilon)$ exists and is invertible, and the matrices $\mathbf{B}^{\rm new}(x,0)=\lim\limits_{\varepsilon\to 0}\mathbf{B}^{\rm new}(x,\varepsilon)$ and $\mathbf{M}^{\rm new}(0)=\lim\limits_{\varepsilon\to 0}\mathbf{M}^{\rm new}(\varepsilon)$ exist.

Recall that the eigenvalues of the unscaled systems at infinity are μ_i, $i=1,\ldots,2n$ (cf. §3.1). For the scaled systems, the eigenvalues of the systems at infinity (i.e. $|s|\to\infty$) related to (8.1) are $\nu_i=\varepsilon^{\gamma}\mu_i$ and the eigenvectors are $\theta_i^{\pm}=T(\varepsilon)\zeta_i^{\pm}$. Similarly, in the new adjoint asymptotic system, the eigenvalues and eigenvectors are $-\nu_i$ and $\psi_j^{\pm}=T(\varepsilon)\eta_j^{\pm}$ respectively.

The systems (8.1) and (8.2) are both cast the abstract form

$$\tfrac{\rm d}{{\rm d}s}V=\left[\,\mathbf{F}(\varepsilon)+\varepsilon^{\kappa+\gamma}\mathbf{G}(\varepsilon^{\gamma}s,\varepsilon)\,\right]V, \quad V\in\mathbb{C}^{2n} \tag{8.3}$$

for some $\kappa\in\mathbb{R}$ (in our examples we have that $\kappa=1$) and the matrix $\mathbf{G}$ has the property that $\int_0^{\infty}|\mathbf{G}(x,\varepsilon)|{\rm d}x$ or $\int_{-\infty}^{0}|\mathbf{G}(x,\varepsilon)|{\rm d}x$ is bounded, uniform in ε. To be explicit,

$$\mathbf{F}(\varepsilon)=\mathbf{J}_c^{\rm new}(\varepsilon)^{-1}\,[\,\mp\mathbf{M}^{\rm new}(\varepsilon)+\lim_{x\to\pm\infty}\mathbf{B}^{\rm new}(x,\varepsilon)\,],$$

and

$$\mathbf{G}(x,\varepsilon)=\varepsilon^{-\kappa-\gamma}\mathbf{J}_c^{\rm new}(\varepsilon)^{-1}\,[\,\mathbf{B}^{\rm new}(x,\varepsilon)-\lim_{x\to\pm\infty}\mathbf{B}^{\rm new}(x,\varepsilon)\,].$$

Thus $\lim\limits_{x\to\pm\infty}\mathbf{G}(x,\varepsilon)=0$. Furthermore, all matrices depend continuously (analytically if $\frac{1}{\alpha}$ and γ are integers) on ε, for ε small. If $p=2$ and $\mu_1=\overline{\mu_2}$ for large λ, then the following lemma describes the behaviour of the exponentially decaying solutions of (8.3) in the asymptotic limit $\varepsilon\to 0$.

Lemma 13. *For ε small, let $\mathbf{F}(\varepsilon)$ be a matrix with (semi-)simple eigenvalues $\nu_i(\varepsilon)$, with $\nu_2(\varepsilon)=\overline{\nu_1(\varepsilon)}$, $\Re(\nu_1(\varepsilon))=\Re(\nu_2(\varepsilon))<0$ and $\Re(\nu_i(\varepsilon))\geq 0$ for $i>2$. We denote the eigenvectors corresponding to $\nu_i(\varepsilon)$ by $\theta_i(\varepsilon)$, $i=1,2$. For $i=1,2$, we denote the solution $V(s,\varepsilon)$ of (8.3) with $\lim\limits_{s\to\infty}{\rm e}^{-\nu_i s}V(s,\varepsilon)=\theta_i(\varepsilon)$ by $V_i(s,\varepsilon)$.*

If $\int_0^{\infty}|\mathbf{G}(x,\varepsilon)|{\rm d}x$ is bounded, uniform in ε, then

$$V_i(0,\varepsilon)=\theta_i(\varepsilon)+\mathcal{O}(\varepsilon^{\kappa})|\theta_i(\varepsilon)|, \quad \textit{for } i=1,2 \quad .$$

Proof. Let $i=1,2$. Define $v_i(x,\varepsilon)=\mathrm{e}^{-\nu_i(\varepsilon)x}V_i(x,\varepsilon)-\theta_i(\varepsilon)$. Then $\lim_{x\to\infty} v_i(x,\varepsilon)=0$ and $v_i(x,\varepsilon)$ satisfies the following inhomogeneous linear differential equation

$$-\tfrac{\mathrm{d}}{\mathrm{d}s}v_i(s,\varepsilon)+(\mathbf{F}-\nu_i\,\mathbf{I})v_i(s,\varepsilon)+\varepsilon^{\gamma+\kappa}\mathbf{G}(\varepsilon^\gamma s,\varepsilon)(v_i(s,\varepsilon)+\theta_i)=0,$$

hence

$$v_i(s,\varepsilon)=-\varepsilon^{\gamma+\kappa}\int_s^\infty \mathrm{e}^{\mathbf{F}_i(\varepsilon)(s-t)}\mathbf{G}(\varepsilon^\gamma t,\varepsilon)(v_i(t,\varepsilon)+\theta_i(\varepsilon))\,\mathrm{d}t\,, \tag{8.4}$$

where $\mathbf{F}_i(\varepsilon)=\mathbf{F}(\varepsilon)-\nu_i(\varepsilon)\,\mathbf{I}$. Note that all the eigenvalues of $\mathbf{F}_i(\varepsilon)$ have non-negative real part, since $\nu_2=\overline{\nu_1}$ for ε small. Therefore, there is some $C>0$, such that $|\mathrm{e}^{(s-t)\mathbf{F}_i(\varepsilon)}|<C$ for $t\geq s$, uniform in ε, and (8.4) leads immediately to the following estimate for $|v_i(s,\varepsilon)|$

$$|v_i(s,\varepsilon)|\leq\varepsilon^{\gamma+\kappa}\,C\int_s^\infty|\mathbf{G}(\varepsilon^\gamma\tau,\varepsilon)(v_i(\tau,\varepsilon)+\theta_i(\varepsilon))|\,\mathrm{d}\tau\,.$$

Since $\int_0^\infty|\mathbf{G}(x,\varepsilon)|\mathrm{d}x$ is bounded, we can find some $K>0$ such that $C\int_0^\infty|\mathbf{G}(x,\varepsilon)|\,\mathrm{d}x\leq K$, uniform in ε. Together with the above estimate, this gives

$$|v_i(s,\varepsilon)|\leq\varepsilon^\kappa K\sup_{x\geq0}|v_i(x)|+\varepsilon^\kappa K|\theta_i|.$$

for all $s\geq0$, and so for ε sufficiently small

$$\sup_{x\geq0}|v_i(x)|\leq2\varepsilon^\kappa K|\theta_i|,$$

uniform in ε. Thus also, $|V_i(0,\varepsilon)-\theta_i(\varepsilon)|=|v_i(0,\varepsilon)|=\mathcal{O}(\varepsilon^\kappa)|\theta_i(\varepsilon)|$, for $\varepsilon\to0$, hence

$$V_i(0,\varepsilon)=\theta_i(\varepsilon)+\mathcal{O}(\varepsilon^\kappa)|\theta_i(\varepsilon)|.$$

With this lemma, we can determine the asymptotic behaviour of the entries of the symplectic Evans matrix for large λ , and hence its determinant.

Theorem 3. *Consider a transformation as described in* (8.1) *and* (8.2)*, such that* $\mathbf{J}_c^{\mathrm{new}}(0)=\lim_{\varepsilon\to0}\mathbf{J}_c^{\mathrm{new}}(\varepsilon)$ *exists and is invertible and* $\mathbf{B}^{\mathrm{new}}(x,0)=\lim_{\varepsilon\to0}\mathbf{B}^{\mathrm{new}}(x,\varepsilon)$, $\mathbf{M}^{\mathrm{new}}(0)=\lim_{\varepsilon\to0}\mathbf{M}^{\mathrm{new}}(\varepsilon)$ *both exist. Assume that there is some* $\kappa>0$ *such that* $\mathbf{B}^{\mathrm{new}}(x,\varepsilon)-\lim_{x\to\infty}\mathbf{B}^{\mathrm{new}}(x,\varepsilon)=\mathcal{O}(\varepsilon^{\kappa+\gamma})$*, uniform for* $x\in\mathbb{R}$ *and that* $T(\varepsilon)\zeta_i^+(\varepsilon^{-\frac{1}{\alpha}})$ *are of the same order for* $i=1,2$*. If* $p=2$*, then the determinant of the symplectic Evans matrix satisfies*

$$\det\boldsymbol{\Omega}(\lambda)=1+\mathcal{O}(\lambda^{-\alpha\kappa}+\lambda^{-\alpha}),\qquad\lambda\to\infty.$$

Proof. For $i=1,2$, define $U_i^{\mathrm{new},+}(s,\varepsilon)=T(\varepsilon)U_i^+(\varepsilon^\gamma s)$, $\theta_i^+(\varepsilon)=T(\varepsilon)\zeta_i^+$ and $W_i^{\mathrm{new},-}(s,\varepsilon)=T(\varepsilon)W_i^-(\varepsilon^\gamma s)$, $\psi_i^-(\varepsilon)=T(\varepsilon)\eta_i^-$. Then $U_i^{\mathrm{new},+}(s,\varepsilon)$ and $W_i^{\mathrm{new},-}(s,\varepsilon)$ are solutions of (8.1) and (8.2) respectively with

$$\lim_{s\to\infty}\mathrm{e}^{\nu_i(\varepsilon)}U_i^{\mathrm{new},+}(s,\varepsilon)=\theta_i^+(\varepsilon)\quad\text{and}\quad\lim_{s\to-\infty}\mathrm{e}^{\nu_i(\varepsilon)}W_i^{\mathrm{new},-}(s,\varepsilon)=\psi_i^-(\varepsilon)\,,$$

where $\nu_i(\varepsilon)=\varepsilon^\gamma\mu_i$. From Lemma 13, it follows that

$$U_i^{\mathrm{new},+}(0,\varepsilon)=\theta_i^+(\varepsilon)+\mathcal{O}(\varepsilon^\kappa)|\theta_i(\varepsilon)|\quad\text{and}\quad W_i^{\mathrm{new},-}(0,\varepsilon)=\psi_i^-(\varepsilon)+\mathcal{O}(\varepsilon^\kappa)|\psi_i^-(\varepsilon)|.$$

Hence we get for the entries in the Evans matrix: for each $i,j=1,2$,

$$\begin{aligned}\Omega(W_j^-(0),U_i^+(0))&=\Omega(T(\varepsilon)^{-1}W_j^{\mathrm{new},-}(0,\varepsilon),T(\varepsilon)^{-1}U_i^{\mathrm{new},+}(0,\varepsilon))\\&=\Omega(T(\varepsilon)^{-1}\psi_j^-(\varepsilon),T(\varepsilon)^{-1}\theta_i^+(\varepsilon))\\&\qquad\qquad+\mathcal{O}(\varepsilon^{\kappa+\gamma-1/\alpha})|\mathbf{J}_c^{\mathrm{new}}(\varepsilon)||\theta_i^+(\varepsilon)||\psi_j^-(\varepsilon)|\\&=\varepsilon^{\gamma-1/\alpha}\left[\langle\mathbf{J}_c^{\mathrm{new}}(\varepsilon)\psi_j^-(\varepsilon)\,,\,\theta_i^+(\varepsilon)\rangle+\mathcal{O}(\varepsilon^\kappa)|\theta_i^+(\varepsilon)||\psi_j^-(\varepsilon)|\right]\,.\end{aligned}\tag{8.5}$$

Taking the limit as $\varepsilon \to 0$, we see that (8.1) and (8.2) converge pointwise to an autonomous system of the form

$$[\mathbf{J}_c^{\mathrm{new}}(0)\frac{\mathrm{d}}{\mathrm{d}s} - \mathbf{B}^{\mathrm{new}}(0,0) \pm \mathbf{M}^{\mathrm{new}}(0)]V = 0.$$

Let $\overline{\theta_i^0}$ and $\overline{\psi_j^0}$ be the normalised eigenvectors of the asymptotic system and its adjoint respectively. Because of the symplectic definition of the adjoint equation, we have that $\langle \mathbf{J}_c^{\mathrm{new}}(0)\overline{\psi_j^0}\,,\,\overline{\theta_i^0}\rangle = \delta_{ij}\,K_i$, with K_i being some nonzero constants.

All matrices depend continuously on ε (for ε small), hence

$$\theta_i^{\pm}(\varepsilon) = |\theta_i^{\pm}(\varepsilon)|(\overline{\theta_i^0} + \mathcal{O}(\varepsilon))\,, \quad \psi_j^{\pm}(\varepsilon) = |\psi_j^{\pm}(\varepsilon)|(\overline{\psi_j^0} + \mathcal{O}(\varepsilon))\,,$$

and $\mathbf{J}_c^{\mathrm{new}}(\varepsilon) = \mathbf{J}_c^{\mathrm{new}}(0) + \mathcal{O}(\varepsilon)$.

Combining these asymptotic results with equation (8.5), we get

$$\Omega(W_j^-(0), U_i^+(0)) = \varepsilon^{\gamma-1/\alpha}\,|\theta_i^+(\varepsilon)||\psi_j^-(\varepsilon)|\;[\,K_i\,\delta_{ij} + \mathcal{O}(\varepsilon^\kappa + \varepsilon)\,]\;. \tag{8.6}$$

Finally we look at $|\theta_i^+(\varepsilon)||\psi_j^-(\varepsilon)|$. By the normalisation of the vectors $\zeta_i^{\pm}$ and $\eta_j^{\pm}$ and using the ε asymptotics of $\theta_i^{\pm}(\varepsilon)$ and $\psi_j^{\pm}(\varepsilon)$, we have

$$\delta_{ij} = \Omega(\eta_i^{\pm}, \zeta_j^{\pm}) = \varepsilon^{\gamma-1/\alpha}\langle \mathbf{J}_c^{\mathrm{new}}(\varepsilon)\psi_j^{\pm}(\varepsilon)\,,\,\theta_i^{\pm}(\varepsilon)\rangle = \varepsilon^{\gamma-1/\alpha}\,|\theta_i^{\pm}(\varepsilon)||\psi_j^{\pm}(\varepsilon)|\,[\delta_{ij}\,K_i + \mathcal{O}(\varepsilon)].$$

In other words, $|\theta_i^{\pm}(\varepsilon)||\psi_i^{\pm}(\varepsilon)| = \varepsilon^{1/\alpha-\gamma}\,[1/K_i + \mathcal{O}(\varepsilon)]$.

Combining this expression with (8.6) and using the assumption that $\theta_i^{\pm}(0)$ are of the same order, we can conclude that $\Omega(W_j^-(0), U_i^+(0)) = \delta_{ij} + \mathcal{O}(\varepsilon^\kappa + \varepsilon)$, for $i,j \in \{1,2\}$. Now, using the definition of the Evans matrix and the relation between λ and ε, we get the result as stated in the theorem.

9. Intermezzo on analyticity of the Evans matrix

In constructing an Evans function framework for Hamiltonian PDEs, there are two competing issues: analyticity and symplecticity. Applying Hodge duality to the exterior algebra representation (5.10) of the Evans function lead to a dual symplectic framework for the Evans function. However, the symplectic Evans matrix involves individual vector-valued solutions of the linear system which are analytic for all $\lambda \in \Lambda$, but they are not in general analytic for all $\lambda \in \mathbb{C}_+$. In particular, it may not be analytic at those values of $\lambda \in \mathbb{C}$ where $\mathbf{A}^{\pm}(\lambda;\cdot)$ has a multiple eigenvalue.

On the other hand, when the Evans function is formulated as a wedge product between forms $Y^+(x,\lambda;\cdot) \in \bigwedge^p(\mathbb{C}^{2n})$ and $Y^-(x,\lambda;\cdot) \in \bigwedge^{2n-p}(\mathbb{C}^{2n})$ as in (5.10), the Evans function *is analytic* at these distinguished values of λ (cf. [2,23]). To see this, let Λ^+ be any simply-connected subset of $\mathbb{C}_+$ that includes Λ and may also include values of $\lambda \in \mathbb{C}_+$ where $\mathbf{A}^{\pm}(\lambda;\cdot)$ has multiple eigenvalues, but retaining the condition that there is a split of the spectrum with p eigenvalues having negative real part and $2n-p$ eigenvalues with non-negative real part (see [23,30] for more general conditions for this splitting). Then $Y^+(x,\lambda;\cdot) \in \bigwedge^p(\mathbb{C}^{2n})$ and $Y^-(x,\lambda;\cdot) \in \bigwedge^{2n-p}(\mathbb{C}^{2n})$ are analytic functions of λ for all $\lambda \in \Lambda^+$. This follows since

$$\lim_{x\to+\infty} \mathrm{e}^{-\sigma_+(\lambda;\cdot)x}Y^+(x,\lambda;\cdot) = Y_\infty^+(\lambda;\cdot) \in \textstyle\bigwedge^p(\mathbb{C}^{2n})\,,$$

and

$$\lim_{x\to-\infty} \mathrm{e}^{-\sigma_-(\lambda;\cdot)x}Y^-(x,\lambda;\cdot) = Y_\infty^-(\lambda;\cdot) \in \textstyle\bigwedge^{2n-p}(\mathbb{C}^{2n})\,.$$

Both $\sigma_+(\lambda;\cdot)$ and $\sigma_-(\lambda;\cdot)$ are analytic functions for all $\lambda \in \Lambda^+$, because $\sigma_+(\lambda;\cdot)$ and $\sigma_-(\lambda;\cdot)$ are *simple* eigenvalues of $\mathbf{A}^{\pm}(\lambda;\cdot)$ restricted to $\bigwedge^p(\mathbb{C}^{2n})$ and $\bigwedge^{2n-p}(\mathbb{C}^{2n})$ respectively, and the limits are uniform on compact subsets of Λ^+ (cf. [42]).

The left-hand side of (5.25) is in general analytic, and so the right-hand side is analytic, but it is a product of two functions each of which may not be analytic for all $\lambda \in \Lambda^+$. The part of the proof of Lemma 4 where analyticity may be lost (if the enlarged set Λ^+ is used instead of Λ) is when a basis is extracted for the p-plane characterized by the form $Y^+(x,\lambda;\cdot)$. Implicit in the evaluation of the inner product on $\bigwedge^p(\mathbb{C}^{2n})$ in (5.22) is the extraction of a basis which depends analytically on λ.

In other words, $\mathcal{Y}^-(x,\lambda;\cdot) \wedge Y^+(x,\lambda;\cdot)$ in (5.22) is an analytic function of λ for all $\lambda \in \Lambda^+$, whereas, *a priori*, it is not immediate that $\boldsymbol{\Omega}(\lambda;\cdot)$ can be analytically continued from Λ to Λ^+. A basic question related to this is: given a decomposable form, say $Y^+(x,\lambda;\cdot) \in \bigwedge^p(\mathbb{C}^{2n})$, which is analytic for all $\lambda \in \Lambda^+$, can a basis for the p-plane characterized by $Y^+(x,\lambda;\cdot)$ be constructed which is also analytic for all $\lambda \in \Lambda^+$? If so, then the Evans matrix *constructed with this basis* is analytic for all $\lambda \in \Lambda^+$. However, this does not imply that $\boldsymbol{\Omega}(\lambda;\cdot)$ is analytic for all $\lambda \in \Lambda^+$, because $\boldsymbol{\Omega}(\lambda;\cdot)$ is constructed using a particular subset of solutions of the adjoint equation. Analyticity of $\boldsymbol{\Omega}(\lambda;\cdot)$ for all $\lambda \in \Lambda^+$ would require further that the matrix $\mathbf{E}(\lambda;\cdot)$ defined in (5.24) also be analytic for all $\lambda \in \Lambda^+$.

Although these issues regarding analyticity are important and interesting, we will see in the examples that the distinction between the sets Λ and Λ^+ is not important – at least for the analysis of *instability*. In examples the distinguished points $\lambda \in \mathbb{C}$ where $\mathbf{A}^{\pm}(\lambda;\cdot)$ has multiple eigenvalues are isolated branch points and can be easily excised (see Figures 10.2 and 10.3 for example). However, for consideration of *stability* these distinguished points have to be studied to ensure they don't correspond to unstable eigenvalues, although in general it would be rare – but not impossible – for such points to correspond to unstable eigenvalues.

On the other hand, an interesting approach to maintaining symplecticity and analyticity is to work with a symplectictification of the forms $Y^+(x,\lambda;\cdot)$ and $Y^-(x,\lambda;\cdot)$. In other words the symplectic structures ω, κ and Ω induce *pre-symplectic* structures on $\bigwedge^p(\mathbb{C}^{2n})$ and $\bigwedge^{2n-p}(\mathbb{C}^{2n})$, and so the linear systems (5.1) can be restricted to linear systems on $\bigwedge^p(\mathbb{C}^{2n})$ and $\bigwedge^{2n-p}(\mathbb{C}^{2n})$ with an induced multi-symplectic structure. The formulation and analysis from this point of view is intriguing, but is outside the scope of the present work.

10. Instability of solitary waves for a Boussinesq model

The Boussinesq equations are a class of model partial differential equations for dispersive shallow-water waves, nonlinear strings and condensed matter physics. In this section the theory of §§2-8 will be applied to the stability of a class of solitary waves of the following Boussinesq model:

$$u_{tt} = \partial_{xx}(u - u^2 - u_{xx}), \quad x \in \mathbb{R}, \quad t > 0, \tag{10.1}$$

where $u(x,t)$ is a scalar-valued function. This Boussinesq equation is sometimes called the "good" Boussinesq equation because it is well-posed and there is a good local existence theory for the Cauchy problem and in some cases a global theory (cf. Bona & Sachs [10] and references therein).

For each c with $|c| < 1$ there is an explicit one-parameter family of solitary waves of equation (10.1) decaying to 0 as $x \to \pm\infty$ of the form

$$u(x,t) = \frac{3}{2}(1-c^2)\operatorname{sech}^2\left(\frac{1}{2}\sqrt{1-c^2}(x-ct)\right). \tag{10.2}$$

In Bona & Sachs [10] it is proved that this family of solitary waves is stable when $\frac{1}{2} < |c| < 1$. For the case $|c| \le \frac{1}{2}$, Alexander & Sachs [4] analyse the Evans function and show that there exists an unstable eigenvalue for all $|c| < \frac{1}{2}$. The analysis in [4] involves a computer-assisted calculation of the leading terms in the Taylor expansion of the Evans function. Alexander & Sachs conjectured that the convexity of the momentum and the Evans function were related but a precise link was not established.

We extend these instability results in several directions. First, the Boussinesq equation (10.1) has additional symmetry (associated with conservation of $\int u\,\mathrm{d}x$). By including the additional symmetry, we extend the solitary-wave state to a two-parameter family of waves. These waves are biasymptotic to a nonzero constant state at infinity. Secondly, we show that the *value* of this constant state at infinity affects the linear stability, and thirdly, we establish the link between the momentum (and other invariants) and the Evans function as conjectured in [4]. In fact all these results are deduced from the abstract framework of §§2-8.

The Boussinesq equation (10.1) can be formulated as a classical Hamiltonian system on an infinite-dimensional phase space (cf. Alexander & Sachs [4], equations (1.2)-(1.3)). We take this a step further and give a new formulation of equation (10.1) as a Hamiltonian system with a multi-symplectic structure. First the equivalence between the scalar PDE (10.1) and a multi-symplectic systems formulation is established.

By introducing new functions $v(x,t)$, $w(x,t)$ and $z(x,t)$ defined by

$$v = u_x, \quad w_{xx} = u_t, \quad \text{and} \quad z = w_x,$$

the scalar PDE (10.1) can be written as the first-order system of PDEs

$$\begin{aligned} 0 &= w_t + v_x - u + u^2 \\ 0 &= -u_x + v \\ 0 &= -u_t + z_x \\ 0 &= -w_x + z. \end{aligned} \tag{10.3}$$

Suppose $u(x,t)$ is a smooth bounded function, which is asymptotic to some constant value as $|x| \to \infty$, and its first, second, and third derivatives decay to 0 as $|x| \to \infty$. Then the scalar PDE (10.1) and the system (10.3) are equivalent. It is straightforward to see that if (u,v,w,z) is a solution of (10.3), then u is a solution of (10.1). The converse – if u is a solution of (10.1), then

$$(u, u_x, \int_{-\infty}^{x}\int_{-\infty}^{y} u_t(s,t)\mathrm{d}s\mathrm{d}y + (u(-\infty) - u^2(-\infty))t, \int_{-\infty}^{x} u_t(s,t)\mathrm{d}s)$$

is a solution of (10.3) – takes a bit more effort. We see immediately from the definitions that $v = u_x$, $z = w_x$ and $z_x = u_t$. To satisfy the third equation of (10.3), we require the following expression to vanish:

$$w_t + v_x - u + u^2 = \int_{-\infty}^{x}\int_{-\infty}^{\xi} u_{tt}(\zeta,t)\mathrm{d}\zeta\mathrm{d}\xi + u_{xx} - u + u^2 + (u(-\infty) - u^2(-\infty))\,. \tag{10.4}$$

By integrating (10.1) once, we get $\int_{-\infty}^{x} u_{tt}(\xi,t)\mathrm{d}\xi = \partial_x(-u_{xx} + u - u^2)$. (Here use that u_x and u_{xxx} converge to 0, if $|x| \to \infty$.) Another integration gives

$$\int_{-\infty}^{x}\int_{-\infty}^{\xi} u_{tt}(\zeta,t)\mathrm{d}\zeta\mathrm{d}\xi = (-u_{xx} + u - u^2) - (u(-\infty) - u^2(-\infty))\,.$$

Substitution into (10.4) shows that the right-hand side vanishes and hence the third equality of (10.3) is satisfied. Note that the solutions of (10.3) are not unique. One can always add an arbitrary constant to w. We will use this symmetry property later on.

The multi-symplecticity of the system (10.3) is established as follows. Let $\langle\cdot,\cdot\rangle$ be a standard inner product on $\mathbb{R}^4$ and define

$$Z = \begin{pmatrix} u \\ v \\ w \\ z \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix},$$

and $S(Z)=\frac{1}{2}u^2-\frac{1}{2}v^2-\frac{1}{2}z^2-\frac{1}{3}u^3$. Then the system (10.3) can be written as

$$\mathbf{M}Z_t+\mathbf{K}Z_x=\nabla S(Z)\,, \quad Z\in\mathbb{R}^4\,. \tag{10.5}$$

The gradient of S is with respect to the real inner product $\langle\cdot,\cdot\rangle$ in $\mathbb{R}^4$ and the matrices $\mathbf{M}$ and $\mathbf{K}$ are skew symmetric and constant and therefore can be associated with two-forms via

$$\omega(U,V)=\langle\mathbf{M}U,V\rangle\,, \quad \text{and} \quad \kappa(U,V)=\langle\mathbf{K}U,V\rangle\,, \quad \text{for any } U,V\in\mathbb{R}^4\,.$$

The system (10.5) is translation invariant in x and t: $Z(x+\varepsilon_1,t)$ and $Z(x,t+\varepsilon_2)$, for any $\varepsilon_1\in\mathbb{R}$ and $\varepsilon_2\in\mathbb{R}$, are solutions of (10.5) whenever $Z(x,t)$ is a solution. There is also an affine symmetry due to the fact that $Z(x,t)+\varepsilon_3(0,0,1,0)^T$, for any $\varepsilon_3\in\mathbb{R}$, is a solution when $Z(x,t)$ is a solution (constant shift in the w coordinate in (10.3)). Therefore we take $q=1$ and define $\mathcal{G}$ to be this one-dimensional affine group.

Remark 5. There are other symmetries of the system (10.5); for example, the action

$$(u,v,w,z)\mapsto(u,v,w+\varepsilon_4 x,z+\varepsilon_4)$$

corresponds to a one-parameter symmetry of equation (10.5). However, this group acts on the dependent and independent variables, instead of on the phase-space $\mathbb{R}^4$ only, and therefore this subgroup is not included in $\mathcal{G}$ since it does not commute with translations in the x-variable.

The Lie algebra $\mathfrak{g}$ of $\mathcal{G}$ is spanned by the generator $\xi_1(Z)=\mathbf{e}_3$, with $\mathbf{e}_3=(0,0,1,0)^T$. Thus $\mathbf{M}\xi_1(Z)=\nabla P_1(Z)$, where $P_1(Z)=(\mathbf{e}_1,Z)$ and $\mathbf{e}_1=(1,0,0,0)^T$. The generator related to the x-translation symmetry is $\xi_0(Z)=Z_x$. Constants of motion associated with the generators ξ_0 and ξ_1 are

$$\mathcal{I}(Z)=\tfrac{1}{2}\textstyle\int\omega(Z_x,Z)\mathrm{d}x=\tfrac{1}{2}\int(uw_x-u_xw)\,\mathrm{d}x$$

and the integral of the density $P_1(Z)$ where

$$P_1(Z)=\omega(\xi_1(Z),Z)=\langle\mathbf{e}_1,Z\rangle=u\,.$$

The constants of motion $\mathcal{I}$ and $\int P_1$ are well defined only if $uw_x-u_xw\in L^1(\mathbb{R})$ and $u\in L^1(\mathbb{R})$. These invariants are associated with the temporal symplectic structure. The functional associated with $\xi_1(Z)$ and the *spatial* symplectic structure is defined as follows:

$$\mathbf{K}\xi_1(Z)=-\mathbf{e}_4=(0,0,0,-1)^T=\nabla Q_1(Z)\,,$$

and therefore

$$Q_1(Z)=\kappa(\xi_1(Z),Z)=-\langle\mathbf{e}_4,Z\rangle=-z\,.$$

In the sequel we will show that the Boussinesq equation (10.1) has a two-parameter family of solitary wave solutions $u(x,t)=\mathcal{T}_{ct}\widetilde{u}(x;a,c)=\widetilde{u}(x-ct;a,c)$, where

$$\widetilde{u}(x;a,c)=\frac{1}{2}(1-c^2)-2\delta^2+6\delta^2\,\mathrm{sech}^2(\delta x), \quad \delta=\tfrac{1}{2}\sqrt{\sqrt{1+4a}-c^2}, \quad \text{and} \quad a>\tfrac{1}{4}(c^4-1).$$

In the multi-symplectic system, this two-parameter family can be embedded in a three-parameter family of solutions $\widetilde{Z}(x;a,b,c)$, see Lemma 14 for more details.

In the following sub-sections we will establish the behaviour of the Evans function, associated with the linearization of the system (10.5) about this family of waves, for small λ and large λ. These results combined with the theory of §§2-8 will lead to a proof of the following theorem.

Theorem 4. *Suppose $a>-\frac{1}{4}$ and $c^2<\frac{1}{4}\sqrt{1+4a}$. Then the multi-parameter family of solitary waves $\widetilde{u}(x;a,b,c)$ is linearly unstable.*

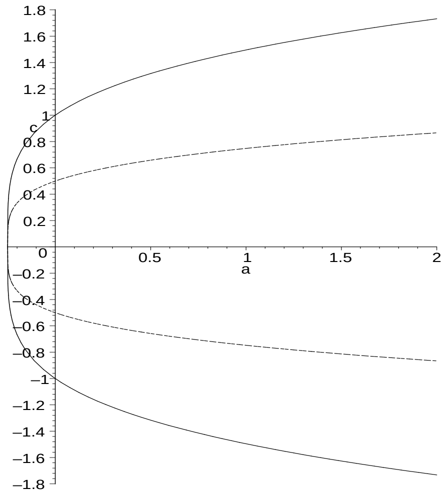

Fig. 10.1. The family of solitary waves is unstable for parameter values in the region bounded by the dashed curves. Solitary waves exist for all values of (a, c) within the larger region bounded by the solid curve.

In Figure 10.1, we have sketched the area of linear instability in the a-c plane, embedded in the region for which a member of the family of solitary waves exists. Solitary waves with vanishing tails correspond to the vertical line $a = 0$, and it is this line which corresponds to the results of [4].

To prove Theorem 4 we have to verify

1. the existence of the family of solitary waves (the hypotheses in §2.4),
2. the hypotheses on the systems at infinity (the hypotheses in §3.4),
3. the behaviour of the symplectic Evans matrix for λ near 0 ($\left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \frac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right]$ and χ_{00}^-),
4. the asymptotic behaviour of the symplectic Evans matrix as λ tends to infinity along the real axis (the hypothesis in Theorem 3).

These properties will be verified in the following subsections and we will end the section with a proof of Theorem 4.

10.1. A two-parameter family of solitary waves

First we prove the existence of a two-parameter family of solitary waves, i.e., we verify the hypotheses in §2.4.

Lemma 14. *The system* (10.5) *satisfies*

1. *a multi-symplectic formulation on* $\mathbb{R}^4$ *with* $\mathbf{J}_c$ *non-degenerate for all* $c \in \mathbb{R}$ *and* $S : \mathbb{R}^4 \to \mathbb{R}$ *is a smooth function. The system* (10.5) *is* $\mathcal{G}$*-equivariant with* $\mathcal{G}$ *a* 1*-dimensional affine subgroup of the Euclidean group on* $\mathbb{R}^4$ *;*
2. *for all* $a > -\frac{1}{4}$, $b \in \mathbb{R}$, *and* $c^2 < \sqrt{1+4a}$ *there exists a solitary-wave shape* $\widetilde{Z}(x; a, b, c)$ *satisfying* $\mathbf{K}\widetilde{Z}_x - c\mathbf{M}\widetilde{Z}_x = \nabla S(\widetilde{Z}) - a\nabla P_1(\widetilde{Z}) - b\nabla Q_1(\widetilde{Z})$*;*
3. *the solitary wave shape* $\widetilde{Z}$ *is bounded. The asymptotic shapes are given by* $Z_0^-(a,b) = (\frac{1}{2}(1 - \sqrt{1+4a}), 0, 0, -b)$ *and* $Z_0^+(a,b,c) = (\frac{1}{2}(1 - \sqrt{1+4a}), 0, -6c\sqrt{\sqrt{1+4a} - c^2}, -b)$*;*
4. *the invariant manifold at infinity* $\mathcal{M}_\infty(a,b)$ *is non-degenerate;*

5. *the derivative of the relative equilibrium shape* $\widetilde{Z}_x$ *is exponentially decaying: for* $\beta = \sqrt{\sqrt{1+4a} - c^2}$ *there exist vectors* $\Psi^{\pm}(a,b,c) \neq 0$ *such that*

$$\lim_{x\to\pm\infty} \mathrm{e}^{\pm\beta x}\widetilde{Z}_x = \Psi^{\pm} = 6\beta^2\,(\mp\beta, \beta^2, -c, \pm c\beta)^T$$

and $\lim_{x\to\pm\infty} \partial_x \left(\mathrm{e}^{\pm\beta x}\widetilde{Z}_x\right) = 0\,.$

Proof. The matrix $\mathbf{J}_c = \mathbf{K} - c\mathbf{M} = \begin{pmatrix} 0 & 1 & -c & 0 \\ -1 & 0 & 0 & 0 \\ c & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$ is invertible since $\det(\mathbf{J}_c) = 1$, for any $c \in \mathbb{R}$.

A shape of the solitary wave must satisfy

$$(\mathbf{K} - c\mathbf{M})\,\widetilde{Z}_x = \nabla S(\widetilde{Z}) - a\nabla P_1(\widetilde{Z}) - b\nabla Q_1(\widetilde{Z})\,. \tag{10.6}$$

Explicitly this means

$$\begin{cases} 0 = -cw_x + v_x - u + u^2 - a \\ 0 = -u_x + v \\ 0 = cu_x + z_x \\ 0 = -w_x + z + b \end{cases} \Leftrightarrow \begin{cases} 0 = c^2 u - c(\mathsf{const} + b) + u_{xx} - u + u^2 - a \\ v = u_x \\ z = -cu + \mathsf{const} \\ w_x = z + b = -cu + b + \mathsf{const} \end{cases} \tag{10.7}$$

with const some integration constant. Using the condition that the shape of a solitary wave must be bounded and that Z_0^- is independent of c, a straightforward calculation gives that $\mathsf{const} = -b + \frac{c}{2}(1 - \sqrt{1+4a})$ and

$$\widetilde{Z}(x; a, b, c) = \begin{pmatrix} \frac{1}{2}(1 - c^2 - 4\delta^2) + 6\delta^2\operatorname{sech}^2(\delta x) \\ -12\delta^3\operatorname{sech}^2(\delta x)\tanh(\delta x) \\ -6c\delta(\tanh(\delta x) + 1) \\ -6c\delta^2\operatorname{sech}^2(\delta x) - b \end{pmatrix} \quad \text{with} \quad \delta = \tfrac{1}{2}\sqrt{\sqrt{1+4a} - c^2}\,,$$

is a solution of (10.7) for any $a > -\frac{1}{4}$, $b \in \mathbb{R}$ and $c^2 < \sqrt{1+4a}$.

Since $\lim_{|x|\to\infty} \operatorname{sech}^2(x) = 0$, it follows that

$$Z_0^{\pm} = \lim_{x\to\pm\infty} \widetilde{Z}(x;\cdot) = \begin{pmatrix} \frac{1}{2}(1 - c^2 - 4\delta^2) \\ 0 \\ -12c\delta \text{ or } 0 \\ -b \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(1 - \sqrt{1+4a}) \\ 0 \\ -12c\delta \text{ or } 0 \\ -b \end{pmatrix}.$$

Furthermore, $P_1(Z_0^{\pm}) = \frac{1}{2}(1 - \sqrt{1+4a})$ and $Q_1(Z_0^{\pm}) = b$. Thus

$$\begin{vmatrix} \frac{\partial P_1}{\partial a} & \frac{\partial P_1}{\partial b} \\ \frac{\partial Q_1}{\partial a} & \frac{\partial Q_1}{\partial b} \end{vmatrix} = \begin{vmatrix} -\frac{1}{\sqrt{1+4a}} & 0 \\ 0 & 1 \end{vmatrix} = -\frac{1}{\sqrt{1+4a}} \neq 0$$

and we can conclude that the invariant manifold at infinity $\mathcal{M}_\infty(a,b)$ is nondegenerate.

Finally,

$$\widetilde{Z}_x = \begin{pmatrix} -12\delta^3\operatorname{sech}^2(\delta x)\tanh(\delta x) \\ 24\delta^4\operatorname{sech}^2(\delta x) - 36\delta^4\operatorname{sech}^4(\delta x) \\ -6c\delta^2\operatorname{sech}^2(\delta x) \\ 12c\delta^3\operatorname{sech}^2(\delta x)\tanh(\delta x) \end{pmatrix}, \quad \text{hence} \quad \mathrm{e}^{\pm 2\delta x}\widetilde{Z}_x \to 24\delta^2 \begin{pmatrix} \mp 2\delta \\ 4\delta^2 \\ -c \\ \pm 2c\delta \end{pmatrix},$$

for $x \to \pm\infty$. Thus $\beta = 2\delta = \sqrt{\sqrt{1+4a} - c^2}$. And also

$$\lim_{x\to\pm\infty} \mathrm{e}^{\pm\beta x} \widetilde{Z}_{xx} = \mp 48\delta^4 \begin{pmatrix} \mp 2\delta \\ 4\delta^2 \\ -c \\ \pm 2c\delta \end{pmatrix} = \mp\beta \lim_{x\to\pm\infty} \mathrm{e}^{\pm 2\beta x} \widetilde{Z}_x ,$$

hence $\lim_{x\to\pm\infty} \partial_x \left(\mathrm{e}^{\pm\beta x} \widetilde{Z}_x \right) = 0$.

Remark 6. Note that the u- and v-components of $\widetilde{Z}(x;a,b,c)$ do not depend on b. Hence the solitary wave solution of (10.1) does not depend on the choice of b. When analysing the stability of this wave, we can choose any convenient value of b.

Remark 7. The requirement that $\beta^2 > 0$ ensures that the hypothesis $A \neq 0$ in (4.11) is satisfied since $\beta^2 = \sqrt{1+4a}\, A$ using the result above with A as defined in §4.

10.2. The systems at infinity

Next we analyse the linearised stability equation for $|x| \to \infty$, i.e., we verify the hypotheses in §3.5. Since the symmetry group $\mathcal{G}$ is affine, it follows that $\mathbf{B}^+ = \mathbf{B}^-$. Therefore define $\mathbf{B}^\infty(a,b) = \mathbf{B}^+(a,b,c) = \mathbf{B}^-(a,b)$ and $\mathbf{A}^\infty(\lambda;a,b,c) = \mathbf{A}^+(\lambda;a,b,c) = \mathbf{A}^-(\lambda;a,b,c)$.

First we determine the position of the eigenvalues of $\mathbf{A}^\infty(\lambda;a,b,c)$ as a function of λ and the parameters (a,b,c).

Lemma 15. *Suppose $\lambda \in \mathbb{C}$ and $\Re(\lambda)$ positive. The eigenvalue problem of $\mathbf{A}^\infty(\lambda;a,b,c)$ for $\mu \in \mathbb{C}$ has 2 roots $\mu_1(\lambda;a,b,c)$ and $\mu_2(\lambda;a,b,c)$ with*

$$\Re(\mu_1(\lambda;a,b,c)) \le \Re(\mu_2(\lambda;a,b,c)) < 0 .$$

All other eigenvalues have nonnegative real part. The eigenvalues can only cross the imaginary axis if λ is purely imaginary.

The eigenvalues $\mu_1(\lambda;a,b,c)$, $\mu_2(\lambda;a,b,c)$ are simple for $Re(\lambda) > 0$, except at the isolated point

$$\lambda_0 = \frac{1}{16} \sqrt{(c + \sqrt{9c^2 + 32\delta^2})^2 - 4c^2} \left(3c + \sqrt{9c^2 + 32\delta^2} \right) \in \mathbb{R}.$$

The eigenvalues $\mu_1(\lambda;a,b,c)$ and $\mu_2(\lambda;a,b,c)$ can be extended in an analytic way to $\lambda = 0$ with

$$\lim_{\lambda\to 0} \mu_1(\lambda;a,b,c) = \beta(a,b,c) \quad \text{and} \quad \lim_{\lambda\to 0} \mu_2(\lambda;a,b,c) = 0 .$$

An eigenvector related to an eigenvalue μ can be written as a scalar multiple of

$$\zeta = (\mu^2, \mu^3, \lambda - c\mu, \mu(\lambda - c\mu))^T .$$

Hence the eigenvectors associated with the eigenvalue 0 at $\lambda = 0$ both are a multiple of $(0,0,1,0)^T = \xi_1$. To be explicit, define

$$\boldsymbol{\chi}_{11}^+ = 1, \quad \text{and} \quad \boldsymbol{\chi}_{11}^- = \frac{\sqrt{c^2 + 4\delta^2}}{2} ,$$

then

$$\zeta_2^+(0;a,b,c) = \boldsymbol{\chi}_{11}^+ \xi_1(Z_0^+), \quad \textit{and} \quad \widetilde{\eta}_2^-(0;a,b,c) = \boldsymbol{\chi}_{11}^- \xi_1(Z_0^-)$$

with $\widetilde{\eta}_2^-(0;a,b,c) = \lim_{\lambda\to 0} \lambda\, \eta_2^-(\lambda;a,b,c)$, hence $m_2 = 1$.

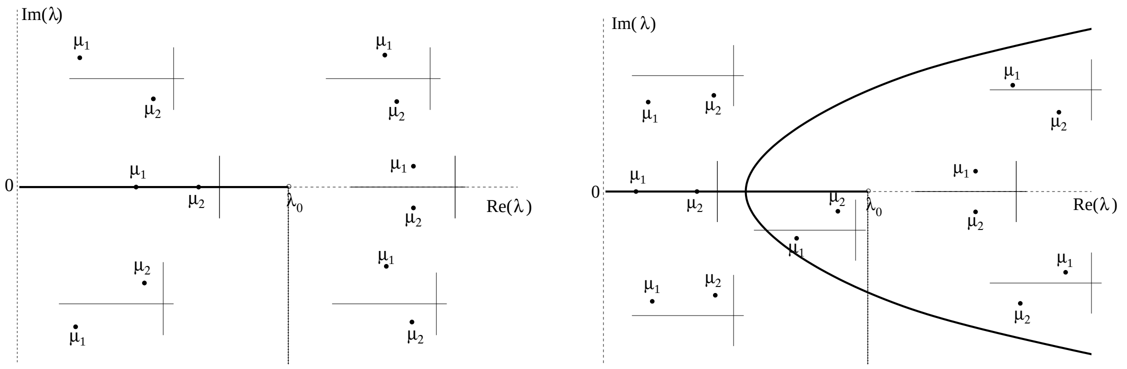

Fig. 10.2. An overview of the position of the eigenvalues μ_1 and μ_2 as a function of λ. The left picture gives the situation for $c > 0$ and the right picture the situation for $c < 0$. The dotted vertical line going down from λ_0 denotes the cut in the complex plane, necessary because of the double eigenvalue at λ_0. Eigenvalues are real, only when λ is on the solid line(s).

Remark 8. Figure 10.2 gives an overview of the position of the eigenvalues μ_1 and μ_2 in the complex right-half λ-plane.

Proof. It is easy to calculate that $\det(\mathbf{A}^\infty(\lambda; a, b, c) - \mu \mathbf{I}) = 0$ implies that

$$\mu^4 - 4\delta^2\mu^2 - 2c\lambda\mu + \lambda^2 = 0. \tag{10.8}$$

We can immediately see that if $\lambda \neq 0$ then $\mu = 0$ is not a root of this equation.

If we denote the roots of (10.8) by μ_1, μ_2, μ_3, and μ_4, then we have for $\lambda \in \mathbb{R}$

$$\mu_1\mu_2\mu_3\mu_4 = \lambda^2 > 0 \quad \text{and} \quad \mu_1 + \mu_2 + \mu_3 + \mu_4 = 0.$$

The first equation implies that for $\lambda \in \mathbb{R}$ we have an even number of roots with positive real part, i.e., 0, 2, or 4 of such roots. From the second equation it follows that we must have both roots with positive and with negative real part. So we can conclude that – for λ real and positive – there are two roots with negative real part and two roots with positive real part.

Furthermore, a short calculation shows that eigenvalues can cross the imaginary axis only if λ is on the imaginary axis. Hence the 2-2 split of the eigenvalues is valid in the full right half of the complex plane.

Furthermore, double negative eigenvalues can occur at only one isolated real value of $\lambda = \lambda_0$, being

$$\lambda_0 = \frac{1}{16}\sqrt{(c + \sqrt{9c^2 + 32\delta^2})^2 - 4c^2}\,\left(3c + \sqrt{9c^2 + 32\delta^2}\right).$$

Since this point is isolated in $\mathbb{C}$, it can not give a sign change in the Evans function on the real axis. Furthermore, for $c > 0$, eigenvalues with negative real part can be real only if $\lambda \in \mathbb{R}$ and for $c < 0$, eigenvalues with negative real part can be real if $\lambda \in \mathbb{R}$ or on the curve $(\lambda - cs)^2 = s^2(s^2 - 4\delta^2 - c^2)$, $|s| > \sqrt{c^2 + 4\delta^2}$, see Figure 10.2. Hence for $\lambda \in \mathbb{R}$ and $\lambda > \lambda_0$, we have $\Im(\mu_1) > 0$.

For small values of λ the eigenvalues have the following Taylor expansions,

$$\begin{aligned}
\mu_1 &= -2\delta + \tfrac{c}{4\delta^2}\lambda + \tfrac{3c^2+4\delta^2}{64\delta^5}\lambda^2 + \mathcal{O}(\lambda^3),\\
\mu_2 &= -\tfrac{c+\sqrt{c^2+4\delta^2}}{4\delta^2}\lambda - \tfrac{c^4+4c^2\delta^2+2\delta^4+c\sqrt{c^2+4\delta^2}(c^2+2\delta^2)}{64\delta^8\sqrt{c^2+4\delta^2}}\lambda^3 + \mathcal{O}(\lambda^5),\\
\mu_3 &= -\tfrac{c-\sqrt{c^2+4\delta^2}}{4\delta^2}\lambda + \tfrac{c^4+4c^2\delta^2+2\delta^4-c\sqrt{c^2+4\delta^2}(c^2+2\delta^2)}{64\delta^8\sqrt{c^2+4\delta^2}}\lambda^3 + \mathcal{O}(\lambda^5),\\
\mu_4 &= 2\delta + \tfrac{c}{4\delta^2}\lambda - \tfrac{3c^2+4\delta^2}{64\delta^5}\lambda^2 + \mathcal{O}(\lambda^3).
\end{aligned} \tag{10.9}$$

A straightforward calculation shows that an eigenvector related to an eigenvalue μ can be written as a scalar multiple of $\zeta = (\mu^2, \mu^3, \lambda - c\mu, \mu(\lambda - c\mu))^T$. Since $\mathbf{A}^+ = \mathbf{A}^-$, we can drop the $\pm$ subscript for the eigenvectors and with the expression above for the eigenvector associated with an eigenvalue μ, we get for λ near 0:

$$\begin{aligned}
\zeta_1(\lambda; a,b,c) &= C_1\,(-2\delta, 4\delta^2, -c, 2c\delta)^T + \mathcal{O}(\lambda),\\
\eta_1(\lambda; a,b,c) &= \chi_{00}^-\,C_1\,(2\delta, 4\delta^2, -c, -2c\delta)^T + \mathcal{O}(\lambda),\\
\zeta_2(\lambda; a,b,c) &= (0,0,1,0) + \lambda\,\tfrac{\sqrt{c^2+4\delta^2}+c}{4\delta^2\sqrt{c^2+4\delta^2}}\,(1,0,0,-\sqrt{c^2+4\delta^2}) + \mathcal{O}(\lambda^2),\\
\eta_2(\lambda; a,b,c) &= C_2(\lambda)\left[(0,0,1,0) + \lambda\,\tfrac{\sqrt{c^2+4\delta^2}+c}{4\delta^2\sqrt{c^2+4\delta^2}}(-1,0,0,\sqrt{c^2+4\delta^2}) + \mathcal{O}(\lambda^2)\right],
\end{aligned} \tag{10.10}$$

with C_1 such that $\zeta_1(\lambda; a,b,c) = \Psi^+$. The orthonormalisation condition for η_2 gives

$$1 = \Omega(\eta_2, \zeta_2) = 2\lambda C_2(\lambda)\,\frac{1}{\sqrt{c^2+4\delta^2}},$$

thus $C_2(\lambda) = \frac{\sqrt{c^2+4\delta^2}}{2\lambda}$ and $m_2 = 1$.

With this lemma, we have verified the hypotheses of § 3.5:

1. Define the set Λ to be

$$\Lambda = \{\lambda \in \mathbb{C} \mid \Re(\lambda) > 0,\ |\Im(\lambda)| < \Re(\lambda)\} \backslash \{\lambda_0 + iy \mid y \le 0\}.$$

   see also Figure 10.3. This set is open simply-connected and $0 \in \overline{\Lambda}$. The Evans function can not change sign on the real axis at the singular point λ_0.
2. For all $\lambda \in \Lambda$, the spectrum of $\mathbf{A}^\pm(\lambda, a, b, c)$ consists of 2 simple eigenvalues with strictly negative real part and 2 eigenvalues with non-negative real part.
3. The spectrum of the matrix $\mathbf{A}^\pm(0, a, b, c)$ is of the form $\{-\beta,\, 0,\, \beta\}$ and the eigenvalues $\pm\beta$ are simple.
4. The vector spaces $\mathrm{span}\{\zeta_2^+(0,a,b,c)\}$, $\mathrm{span}\{\widetilde{\eta}_2^-(0,a,b,c)\}$ and $\mathrm{span}\{\xi_1(Z_0^\pm(a,b,c))\}$ are isomorphic.

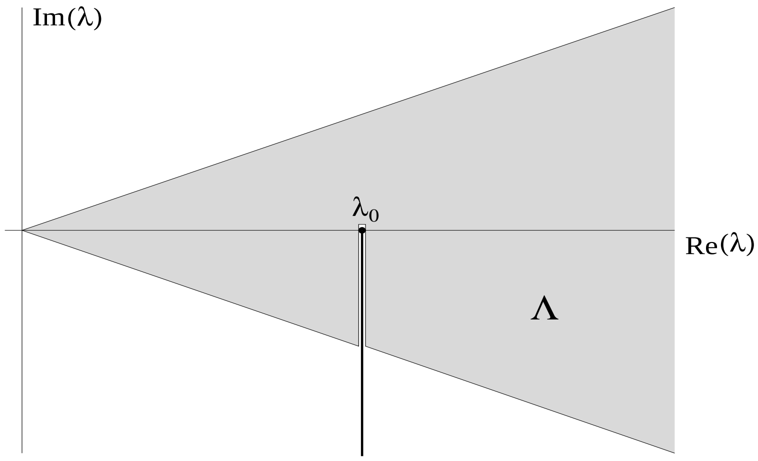

Fig. 10.3. The set Λ in the right half complex plane in which the eigenvalues are analytic.

10.3. The symplectic Evans matrix for λ near 0

In this section we determine the value of χ_{00}^- and $\left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \frac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right]$. With Theorem 2 and the fact that $\partial\mathcal{D} = 1$, this gives the behaviour of the Evans function for λ small.

Lemma 16. *For all $b \in \mathbb{R}$, $a > -\frac{1}{4}$ and $c^2 < \sqrt{1+4a}$, we have*

$$\chi_{00}^- = -\frac{1}{2304\delta^5(c^2+4\delta^2)} = -\frac{1}{2304\delta^5\sqrt{1+4a}}$$

and

$$\left[\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \tfrac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right] = 6\sqrt{\sqrt{1+4a}-c^2}\,\left(\sqrt{1+4a}-4c^2\right).$$

Proof. Using (10.10), we get

$$\zeta_1(0;a,b,c) = C_1\,(-2\delta, 4\delta^2, -c, 2c\delta)^T \quad \text{and} \quad \eta_1(0;a,b,c) = \chi_{00}^-\,C_1\,(2\delta, 4\delta^2, -c, -2c\delta)^T.$$

From $\zeta_1 = \Psi^+ = 6\beta^2\,(-\beta, \beta^2, -c, c\beta)^T$, it follows that $C_1 = 24\delta^2$. Furthermore, by the orthonormalisation condition

$$1 = \Omega(\eta_1, \zeta_1) = -4\delta(4\delta^2+c^2)\,C_1^2\,\chi_{00}^- = -2304\delta^5(4\delta^2+c^2)\chi_{00}^-,$$

we get $\chi_{00}^- = -\frac{1}{2304\delta^5(c^2+4\delta^2)}$.

Next we calculate $\left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \frac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right]$. By definition,

$$\mathcal{I}(Z) = \frac{1}{2}\int_{-\infty}^{\infty}\omega(Z_x, Z)\,\mathrm{d}x = \int_{-\infty}^{\infty}uw_x\,\mathrm{d}x - \frac{1}{2}uw\Big|_{-\infty}^{\infty}.$$

Using the expressions for the solitary wave as given in Lemma 14, we get

$$\begin{aligned}
\mathcal{I}(\widetilde{Z}) &= 3c\delta(1-c^2+20\delta^2)\int_{-\infty}^{\infty}\operatorname{sech}^2(s)\,\mathrm{d}s - 36c\delta^3\int_{-\infty}^{\infty}\operatorname{sech}^4(s)\,\mathrm{d}s \\
&\quad + 72c\delta^3\int_{-\infty}^{\infty}\tanh(s)\operatorname{sech}^2(s)\,\mathrm{d}s \\
&= 3\delta c(1-c^2+12\delta^2) \\
&= \tfrac{3}{2}c\sqrt{\sqrt{1+4a}-c^2}(3\sqrt{1+4a}+1-4c^2).
\end{aligned}$$

By differentiating this expression with respect to c, we get

$$\tfrac{\partial}{\partial c}\mathcal{I}(\widetilde{Z}(x;a,b,c)) = \tfrac{3}{2\sqrt{\sqrt{1+4a}-c^2}}\,(16c^4 - 2c^2(1+9\sqrt{1+4a}) + \sqrt{1+4a}(1+3\sqrt{1+4a})).$$

Furthermore,

$$\partial_c Z_0^+ = \left(0, 0, \frac{6}{\sqrt{\sqrt{1+4a}-c^2}}\,\left(2c^2-\sqrt{1+4a}\right), 0\right).$$

Hence

$$\omega(Z_0^+, \partial_c Z_0^+) = \frac{3\left(1-\sqrt{1+4a}\right)\left(\sqrt{1+4a}-2c^2\right)}{\sqrt{\sqrt{1+4a}-c^2}}$$

and after some reordering of the terms, we can conclude that

$$\left[\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \tfrac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right] = 6\sqrt{\sqrt{1+4a}-c^2}\,\left(\sqrt{1+4a}-4c^2\right).$$

Remark 9. By using the expressions for ζ_2^+ and η_2^- in (10.10), we can verify that $\partial\mathcal{D} = 1$. Indeed, from (10.10) it follows that

$$\frac{\partial}{\partial\lambda}\bigg|_{\lambda=0} \zeta_2^+(\lambda; a, b, c) = \frac{\sqrt{c^2+4\delta^2}+c}{4\delta^2\sqrt{c^2+4\delta^2}}\,(1, 0, 0, -\sqrt{c^2+4\delta^2})^T$$

and

$$\frac{\partial}{\partial\lambda}\bigg|_{\lambda=0} \tilde{\eta}_2^-(\lambda; a, b, c) = \frac{\sqrt{c^2+4\delta^2}+c}{8\delta^2}\,(-1, 0, 0, \sqrt{c^2+4\delta^2})^T$$

By substituting this in the definition (6.2) for $\partial\mathcal{D}$, we see that $\partial\mathcal{D} = 1$.

10.4. The symplectic Evans function for large $|\lambda|$

First we determine the asymptotic behaviour of the eigenvalues. For large values of λ, we have that $\det(\mathbf{B}^\infty(a,b) - \lambda\mathbf{M} - \mu\mathbf{J}_c) = 0$ if

$$\mu = \mu_0\sqrt{\lambda} - \tfrac{c}{2}\mu_0^2 + \tfrac{c^2-2}{8}\mu_0^3\,\tfrac{1}{\sqrt{\lambda}} + \mathcal{O}(\tfrac{1}{\lambda}), \quad \text{with} \quad \mu_0^4 = -1.$$

Following §8, we define $\alpha = \frac{1}{2}$ and a small parameter $\varepsilon = \lambda^{-\alpha} = \frac{1}{\sqrt{\lambda}}$. Furthermore, we introduce a new "space" variable $s = \frac{x}{\varepsilon}$ (i.e., $\gamma = 1$), a selfadjoint transformation $T(\varepsilon) = \operatorname{diag}(1, \varepsilon, 1, \varepsilon)$ and transformed vectors $U^{\text{new}}(s) = T(\varepsilon)U(\varepsilon s)$ and $W^{\text{new}}(s) = T(\varepsilon)W(\varepsilon s)$. This implies that the linearized and the adjoint systems transform to equations (8.1) and (8.2) with

$$\mathbf{J}_c^{\text{new}}(\varepsilon) = \mathbf{K} - c\varepsilon\mathbf{M}, \quad \mathbf{M}^{\text{new}}(\varepsilon) = \mathbf{M}, \quad \text{and} \quad \mathbf{B}^{\text{new}}(s,\varepsilon) = \operatorname{diag}(\varepsilon^2(1-2\tilde{u}(s)), -1, 0, -1).$$

Hence $\mathbf{J}_c^{\text{new}}(0) = \mathbf{K}$, which is an invertible matrix. Also, $\mathbf{M}^{\text{new}}(0) = \mathbf{M}$ and

$$\mathbf{B}^{\text{new}}(s,\varepsilon) - \lim_{\sigma\to\pm\infty}\mathbf{B}^{\text{new}}(\sigma,\varepsilon) = \varepsilon^2(1 - \sqrt{1+4a} - 2\tilde{u}(s))\ \operatorname{diag}(1,0,0,0)\,.$$

This means that we satisfy the conditions of Theorem 3 with $\kappa = 1$ and we can conclude that

$$\det \boldsymbol{\Omega}(\lambda; a, b, c) = 1 + \mathcal{O}(\frac{1}{\sqrt{\lambda}}), \qquad \lambda \to \infty.$$

10.5. Proof of Theorem 4

To prove Theorem 4, we will invoke Theorem 1. In the previous subsections we have shown that

$$d_\infty = 1, \quad \chi_{00}^- = -\frac{1}{2304\delta^5(c^2+4\delta^2)}$$

and

$$\left[\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \tfrac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right] = 6\sqrt{\sqrt{1+4a} - c^2}\,\left(\sqrt{1+4a} - 4c^2\right).$$

This implies that $\left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \frac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right] > 0$ if $c < 2 < \frac{1}{4}\sqrt{1+4a}$. Since $\chi_{00}^- < 0$, we can conclude that the solitary wave solution is unstable if $c < 2 < \frac{1}{4}\sqrt{1+4a}$.

10.6. Appendix: the kernel of $[\mathbf{J}_c D_x - \mathbf{B}(x;a,b,c)]$ for the Boussinesq model

In §10.4 we have seen that $\mathbf{B}(x;a,b,c) = \mathrm{diag}(1-2\tilde{u}(x;a,b,c)\,,\,-1\,,\,0\,,\,-1)$. The equivariance of (10.1) under a shift in the w-coordinate and under translations in the x-direction, gives two bounded solutions in the kernel:

$$Z_1^0(x) = \tilde{Z}_x = \xi_0(\tilde{Z}) \quad \text{and} \quad Z_2^0(x) = \mathbf{e}_3 = \xi_1(\tilde{Z}).$$

It is possible to find the other solutions in the kernel as well.

Lemma 17. *The solutions in the kernel of $\mathbf{J}_c D_x - \mathbf{B}(x;a,b,c)$ are spanned by*

$$Z_1^0(x) = \tilde{Z}_x(x), \quad Z_2^0(x) = \mathbf{e}_3, \quad Z_3^0(x) = \begin{pmatrix} g(x)\tilde{u}_x(x) \\ g(x)\tilde{u}_{xx}(x) + g_x(x)\tilde{u}_x(x) \\ c\int_0^x g(s)\tilde{u}_x(s)\,\mathrm{d}s + x \\ cg(x)\tilde{u}_x(x) + 1 \end{pmatrix},$$

and

$$Z_4^0(x) = \begin{pmatrix} h(x)\tilde{u}_x(x) \\ h(x)\tilde{u}_{xx}(x) + h_x(x)\tilde{u}_x(x) \\ c\int_0^x h(s)\tilde{u}_x(s)\,\mathrm{d}s \\ c(x)h\tilde{u}_x(x) \end{pmatrix}$$

with $g(x) = -c\int_0^x \frac{\hat{u}(s)}{(\hat{u}_x(s))^2}\,\mathrm{d}s$ and $h(x) = \int_0^x \frac{1}{(\tilde{u}_x(s))^2}\,\mathrm{d}s$. Recall that $\hat{Z}(x) = \tilde{Z}(x) - Z_0^-$.

The solution $Z_1^0(x)$ is exponentially decaying for $|x|\to\infty$, the solution $Z_4^0(x)$ is exponentially growing for $|x|\to\infty$, and the solutions $Z_2^0(x)$ and $Z_3^0(x)$ are not exponentially growing nor exponentially decaying for $|x|\to\infty$.

Proof. We have

$$(\mathbf{J}_c D_x - \mathbf{B}(x;a,b,c))\,Z = 0 \Leftrightarrow \begin{cases} 0 = c^2 u + cb_0 + u_{xx} - u + 2\tilde{u}u \\ v = u_x \\ z = cu + b_0 \\ w_x = z = cu + b_0 \end{cases},$$

for any constant b_0. By substituting $u(x) = k(x)\,\tilde{u}_x(x)$ into the first equation above, we get

$$\frac{\mathrm{d}}{\mathrm{d}x}(k_x(\tilde{u}_x)^2 + cb_0\tilde{u}) = 0,$$

which implies that

$$k(x) = -cb_0\int_0^x \frac{\hat{u}(s)}{(\hat{u}_x(s))^2}\,\mathrm{d}s + b_1\int_0^x \frac{1}{(\hat{u}_x(s))^2}\,\mathrm{d}s + b_2 = b_0 g(x) + b_1 h(x) + b_2,$$

for any constants b_1 and b_2.

The solution $h(x)\tilde{u}_x(x)$ grows exponentially for $|x|\to\infty$. Furthermore, $g(x)\tilde{u}_x(x) = \frac{c}{4\delta^2}$ for $|x|\to\infty$. Hence

$$Z_3^0(x) = (1+\frac{c^2}{4\delta^2})x\mathbf{e}_3 + \frac{c}{8\delta^2}\begin{pmatrix} 2 \\ 0 \\ \mp 3c \\ 2c \end{pmatrix} + o(1), \quad x\to\pm\infty. \tag{10.11}$$

11. Oscillatory solitons and NLS models

The nonlinear Schrödinger equation (NLS) and coupled NLS models and their generalisations are an important class of model equations for wave packets in electromagnetics, water waves, nonlinear optics, vortex filament dynamics and many other areas (cf. [8,45–47,43,24,28,40,12,5, 33,30,34] and references therein). Many of these models can be written in the general form

$$\mathrm{i}a_1\frac{\partial A}{\partial t}+\mathrm{i}a_2\frac{\partial A}{\partial x}+a_3\frac{\partial^2 A}{\partial x^2}=\frac{\delta H}{\delta\overline{A}} \tag{11.1}$$

$$\mathrm{i}b_1\frac{\partial B}{\partial t}+\mathrm{i}b_2\frac{\partial B}{\partial x}+b_3\frac{\partial^2 B}{\partial x^2}=\frac{\delta H}{\delta\overline{B}} \tag{11.2}$$

where a_1,a_2,a_3,b_1,b_2,b_3 are real parameters, $A(x,t)$ and $B(x,t)$ are complex-valued functions of $(x,t)\in\mathbb{R}^2$ and $H(A,\overline{A},B,\overline{B})$ is a real-valued function. The classical cubic nonlinear Schrödinger equation is recovered by taking $B=0$ and $H=\frac{1}{2}a_4|A|^4$ with $a_4\in\mathbb{R}$. The $\chi^{(2)}$ coupled NLS model in optics is recovered by taking

$$H(A,\overline{A},B,\overline{B})=a_4|A|^2+b_4|B|^2+\mathrm{Re}(\alpha\overline{A}^2B)\,, \tag{11.3}$$

where a_4,b_4 are real parameters and $\alpha\in\mathbb{C}$. Similarly, the classical coupled NLS with cubic coupling is recovered by taking

$$H(A,\overline{A},B,\overline{B})=a_4|A|^2+b_4|B|^2+\tfrac{1}{2}a_5|A|^4+\tfrac{1}{2}b_5|B|^4+\beta|A|^2|B|^2\,, \tag{11.4}$$

where a_4,b_4,a_5,b_5 and β are real parameters.

Another way to view these systems is as spatial perturbations of classical two degree-of-freedom Hamiltonian systems. In other words, the *kinetic part* of (11.1) – the spatially-independent part – is essentially a Hamiltonian system on $\mathbb{R}^4$, when taking real coordinates for

$$\mathrm{i}a_1\frac{\partial A}{\partial t}=\frac{\delta H}{\delta\overline{A}}\quad\text{and}\quad\mathrm{i}b_1\frac{\partial B}{\partial t}=\frac{\delta H}{\delta\overline{B}}\,, \tag{11.5}$$

with the a_2,b_2,a_3,b_3 terms bringing in spatial inhomogeneity.

11.1. Multi-symplectifying coupled NLS models

It is well known that NLS models and coupled NLS models are infinite-dimensional Hamiltonian systems. Indeed, the recognition that the NLS equation – particularly the cubic NLS equation – is an infinite-dimensional Hamiltonian system played a part in the discovery of integrability of cubic NLS by Zakharov & Shabat [46]. Since then, this *temporal* symplectic structure has formed the basis for much of the stability analyses for NLS (e.g. [24,34,43] and references therein). On the other hand, writing NLS in the form

$$\psi_{xx}=-\mathrm{i}\psi_t-W'(|\psi|^2)\psi\,,$$

it is possible to view and formulate NLS as a Hamiltonian system with respect to the x-variable. For example, Jones [28] formulates the linear stability problem for solitary waves using this spatial dynamics formulation – but coordinates are chosen that eliminate the temporal Hamiltonian structure. Bridges [12] formulates NLS and coupled NLS models as spatial evolution equations with a symplectic structure, but this formulation is valid only for NLS or coupled NLS restricted to a moving frame.

In this section, it is shown that systems of the form (11.1) have a new formulation as a Hamiltonian system on a multi-symplectic structure, and therefore the framework of this paper applies to this wide class of these NLS-type models.

The system (11.1) is multi-symplectified as follows. Introduce real coordinates $(q,p) \in \mathbb{R}^4 \times \mathbb{R}^4$, defined by

$$A = q_1 + \mathrm{i}q_2\,, \quad B = q_3 + \mathrm{i}q_4\,, \tag{11.6}$$

$$p_1 + \mathrm{i}p_2 = a_3 A_x + \tfrac{1}{2}\mathrm{i}a_2 A\,, \quad p_3 + \mathrm{i}p_4 = b_3 B_x + \tfrac{1}{2}\mathrm{i}b_2 B\,, \tag{11.7}$$

and

$$W(q,p) = \tfrac{1}{2}H(q_1 + \mathrm{i}q_2, q_1 - \mathrm{i}q_2, q_3 + \mathrm{i}q_4, q_3 - \mathrm{i}q_4)\,. \tag{11.8}$$

Then the system (11.1) is equivalent to

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z)\,, \quad Z \in \mathbb{R}^8\,, \tag{11.9}$$

with

$$\mathbf{M} = \begin{pmatrix} -a_1\mathbf{J}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & -b_1\mathbf{J}_2 & \mathbf{0}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} \mathbf{0}_2 & \mathbf{0}_2 & -\mathbf{I}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & -\mathbf{I}_2 \\ \mathbf{I}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{I}_2 & \mathbf{0}_2 & \mathbf{0}_2 \end{pmatrix},$$

and

$$\begin{aligned} S(Z) = {}& \frac{1}{2a_3}(p_1^2 + p_2^2) + \frac{1}{2b_3}(p_3^2 + p_4^2) + \frac{a_2}{2a_3}(p_1 q_2 - p_2 q_1) + \frac{b_2}{2b_3}(p_3 q_4 - p_4 q_3) \\ & + \frac{a_2^2}{8a_3}(q_1^2 + q_2^2) + \frac{b_2^2}{8b_3}(q_3^2 + q_4^2) - W(q,p)\,. \end{aligned} \tag{11.10}$$

Systems of the form (11.1) are known to have a wide range of solitary wave solutions (cf. [47, 40,37,33,34,39] and references therein). A complete analysis of the stability problem for general systems of the type (11.1) would take us too far afield. To illustrate how the theory of this paper applies to such systems, we apply the theory of §§2-8 to a single NLS model. The approach is similar – albeit more complex – for the coupled systems.

11.2. Dark solitons and cubic NLS

First, we consider the basic form for the NLS equation:

$$\mathrm{i}\psi_t = \psi_{xx} + W'(|\psi|^2)\psi, \quad x \in \mathbb{R}, \quad t > 0\,, \quad \psi \in \mathbb{C}, \tag{11.11}$$

where $W : \mathbb{R} \to \mathbb{R}$ a some smooth function. The most well-known form – and the form for W we will concentrate on for definiteness – is the defocusing cubic NLS equation, for which $W'(z) = k - z$, with k any real constant.

As indicated for the coupled NLS equations, the multi-symplectic form for the single NLS equation is obtained by writing $\psi = q_1 + \mathrm{i}q_2$ and $\psi_x = p_1 + \mathrm{i}p_2$ and $Z = (q_1, q_2, p_1, p_2) \in \mathbb{R}^4$. Define

$$\mathbf{M} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

and $S(Z)=\frac{1}{2}(p_1^2+p_2^2)+\frac{1}{2}W(q_1^2+q_2^2)$. Then (11.11) is equivalent to $\mathbf{M}Z_t+\mathbf{K}Z_x=\nabla S(Z)$. This system is invariant under the action of the rotation group SO(2), given by

$$G_\theta(Z)=\mathbf{R}_\theta Z, \quad \text{with} \quad \mathbf{R}_\theta=\begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Define the generator $\xi_1(Z)=g(Z)=\frac{\mathrm{d}}{\mathrm{d}\theta}G_\theta(Z)|_{\theta=0}=(-q_2,q_1,-p_2,p_1)^T$, then

$$\mathbf{M}g(Z)=\nabla P(Z), \quad \text{with} \quad P(Z)=-\tfrac{1}{2}(q_1^2+q_2^2),$$
$$\mathbf{K}g(Z)=\nabla Q(Z), \quad \text{with} \quad Q(Z)=p_2q_1-p_1q_2.$$

Remark 10. There are other ways to multi-symplectify the NLS equation. For example, write $\psi=\sqrt{2\rho}\,\mathrm{e}^{\mathrm{i}\phi}$, $v=\frac{\rho_x}{2\rho}$, $w=2\rho\phi_x$, and $Y=(\rho,\phi,v,w)\in\mathbb{R}^4$. (For the equivalence we need that $Y\in(\mathbb{R}^+,\mathbb{R}^3)$ or $Y\in(\mathbb{R}^+,\mathbb{S}^1,\mathbb{R}^2)$.) Define $\tilde S(Y)=\frac{1}{2}W(2\rho)+v^2\rho+\frac{w^2}{4\rho}$. Then (11.11) is equivalent to $\mathbf{M}Y_t+\mathbf{K}Y_x=\nabla\tilde S(Y)$. This system is invariant under the action an affine group, given by

$$\tilde G_\theta(Y)=Y+\theta V, \quad \text{with} \quad V=(0,1,0,0).$$

Define the generator $\tilde g(Y)=\frac{\mathrm{d}}{\mathrm{d}\theta}\tilde G_\theta(Y)=V$, then $\mathbf{M}\tilde g(Y)=\nabla\tilde P(Y)$, with $\tilde P(Y)=-\rho$ and $\mathbf{K}\tilde g(Y)=\nabla\tilde Q(Y)$, with $\tilde Q(Y)=w$.

By checking the hypotheses of §§2, 3 and Theorem 1, we will prove the following Theorem.

Theorem 5. *The cubic defocusing nonlinear Schrödinger equation* (11.11) *with* $W'(z)=k-z$ *has a family of solitary wave solutions for* $(2b+c)^2<2(a+k-b^2)$, *the socalled dark solitons:*

$$\overline{\psi}(x,t)=\frac{\mathrm{e}^{\mathrm{i}(at+bx+\phi_0)}}{\rho_0}\left[\rho_0^2-2B^2(\tanh(B(x-ct))+1)+\mathrm{i}B(2b+c)(\tanh(B(x-ct))+1)\right],$$

with $\rho_0^2=a+k-b^2$, $4B^2=2\rho_0^2-(2b+c)^2$. *The shape of the wave has a phase shift of* $2\arctan\left(\frac{2B}{2b+c}\right)$ *as* x *goes from* $-\infty$ *to* $+\infty$.

The Evans function for the second symplectic formulation related to these dark solitons is of the following form for λ *small:*

$$D(\lambda;a,b,c)=\lambda^2\frac{\rho_0^2}{128\,B^6}+\mathcal{O}(\lambda^3). \tag{11.12}$$

Remark 11. Since the Evans function for λ large is positive, this suggests that the entire family of dark solitons is stable. Indeed, in [46] it is proved that in the case $a=0=b$, the family of waves are stable (quoted from [33]).

Remark 12. The result (11.12) can also be compared with the recent results of Kapitula & Rubin [29]. They compute the Evans function for cubic NLS linearized about a dark solitary wave using the fact that the spectral problem can be analyzed exactly. For the linearization about dark solitary waves (with $a=b=0$) they show that the Evans function satisfies

$$D(0)=D'(0)=D''(0)=0 \quad \text{and} \quad D'''(0)>0,$$

which appears to disagree with (11.12). However they agree once it is noted that the two Evans functions are related by a scale factor. In §7 we defined

$$\det(\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p}))=\lambda^{p-1}\det(\boldsymbol{\Omega}(\lambda;\mathbf{p})),$$

and used the Evans matrix without tildes in the definition of the Evans function: $D(\lambda;\mathbf{p}) = \det(\boldsymbol{\Omega}(\lambda;\mathbf{p}))$. Although [29] define the Evans function in a different way, it is related to $\det(\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p}))$. This can be seen by noting that in $\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p})$ the adjoint functions are analytic as in [29], whereas in $\boldsymbol{\Omega}(\lambda;\mathbf{p})$ the adjoint functions are meromorphic. Since $p=2$ for this example, it is apparent that an Evans function based on $\det(\widetilde{\boldsymbol{\Omega}}(\lambda;\mathbf{p}))$ will have order λ^3 as $\lambda \to 0$ and will have the same slope as the Evans function based on $\boldsymbol{\Omega}(\lambda;\mathbf{p})$.

Remark 13. There are also families of solitary wave solutions of the *focussing* NLS equation, the socalled bright solitons. However, all these solutions converge to 0 for $|x| \to \infty$ and the zero state is a fixed point of the symmetry group. As a consequence, there are eigenvalues of the linearized system at infinity which are related to the symmetry, but are not equal to zero. The stability analysis of this form of wave requires a modification of the hypotheses in §§2-3. However, it is possible to modify the development here for the analysis of this class of waves, and this will be considered in future work.

11.3. A multi-parameter family of dark solitons

First we prove the existence of a three-parameter family of solitary wave solutions, i.e., we verify the hypotheses in §2.4.

Lemma 18. *The system* (11.11) *satisfies*

1. *a multi-symplectic formulation on* $\mathbb{R}^4$ *with* $\mathbf{J}_c$ *non-degenerate and* $S:\mathbb{R}^4 \to \mathbb{R}$ *smooth. The multi-symplectic formulation of* (11.11) *is* $\mathcal{G}$*-equivariant with* $\mathcal{G}$ *a* 1*-dimensional rotation group on* $\mathbb{R}^4$*;*
2. *for all* $a,\ b,\ c \in \mathbb{R}$*, satisfying* $(2b+c)^2 < 2(a+k-b^2)$*, there exists a solitary wave shape* $\widetilde{Z}(x;a,b,c)$ *satisfying* $\mathbf{K}\widetilde{Z}_x - c\mathbf{M}\widetilde{Z}_x = \nabla S(\widetilde{Z}) - a\nabla P(\widetilde{Z}) - b\nabla Q(\widetilde{Z})$*;*
3. *the solitary wave shape* $\widetilde{Z}$ *has a state at infinity given by* $Z_0^-(a,b) = (\rho_0, 0, 0, -b\rho_0)^T$ *and* $Z_0^+(a,b,c) = G_\gamma Z_0^-$*, where* $\gamma = 2\arctan\left(\frac{2B}{2b+c}\right)$*,* $\rho_0^2 = (a+k-b^2)$ *and* $4B^2 = 2\rho_0^2 - (2b+c)^2$*;*
4. *the invariant manifold at infinity* $\mathcal{M}_\infty(a,b)$ *is non-degenerate;*
5. *the derivative of the relative equilibrium shape* $\widetilde{Z}_x$ *is exponentially decaying: for* $\beta = 2B = \sqrt{2\rho_0^2 - (2b+c)^2}$ *there exist vectors* $\Psi^\pm(a,b,c) \neq 0$ *such that*

$$\lim_{x\to\pm\infty} \mathrm{e}^{\pm\beta x}\widetilde{Z}_x = \Psi^\pm = \frac{4B^2}{\rho_0}\left(-2B, (2b+c), -b(2b+c) \pm 4B^2, -2bB \mp 2B(2b+c)\right)^T$$

and $\lim_{x\to\pm\infty} \partial_x\left(\mathrm{e}^{\pm\beta x}\widetilde{Z}_x\right) = 0$.

Proof. The matrix $\mathbf{J}_c = \mathbf{K} - c\mathbf{M}$ is invertible, since $\det(\mathbf{J}_c) = 1$ for any $c \in \mathbb{R}$.

The shape of a solitary wave solution must satisfy

$$(\mathbf{K} - c\mathbf{M})\,\widetilde{Z}_x = \nabla S(\widetilde{Z}) - a\nabla P(\widetilde{Z}) - b\nabla Q(\widetilde{Z})\,. \tag{11.13}$$

Explicitly this means

$$\begin{aligned} 0 &= (p_1)_x - c(q_2)_x + q_1(k+a-q_1^2-q_2^2) - bp_2 \\ 0 &= (p_2)_x + c(q_1)_x + q_2(k+a-q_1^2-q_2^2) + bp_1 \\ p_1 &= (q_1)_x - bq_2 \\ p_2 &= (q_2)_x + bq_1 \end{aligned} \tag{11.14}$$

By substituting the last two equations in the first two and writing $q_1 + \mathrm{i}q_2 = \rho\,\mathrm{e}^{\mathrm{i}\phi}$, we get

$$\begin{aligned} 0 &= \rho W'(\rho^2) + \rho(a + cb + \tfrac{c^2}{4}) - \tfrac{\mathsf{constant}^2}{\rho^3} + \rho_{xx} \\ \phi_x &= -\tfrac{c}{2} - b + \tfrac{\mathsf{constant}}{\rho^2}, \end{aligned}$$

where $\mathsf{constant}$ is an integration constant. In order to have a limit for $\widetilde{Z}(x)$ as $|x| \to \infty$, we choose $\mathsf{constant}$ such that $\lim_{x\to-\infty} \phi_x = 0$, hence $\mathsf{constant} = \frac{1}{2}\rho_0^2(c+2b)$.

By taking $W'(z) = k - z$, it can be shown that ρ and ϕ are given by

$$\begin{aligned}\rho^2(x) &= \tfrac{1}{2}\left((2b+c)^2 + 4B^2\tanh^2(Bx)\right) = \rho_0^2 - 2B^2\operatorname{sech}^2(Bx)\\ \phi(x) &= \phi_0 + \arctan\left(\tfrac{2B}{2b+c}\right) + \arctan\left(\tfrac{2B\tanh(Bx)}{2b+c}\right)\end{aligned}$$

with $\rho_0^2 = (a+k-b^2)$, $4B^2 = 2\rho_0^2 - (2b+c)^2$ and the condition $(2b+c)^2 < 2\rho_0^2$ must be satisfied. This implies that

$$\widetilde{Z}(x;a,b,c) = \begin{pmatrix} \rho_0 - \frac{2B^2}{\rho_0}(\tanh(Bx)+1)\\ \frac{B(2b+c)}{\rho_0}(\tanh(Bx)+1)\\ -\frac{2B^3}{\rho_0}\operatorname{sech}^2(Bx) - \frac{bB(2b+c)}{\rho_0}(\tanh(Bx)+1)\\ \frac{B^2(2b+c)}{\rho_0}\operatorname{sech}^2(Bx) + b\rho_0 - \frac{2bB^2}{\rho_0}(\tanh(Bx)+1)\end{pmatrix} \tag{11.15}$$

is a solution of (11.14) for any $a,\, b,\, c \in \mathbb{R}$ which satisfy $(2b+c)^2 < 2(a+k-b^2)$.

Since $\lim_{x\to\pm\infty} \tanh(Bx) = \pm 1$, it follows that

$$Z_0^- = \lim_{x\to-\infty} \widetilde{Z}(x;a,b,c) = \begin{pmatrix}\rho_0\\ 0\\ 0\\ b\rho_0\end{pmatrix} \quad \text{and} \quad Z_0^+ = \lim_{x\to\infty} \widetilde{Z}(x;a,b,c) = \begin{pmatrix}\rho_0 - \frac{4B^2}{\rho_0}\\ \frac{2B(2b+c)}{\rho_0}\\ -\frac{2Bb(2b+c)}{\rho_0}\\ b\left(\rho_0 - \frac{4B^2}{\rho_0}\right)\end{pmatrix}$$

Hence there is a phase shift of $2\arctan\left(\frac{2B}{2b+c}\right)$ between Z_0^- and Z_0^+.

Furthermore, $P(Z_0^\pm) = -\frac{1}{2}\rho_0^2 = -\frac{1}{2}(a+k-b^2)$ and $Q(Z_0^\pm) = b\rho_0^2 = b(a+k-b^2)$. Thus

$$\begin{vmatrix}\frac{\partial P}{\partial a} & \frac{\partial P}{\partial b}\\ \frac{\partial Q}{\partial a} & \frac{\partial Q}{\partial b}\end{vmatrix} = \begin{vmatrix}-\frac{1}{2} & b\\ b & a+k-3b^2\end{vmatrix} = -\frac{1}{2}(a+k-b^2) < -\frac{1}{4}(2b+c)^2 < 0$$

and we can conclude that the invariant manifold at infinity $\mathcal{M}_\infty(a,b)$ is nondegenerate.

Finally,

$$\widetilde{Z}_x = \frac{B^2\operatorname{sech}^2(Bx)}{\rho_0}\begin{pmatrix}-2B\\ 2b+c\\ 4B^2\tanh(Bx) - b(2b+c)\\ -2B(2b+c)\tanh(Bx) - 2bB\end{pmatrix}.$$

Thus $\lim_{x\to\pm\infty} \mathrm{e}^{\pm 2Bx}\widetilde{Z}_x(x) = \frac{4B^2}{\rho_0}(-2B, (2b+c), -b(2b+c) \pm 4B^2, -2bB \mp 2B(2b+c))^T$. Thus $\beta = 2B = \sqrt{2(a+k-b^2) - (2b+c)^2}$. And also

$$\lim_{x\to\pm\infty} \mathrm{e}^{\pm\beta x}\widetilde{Z}_{xx} = \mp\frac{8B^3}{\rho_0}(-2B, (2b+c), -b(2b+c) \pm 4B^2, -2bB \mp 2B(2b+c))^T.$$

Hence $\lim_{x\to\pm\infty} \mathrm{e}^{\pm\beta x}\widetilde{Z}_{xx} = \mp\beta \lim_{x\to\pm\infty} \mathrm{e}^{\pm\beta x}\widetilde{Z}_x$, so $\lim_{x\to\pm\infty} \partial_x\left(\mathrm{e}^{\pm\beta x}\widetilde{Z}_x\right) = 0$.

11.4. The systems at infinity

Next we analyse the linearised stability equation for $|x| \to \infty$, i.e., we verify the hypotheses in §3.5. First we determine the position of the eigenvalues of $\mathbf{A}^{\pm}(\lambda; a, b, c)$ as a function of λ and the parameters (a, b, c).

Lemma 19. *Suppose $\lambda \in \mathbb{C}$ and $\Re(\lambda)$ positive. The eigenvalue problem of $\mathbf{A}^{\pm}(\lambda; a, b, c)$ for $\mu \in \mathbb{C}$ has 2 roots $\mu_1(\lambda; a, b, c)$ and $\mu_2(\lambda; a, b, c)$ with*

$$\Re(\mu_1(\lambda; a, b, c)) \leq \Re(\mu_2(\lambda; a, b, c)) < 0\,.$$

All other eigenvalues have nonnegative real part. The eigenvalues can only cross the imaginary axis if λ is purely imaginary.

The eigenvalues $\mu_1(\lambda; a, b, c)$, $\mu_2(\lambda; a, b, c)$ are simple for $Re(\lambda) > 0$, except at one isolated real point,

$$\lambda_0 = \frac{1}{16}\sqrt{(\sqrt{9(c+2b)^2 + 32B^2} + c + 2b)^2 - 4(c+2b)^2}\left(3(c+2b) + \sqrt{9(c+2b)^2 + 32B^2}\right).$$

The eigenvalues $\mu_1(\lambda; a, b, c)$ and $\mu_2(\lambda; a, b, c)$ can be extended in an analytic way to $\lambda = 0$ with

$$\lim_{\lambda \to 0} \mu_1(\lambda; a, b, c) = \beta(a, b, c) \quad \text{and} \quad \lim_{\lambda \to 0} \mu_2(\lambda; a, b, c) = 0\,.$$

An eigenvector related to an eigenvalue μ can be written as a scalar multiple of

$$\zeta = (\mu^2, \lambda - (2b+c)\mu, \mu^3 - b\lambda + b(2b+c)\mu, b\mu^2 + \mu\lambda - \mu^2(2b+c)) \tag{11.16}$$

Hence the eigenvectors associated with the eigenvalue 0 at $\lambda = 0$ both are a multiple of $(0, 1, -b, 0)^T = \xi_1(Z_0^-)$.

Remark 14. Figure 10.2 gives a representative qualitative overview of the position of the eigenvalues μ_1 and μ_2 in the complex right-half λ-plane.

Proof. Linearising the system, we get

$$D^2S(\widetilde{Z}) - aD^2P(\widetilde{Z}) - bD^2Q(\widetilde{Z}) = \begin{pmatrix} k + a - 3\widetilde{q}_1^2 - \widetilde{q}_2^2 & -2\widetilde{q}_1\widetilde{q}_2 & 0 & -b \\ -2\widetilde{q}_1\widetilde{q}_2 & k + a - \widetilde{q}_1^2 - 3\widetilde{q}_2^2 & b & 0 \\ 0 & b & 1 & 0 \\ -b & 0 & 0 & 1 \end{pmatrix},$$

thus

$$D^2S(Z_0^-) - aD^2P(Z_0^-) - bD^2Q(Z_0^-) = \begin{pmatrix} b^2 - 2\rho_0^2 & 0 & 0 & -b \\ 0 & b^2 & b & 0 \\ 0 & b & 1 & 0 \\ -b & 0 & 0 & 1 \end{pmatrix}.$$

And

$$\det(\mathbf{B}^-\mu\mathbf{J} - \lambda\mathbf{M}) = \mu^4 - \mu^2(2\rho_0^2 - (c+2b)^2) - 2\lambda\mu(c+2b) + \lambda^2.$$

This polynomial is very similar to the one in the Boussinesq equation, with $4\delta^2$ replaced by $2\rho_0^2 - (c+2b)^2 = 4B^2$ and c replaced by $(c+2b)$. So for λ small, the eigenvalues are:

$$\begin{aligned}
\mu_1 &= -2B + \tfrac{(2b+c)}{4B^2}\lambda + \tfrac{3(2b+c)^2+4B^2}{64B^5}\lambda^2 + \mathcal{O}(\lambda^3), \\
\mu_2 &= -\tfrac{(2b+c)+\sqrt{(2b+c)^2+4B^2}}{4B^2}\lambda + \mathcal{O}(\lambda^3), \\
\mu_3 &= -\tfrac{(2b+c)-\sqrt{(2b+c)^2+4B^2}}{4B^2}\lambda + \mathcal{O}(\lambda^3), \\
\mu_4 &= 2B + \tfrac{(2b+c)}{4B^2}\lambda - \tfrac{3(2b+c)^2+4B^2}{64B^5}\lambda^2 + \mathcal{O}(\lambda^3).
\end{aligned}$$

It is straightforward to verify that the eigenvectors are given by

$$\zeta = (\mu^2, \lambda - (2b+c)\mu, \mu^3 - b\lambda + b(2b+c)\mu, b\mu^2 + \mu\lambda - \mu^2(2b+c))^T.$$

And the appropriate set Λ is

$$\Lambda = \{\lambda \in \mathbb{C} \mid \Re(\lambda) > 0,\ |\Im(\lambda)| < \Re(\lambda)\} \backslash \{\lambda_0 + iy \mid y \leq 0\},$$

with

$$\lambda_0 = \frac{1}{16}\sqrt{(\sqrt{9(c+2b)^2 + 32B^2} + c + 2b)^2 - 4(c+2b)^2}\left(3(c+2b) + \sqrt{9(c+2b)^2 + 32B^2}\right).$$

Also, $\xi_1(Z_0^-) = \rho_0(0,1,-b,0)^T$ and $\zeta_2^-(0,a,b,c) = -\frac{(c+2b)+\sqrt{2}\rho_0}{4B^2}(0,1,-b,0)^T$, thus the vector spaces $\mathrm{span}\{\zeta_2(0,a,b,c)\}$ and $\mathrm{span}\{\xi_1(Z_0^-(a,b,c))\}$ are isomorphic.

With this lemma and section 4, we have verified the hypotheses of §3.5:

1. Define the set Λ to be
$$\Lambda = \{\lambda \in \mathbb{C} \mid \Re(\lambda) > 0,\ |\Im(\lambda)| < \Re(\lambda)\} \backslash \{\lambda_0 + iy \mid y \leq 0\},$$
(compare with Figure 10.3). This set is open simply-connected and $0 \in \overline{\Lambda}$. The Evans function can not change sign on the real axis at the singular point λ_0.
2. For all $\lambda \in \Lambda$, the spectrum of $\mathbf{A}^{\pm}(\lambda, a, b, c)$ consists of 2 simple eigenvalues with strictly negative real part and 2 eigenvalues with non-negative real part.
3. The spectrum of the matrix $\mathbf{A}^{\pm}(0, a, b, c)$ is of the form $\{-\beta,\ 0,\ \beta\}$ and the eigenvalues $\pm\beta$ are simple.
4. The vector spaces $\mathrm{span}\{\zeta_2^+(0,a,b,c)\}$ and $\mathrm{span}\{\xi_1(Z_0^+(a,b,c))\}$ are isomorphic as well as the vector spaces $\mathrm{span}\{\widetilde{\eta}_2^-(0,a,b,c)\}$ and $\mathrm{span}\{\xi_1(Z_0^-(a,b,c))\}$.

11.5. The symplectic Evans matrix for λ near zero

In this section we determine the value of χ_{00}^- and $\left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \frac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right]$. With Theorem 2 and the fact that $\partial\mathcal{D} = 1$, this gives the behaviour of the Evans function for λ small.

Lemma 20. *For all a, b, $c \in \mathbb{R}$, which satisfy $(2b+c)^2 < 2(a+k-b^2)$, we have*

$$\boldsymbol{\chi}_{00}^- = \frac{\rho_0^2}{256B^7} \quad \textit{and} \quad \left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \frac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right] = 2B$$

Proof. By definition,

$$\boldsymbol{\chi}_{00}^- = \left[\Omega(\Psi^-, DG_\gamma(Z_0^-)^T\Psi^+)\right]^{-1}.$$

From Lemma 18 it follows that

$$\Psi^- = \frac{4B^2}{\rho_0}(-2B, (2b+c), -b(2b+c) - 4B^2, 2B(b+c))^T$$

and

$$DG_\gamma(Z_0^-)^T\Psi^+ = \frac{4B^2}{\rho_0}(2B, (2b+c), -b(2b+c) - 4B^2, -2B(b+c))^T,$$

hence $\Omega(\Psi^-, DG_\gamma(Z_0^-)^T\Psi^+) = \frac{256B^7}{\rho_0^2}$, thus $\boldsymbol{\chi}_{00}^- = \frac{\rho_0^2}{256B^7}$.

Next we calculate $\left[\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \frac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right]$. By definition,

$$\mathcal{I}(Z) = \frac{1}{2}\int_{-\infty}^{\infty}\omega(Z, Z_x)\,\mathrm{d}x = -\int_{-\infty}^{\infty}(q_1)_x q_2\,\mathrm{d}x + \frac{1}{2}\lim_{x\to\infty}(q_2(x)q_1(x) - q_2(-x)q_1(-x)).$$

Using the expressions for the solitary wave as given in (11.15), we get

$$\begin{aligned}\mathcal{I}(\widetilde{Z}) &= \int_{-\infty}^{\infty} \tfrac{2B^4(2b+c)}{\rho_0^2}(\tanh(Bx)+1)\mathrm{sech}^2(Bx)\,\mathrm{d}x + \tfrac{1}{2}\lim_{x\to\infty} q_2(x)q_1(x)\\ &= \tfrac{2B^3(2b+c)}{\rho_0^2}(\tanh(Bx)+1)\Big|_{-\infty}^{\infty} + \tfrac{1}{2}(\rho_0 - \tfrac{4B^2}{\rho_0})\tfrac{2B(2b+c)}{\rho_0}\\ &= B(2b+c)\end{aligned}$$

By differentiating this expression with respect to c, we get

$$\tfrac{\partial}{\partial c}\mathcal{I}(\widetilde{Z}(x;a,b,c)) = \tfrac{\rho_0^2-(2b+c)^2}{2B} = \tfrac{4B^2-(2b+c)^2}{4B}$$

Furthermore,

$$\partial_c Z_0^+ = \frac{2}{\rho_0}\begin{pmatrix}(2b+c)\\ -\frac{(2b+c)^2-\rho_0^2}{2B}\\ \frac{b((2b+c)^2-\rho_0^2)}{2B}\\ b(2b+c)\end{pmatrix} \quad \text{and} \quad Z_0^+ = \begin{pmatrix}\rho_0 - \frac{4B^2}{\rho_0}\\ \frac{2B(2b+c)}{\rho_0}\\ -\frac{2Bb(2b+c)}{\rho_0}\\ b\left(\rho_0 - \frac{4B^2}{\rho_0}\right)\end{pmatrix}$$

Hence

$$\omega(Z_0^+, \partial_c Z_0^+) = -\frac{\rho_0^2}{B}$$

and we can conclude that

$$\left[\tfrac{\mathrm{d}}{\mathrm{d}c}\mathcal{I}(\widetilde{Z}(x)) - \tfrac{1}{2}\,\omega(Z_0^+, \partial_c Z_0^+)\right] = \frac{\rho_0^2-(2b+c)^2}{2B} + \frac{\rho_0^2}{2B} = \frac{4B^2}{2B} = 2B.$$

11.6. The symplectic Evans function for large $|\lambda|$

The last thing we have to analyse is the Evans function for λ large. Just as with the Boussinesq equations, we get with $\lambda = \nu^2$ and

$$\mu = \mu_0\nu + (\frac{c}{2}+b)\mu_0^2 + \frac{(c+2b)^2-2}{8}\mu_0^3\frac{1}{\nu} + \mathcal{O}\left(\frac{1}{\nu^2}\right), \quad \nu\to\infty.$$

Following §8, we define $\alpha = \frac{1}{2}$ and a small parameter $\varepsilon = \lambda^{-\alpha} = \frac{1}{\sqrt{\lambda}}$. Furthermore, we introduce a new "space" variable $s = \frac{x}{\varepsilon}$ (i.e., $\gamma = 1$), a selfadjoint transformation $T(\varepsilon) = \mathrm{diag}(1,1,\frac{1}{\varepsilon},\frac{1}{\varepsilon})$ and transformed vectors $U^{\mathrm{new}}(s) = T(\varepsilon)U(\varepsilon s)$ and $W^{\mathrm{new}}(s) = T(\varepsilon)W(\varepsilon s)$. This implies that the linearized and the adjoint systems transform to equations (8.1) and (8.2) with

$$\mathbf{J}_c^{\mathrm{new}}(\varepsilon) = \mathbf{K} - c\varepsilon\mathbf{M}, \quad \mathbf{M}^{\mathrm{new}}(\varepsilon) = \mathbf{M}$$

and

$$\mathbf{B}^{\mathrm{new}}(s,\varepsilon) = \begin{pmatrix}\varepsilon^2(k+a-3q_1^2(\varepsilon s)-q_2^2(\varepsilon s)) & -2\varepsilon^2 q_1(\varepsilon s)q_2(\varepsilon s) & 0 & -\varepsilon b\\ -2\varepsilon^2 q_1(\varepsilon s)q_2(\varepsilon s) & \varepsilon^2(k+a-q_1^2(\varepsilon s)-3q_2^2(\varepsilon s)) & \varepsilon b & 0\\ 0 & \varepsilon b & 1 & 0\\ -\varepsilon b & 0 & 0 & 1\end{pmatrix}.$$

Hence $\mathbf{J}_c^{\mathrm{new}}(0) = \mathbf{K}$, which is an invertible matrix. Also, $\mathbf{M}^{\mathrm{new}}(0) = \mathbf{M}$ and $\mathbf{B}^{\mathrm{new}}(s,\varepsilon) - \lim_{\sigma\to\pm\infty}\mathbf{B}^{\mathrm{new}}(\sigma,\varepsilon) = \mathcal{O}(\varepsilon^2\mathrm{e}^{\mp\varepsilon Bs})$, $s\to\pm\infty$.

This means that we satisfy the conditions of Theorem 3 with $\kappa = 1$ and we can conclude that

$$\det\boldsymbol{\Omega}(\lambda;a,b,c) = 1 + \mathcal{O}\left(\frac{1}{\sqrt{\lambda}}\right), \qquad \lambda\to\infty.$$

12. Oscillatory fronts and a complex nonlinear wave equation

A class of nonlinear PDEs that arises in many applications is complex nonlinear wave equations of the form

$$\frac{\partial^2 A}{\partial t^2} = \frac{\partial^2 A}{\partial x^2} + W'(|A|^2)A\,, \quad x \in \mathbb{R}, \quad A \in \mathbb{C}, \quad t > 0\,, \tag{12.1}$$

where $W : \mathbb{R} \to \mathbb{R}$ is some smooth function. The form for $W'(|A|^2)$ most commonly encountered is when $W'(|A|^2)$ is a quadratic function; an example of this form is

$$W'(|A|^2) = \alpha(1 - |A|^2) \quad \text{with} \quad \alpha = \pm 1\,. \tag{12.2}$$

Some applications where a quadratic function $W'(|A|^2)$ appears in a such model equation are wave packets in baroclinic instability [41] and the nonlinear Kelvin-Helmholtz instability [44,7]. In this section, we give a new formulation of equation (12.1) as an equivariant Hamiltonian system on a multi-symplectic structure, and sketch how the theory of this paper applies to the stability of oscillatory fronts. As far as we are aware, the stability of oscillatory fronts for this type of equation has never been studied.

The system (12.1) has several novel features: it has a family of oscillating fronts which are biasymptotic, as $x \to \pm\infty$, to an invariant manifold diffeomorphic to a circle, and this manifold may be linearly unstable. Moreover, the full symmetry group for the system is non-abelian although the manifold at infinity which we will consider is associated with an abelian subgroup. A complete analysis of the linear stability problem for the oscillatory fronts is not given as the details would take us too far afield, but a sketch will be given of how the theory proceeds.

To reformulate (12.1) as a multi-symplectic system, first introduce complex functions B, C and D defined by

$$B = A_t + D_x\,, \quad C = A_x + D_t \quad \text{and} \quad B_x = C_t\,. \tag{12.3}$$

Then the complex wave equation (12.1) can be written

$$B_t = C_x + W'(|A|^2)A\,. \tag{12.4}$$

Now, introduce real coordinates $Z = (q,p) \in \mathbb{R}^4 \times \mathbb{R}^4$,

$$A = q_1 + \mathrm{i}q_2\,, \quad B = q_3 + \mathrm{i}q_4\,, \quad C = p_1 + \mathrm{i}p_2 \quad \text{and} \quad D = p_3 + \mathrm{i}p_4\,.$$

Then (12.3) and (12.4) can be written in the form

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z)\,, \quad Z \in \mathbb{R}^8\,, \tag{12.5}$$

where

$$\mathbf{M} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_2 & \mathbf{0} & \mathbf{0} \\ -\mathbf{I}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_2 \\ \mathbf{0} & \mathbf{0} & -\mathbf{I}_2 & \mathbf{0} \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & -\mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I}_2 \\ \mathbf{I}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$S(Z) = \tfrac{1}{2}(p_1^2 + p_2^2) - \tfrac{1}{2}(q_3^2 + q_4^2) + \tfrac{1}{2}W(q_1^2 + q_2^2)\,.$$

The two forms associated with $\mathbf{M}$ and $\mathbf{K}$ are

$$\omega = \mathbf{d}q_1 \wedge \mathbf{d}q_3 + \mathbf{d}q_2 \wedge \mathbf{d}q_4 + \mathbf{d}p_1 \wedge \mathbf{d}p_3 + \mathbf{d}p_2 \wedge \mathbf{d}p_4\,,$$

$$\kappa = \mathbf{d}p_1 \wedge \mathbf{d}q_1 + \mathbf{d}p_2 \wedge \mathbf{d}q_2 + \mathbf{d}p_3 \wedge \mathbf{d}q_3 + \mathbf{d}p_4 \wedge \mathbf{d}q_4 \,.$$

The function S is a smooth function (as smooth as W) and *both* ω and κ are closed and non-degenerate twoforms on the full phase space: $(\mathbb{R}^8, \omega)$ and $(\mathbb{R}^8, \kappa)$ are symplectic manifolds. Let $\mathcal{C} = \mathbb{R}\backslash\{-1, 1\}$; then $\det(\mathbf{K} - c\mathbf{M}) = (1 - c^2)^4 \neq 0$ for all $c \in \mathcal{C}$.

The basic symmetry group of the system is a three-parameter subgroup of the Euclidean group on $\mathbb{R}^8$. The Lie algebra is spanned by the generators

$$\begin{aligned}
\xi_1(Z) &= \mathbf{J}_2 \oplus \mathbf{J}_2 \oplus \mathbf{J}_2 \oplus \mathbf{J}_2 \quad \text{where} \quad \mathbf{J}_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\
\xi_2(Z) &= \mathbf{e}_7 = (0,0,0,0,0,0,1,0)^T \,, \\
\xi_3(Z) &= \mathbf{e}_8 = (0,0,0,0,0,0,0,1)^T \,.
\end{aligned}$$

The element $\xi_1(Z)$ generates an action of $SO(2)$ on $\mathbb{R}^8$ which is associated with the gauge-invariance of (12.1); that is $e^{i\theta}A$ is a solution of (12.1) for any $\theta \in \mathbb{R}$ whenever A is a solution. The elements $\xi_2(Z)$ and $\xi_3(Z)$ generate a two-parameter group of affine translations, associated with the fact that $S(Z)$ is independent of p_3 and p_4 and hence an arbitrary constant can be added to p_3 and p_4.

This three-parameter group is not abelian. The most interesting abelian subgroup is the subgroup $SO(2)$ generated by $\xi_1(Z)$. Therefore, we take $q = 1$ and $\mathcal{G} = SO(2)$ in which case $\mathfrak{g} = \text{span}\{\xi_1\}$. The functionals associated with the group $\mathcal{G}$ are defined by

$$\mathbf{M}\xi_1(Z) = \nabla P_1(Z) \quad \text{and} \quad \mathbf{K}\xi_1(Z) = \nabla Q_1(Z) \,,$$

from which it follows that

$$P_1(Z) = -q_1q_4 + q_2q_3 - p_1p_4 + p_2p_3 \quad \text{and} \quad Q_1(Z) = q_1p_2 - q_2p_1 + q_3p_4 - q_4p_3 \,. \tag{12.6}$$

We will consider solutions of (12.1) which are biasymptotic to an invariant manifold at infinity, $\mathcal{M}_\infty$, which is identified with an orbit of $\mathcal{G} = SO(2)$. Associated with this invariant manifold at infinity is a two-parameter family of relative equilibria of the form

$$Z(x,t) = G_{\theta(x,t)} Z_0 \,, \quad \theta(x,t) = at + bx + \theta_0 \,, \tag{12.7}$$

where $(Z_0, a, b) \in \mathbb{R}^8 \times \mathbb{R} \times \mathbb{R}$ satisfy

$$\nabla S(Z_0) = a\nabla P_1(Z_0) + b\nabla Q_1(Z_0) \quad \text{and} \quad P_1(Z_0) = \mathcal{P}_1 \,, \quad Q_1(Z_0) = \mathcal{Q}_1 \,, \tag{12.8}$$

with $(\mathcal{P}_1, \mathcal{Q}_1) \in \mathbb{R}^2$ specified.

There are two issues regarding this family of relative equilibria at infinity. First, its properties as the asymptotic state associated with the front and second, the stability of this family in the time-dependent problem. It turns out that this family has a rich bifurcation and stability structure. Before considering this structure, we establish the properties of this family of relative equilibria.

With the hypotheses $a^2 \neq b^2$ and $c^2 \neq 1$, a straightforward calculation leads to

$$Z_0(a,b) = \begin{bmatrix} \mathbf{I}_2 \\ a\mathbf{J}_2 \\ b\mathbf{J}_2 \\ \mathbf{0} \end{bmatrix} \begin{pmatrix} q_1^0 \\ q_2^0 \end{pmatrix} \,, \tag{12.9}$$

with (q_1^0, q_2^0) defined by

$$W'(r_0^2) = b^2 - a^2 \,, \quad \text{with} \quad r_0^2 = (q_1^0)^2 + (q_2^0)^2 \,. \tag{12.10}$$

For example, if $W'(r^2)=\alpha(1-r^2)$ (cf. (12.2)), then $r_0^2=(\alpha-b^2+a^2)/\alpha$.

When evaluated on the solution (12.9) and (12.10), the functionals $P_1(Z_0)$ and $Q_1(Z_0)$ take the form

$$P_1(Z_0)=-ar_0^2 \quad \text{and} \quad Q_1(Z_0)=br_0^2\,, \tag{12.11}$$

where r_0 is a function of (a,b) via (12.10). The state (Z_0,a,b) is non-degenerate when the Jacobian of (P_1,Q_1) with respect to (a,b) is non-degenerate. The Jacobian matrix is

$$\begin{bmatrix} \frac{\partial\mathcal{P}_1}{\partial a} & \frac{\partial\mathcal{P}_1}{\partial b} \\ \frac{\partial\mathcal{Q}_1}{\partial a} & \frac{\partial\mathcal{Q}_1}{\partial b}\end{bmatrix} = \frac{1}{W''(r_0^2)}\begin{pmatrix} 2a^2-r_0^2W''(r_0^2) & -2ab \\ -2ab & 2b^2+r_0^2W''(r_0^2)\end{pmatrix},$$

with determinant,

$$\det\begin{bmatrix} \frac{\partial\mathcal{P}_1}{\partial a} & \frac{\partial\mathcal{P}_1}{\partial b} \\ \frac{\partial\mathcal{Q}_1}{\partial a} & \frac{\partial\mathcal{Q}_1}{\partial b}\end{bmatrix} = -\frac{r_0^2}{W''(r_0^2)}(2W'(r_0^2)+r_0^2W''(r_0^2))\,.$$

Although most of the following results go through for a general class of potentials W, for definiteness, we will restrict attention to the potential (12.2). In this case

$$r_0^2=(\alpha+a^2-b^2)/\alpha>0 \quad \text{and} \quad \begin{bmatrix} \frac{\partial\mathcal{P}_1}{\partial a} & \frac{\partial\mathcal{P}_1}{\partial b} \\ \frac{\partial\mathcal{Q}_1}{\partial a} & \frac{\partial\mathcal{Q}_1}{\partial b}\end{bmatrix} = \frac{1}{\alpha}\begin{pmatrix} -\alpha+b^2-3a^2 & 2ab \\ 2ab & \alpha-3b^2+a^2\end{pmatrix},$$

with determinant $r_0^2(-\alpha+3b^2-3a^2)/\alpha$. This completes the description of the two-parameter family of relative equilibria on $\mathcal{M}_\infty(a,b)$.

Now, we assume there exists a travelling oscillating front which is biasymptotic to $\mathcal{M}_\infty$ as $x\to\pm\infty$ of the form

$$Z(x,t)=G_{\theta(x,t)}\mathcal{T}_{ct}\widetilde{Z}(x;a,b,c)\,, \tag{12.12}$$

with $\theta(x,t)=at+bx+\theta_0$. The shape of the front, $\widetilde{Z}(x;a,b,c)$ satisfies equation (2.13) and the asymptotic estimates in equation (2.11) as $x\to\pm\infty$. An advantage of the potential (12.2) is that an explicit solution for such an oscillating front can be constructed:

$$\widetilde{Z}=(\widetilde{q}_1,\widetilde{q}_2,\widetilde{q}_3,\widetilde{q}_4,\widetilde{p}_1,\widetilde{p}_2,\widetilde{p}_3,\widetilde{p}_4)^T\,, \tag{12.13}$$

with

$$\widetilde{q}_1(x;a,b,c)=\frac{\sqrt{2}}{\vartheta}(\ell\cos\varepsilon\tanh\ell(x+x_0)+\tfrac{1}{2}\widetilde{a}\,\sin\varepsilon)\,,$$

$$\widetilde{q}_2(x;a,b,c)=\frac{\sqrt{2}}{\vartheta}(\ell\sin\varepsilon\tanh\ell(x+x_0)-\tfrac{1}{2}\widetilde{a}\,\cos\varepsilon)\,,$$

where, with the hypothesis $c^2\neq 1$,

$$\widetilde{a}=\frac{2(ac+b)}{1-c^2}\,,\quad \widetilde{b}=\frac{\alpha+a^2-b^2}{1-c^2}>0\,,\quad 2\ell^2=\widetilde{b}-\tfrac{1}{2}\widetilde{a}^2>0 \quad \text{and} \quad \vartheta^2=\tfrac{\alpha}{1-c^2}>0\,. \tag{12.14}$$

The parameter ε is chosen so that the limit of $q_1(x;a,b,c)$ and $q_2(x;a,b,c)$ as $x\to-\infty$ is independent of c (hence Z_0^- is independent of c). The other components of the vector $\widetilde{Z}(x;a,b,c)$ are

$$\widetilde{q}_3(x;a,b,c)=-c\frac{\partial\widetilde{q}_1}{\partial x}-a\widetilde{q}_2\,,\quad \widetilde{q}_4(x;a,b,c)=-c\frac{\partial\widetilde{q}_2}{\partial x}+a\widetilde{q}_1\,,$$

$$\widetilde{p}_1(x;a,b,c)=\frac{\partial\widetilde{q}_1}{\partial x}-b\widetilde{q}_2\,,\quad \widetilde{p}_2(x;a,b,c)=\frac{\partial\widetilde{q}_2}{\partial x}+b\widetilde{q}_1\,,$$

and

$$\widetilde{p}_3(x;a,b,c)=\widetilde{p}_4(x;a,b,c)=0\,.$$

This state approaches $\mathcal{M}_\infty$ exponentially with exponent $\beta=2\ell>0$.

The linearization of (12.5) about the oscillating front (12.13) takes the form (3.2) with $U\in\mathbb{C}^8$, and the systems at infinity are associated with the matrices $\mathbf{A}^\pm(\lambda;a,b,c)$ which in this case are 8×8 matrices.

The most important hypothesis to check is the dimension of the stable and unstable subspaces of the matrix $\mathbf{A}^\pm(\lambda;a,b,c)$ as a function of λ. It is not difficult to show that the other hypotheses of §3 go through for this example.

The spectral problem for $\mathbf{A}^\pm(\lambda;a,b,c)$ is defined by the determinant:

$$\varDelta(\mu,\lambda;a,b,c)=\det[\mathbf{B}_0^\pm(a,b,c)-\mu\mathbf{J}_c-\lambda\mathbf{M}]\,. \tag{12.15}$$

Fortunately, this determinant factors into the product of three terms,

$$\varDelta(\mu,\lambda;a,b,c)=\varDelta_1(\mu,\lambda;a,b,c)\,\varDelta_2(\mu,\lambda;a,b,c)\,\varDelta_3(\mu,\lambda;a,b,c)\,,$$

with

$$\begin{aligned}
\varDelta_1(\mu,\lambda;a,b,c)&=(a-b)^2+(\mu-\lambda+c\mu)^2\,,\\
\varDelta_2(\mu,\lambda;a,b,c)&=(a+b)^2+(\mu+\lambda-c\mu)^2\,,\\
\varDelta_3(\mu,\lambda;a,b,c)&=\mu^4+(\lambda-c\mu)^4-2\mu^2(\lambda-c\mu)^2-8ab\mu(\lambda-c\mu)\\
&\qquad+2(-\alpha+3b^2-a^2)\mu^2+2(\alpha-b^2+3a^2)(\lambda-c\mu)^2\,.
\end{aligned}$$

With the hypotheses $c^2\neq 1$ and $a^2\neq b^2$, the first two products lead to the four μ-roots:

$$\mu_{5,6}=\frac{\lambda}{c+1}\pm\mathrm{i}\frac{(a-b)}{c+1}\quad\text{and}\quad\mu_{7,8}=\frac{\lambda}{c-1}\pm\mathrm{i}\frac{(a+b)}{c-1}\,.$$

When $\lambda=0$, these four roots are non-zero and purely imaginary, and they are distinct if $a+bc\neq 0$. These four roots are related to the Euclidean symmetry group of translations in the p_3 and p_4 variables. These symmetries do not commute with the group $\mathcal{G}$, hence they contribute non-zero eigenvalues. When λ is real and positive, these four μ-roots are perturbed off the imaginary axis with the direction of perturbation dependent on the sign of $(c+1)$ and $(c-1)$.

Restricting c to be non-negative, all four roots perturb to the right-half λ-plane when $c>1$. On the other hand, when $0<c<1$, $\mu_{5,6}$ perturb to the right-half λ-plane and $\mu_{7,8}$ perturb to the left-half λ-plane.

The most important μ-roots of $\varDelta(\mu,\lambda;a,b,c)=0$ are the four roots defined by $\varDelta_3(\mu,\lambda;a,b,c)=0$. Using the theory of §4, $\varDelta_3$ can be written in the interesting form

$$\varDelta(\mu,\lambda;a,b,c)=\left(\mu^2-(\lambda-c\mu)^2\right)^2-2\alpha\begin{pmatrix}\lambda-c\mu\\ \mu\end{pmatrix}^T\begin{bmatrix}\partial_a\mathcal{P}_1 & \partial_b\mathcal{P}_1\\ \partial_a\mathcal{Q}_1 & \partial_b\mathcal{Q}_1\end{bmatrix}\begin{pmatrix}\lambda-c\mu\\ \mu\end{pmatrix}\,. \tag{12.16}$$

This form shows the importance of the two-parameter family of relative equilibria in determining the spectral properties of $\mathbf{A}^\pm(\lambda;a,b,c)$.

When $\lambda=0$, $\varDelta_3$ reduces to

$$\varDelta_3(\mu,0;a,b,c)=(1-c^2)^2\mu^2(\mu^2-\beta^2)\,,$$

where

$$\beta^2=\frac{2\alpha}{(1-c^2)^2}\begin{pmatrix}-c\\ 1\end{pmatrix}^T\begin{bmatrix}\partial_a\mathcal{P}_1 & \partial_b\mathcal{P}_1\\ \partial_a\mathcal{Q}_1 & \partial_b\mathcal{Q}_1\end{bmatrix}\begin{pmatrix}-c\\ 1\end{pmatrix}\,. \tag{12.17}$$

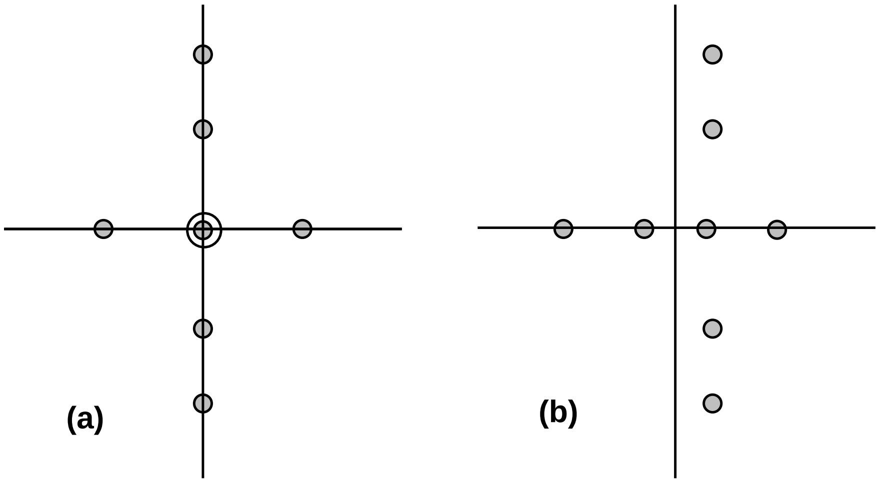

Fig. 12.1. Position of the eigenvalues of $\mathbf{A}^{\pm}(\lambda; a, b, c)$ for $c > 1$ and $a > b > 0$ in the complex μ-plane for (a) $\lambda = 0$ and (b) λ real, positive and small.

In fact, writing out this expression shows that $\beta = 2\ell$, with ℓ defined in equation (12.14). Since $\ell^2 > 0$ by hypothesis, it follows that $\Delta_3(\mu, 0; a, b, c) = 0$ has two zero roots and a pair of hyperbolic real roots. The expression (12.17) is a good example which shows how the hypothesis (4.11) may be constrained to be satisfied automatically by other constraints on c. The requirement $\beta^2 > 0$, which is a requirement for a solitary wave or front to exist in this case, ensures that (4.11) is indeed satisfied.

Therefore for $\lambda = 0$, $c^2 \neq 1$ and $a^2 \neq b^2$ the eight μ-roots of $\Delta(\mu, \lambda; a, b, c) = 0$ take the position shown in Figure 12.1(a).

To establish the positions of the μ-roots when $\lambda \neq 0$ it remains to check how the two eigenvalues 0 perturb. But, their movement for $|\lambda|$ small is dictated precisely by the geometric perturbation theory of §4. In fact, depending on the value of a and b, every region in Figure 4.1, for the position of the two eigenvalues 0, is possible. The most interesting case is when $c > 1$, and for this case there is an open set of (a, b) values for which the two eigenvalues 0 perturb in opposite directions. The qualitative position of all the eigenvalues for this case is shown in Figure 12.1(b). In this case $p = 2$ with $q = 1$. However for other (a, b, c) and λ values it is possible to also have the cases $p = 1$, $p = 3$, $p = 4$, and $p = 0$!

Another intriguing property of this example is the linear stability properties of the asymptotic state. The linear stability of the asymptotic state can be determined using the spectral function $\Delta(\mu, \lambda; a, b, c)$. We say that a member of the family of relative equilibria (12.7)-(12.8) is linearly unstable if – for some $\mu \in \mathrm{i}\mathbb{R}$ – there exists $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > 0$ satisfying $\Delta(\mu, \lambda; a, b, c) = 0$.

The λ-roots of $\Delta_1 = 0$ and $\Delta_2 = 0$ with $\mu \in \mathrm{i}\mathbb{R}$ are purely imaginary for all parameter values. Therefore any instability will be due to the λ-roots of $\Delta_3 = 0$. First, set $\mu = 0$, then

$$\Delta_3(0, \lambda; a, b, c) = \lambda^2(\lambda^2 - 2\alpha\partial_a\mathcal{P}_1)\,.$$

When $\alpha\partial_a\mathcal{P}_1$ is positive (i.e. $3a^2 < b^2 - \alpha$) the state is unstable to a class of perturbations which is spatially homogeneous. It is easy to check that there are open regions in the (a, b) plane for which this property is satisfied (cf. Figure 12.2, region II). A second class of instabilities – the modulational instability – corresponds to $\mu = \mathrm{i}\nu$ with $\nu \in \mathbb{R}$ and $|\nu|$ small. In this case

$$\Delta_3(\mathrm{i}\nu, \lambda; a, b, c) = -2\alpha\left[(\lambda - \mathrm{i}c\nu)^2\partial_a\mathcal{P}_1 + \mathrm{i}\nu(\lambda - \mathrm{i}c\nu)(\partial_b\mathcal{P}_1 + \partial_a\mathcal{Q}_1) - \nu^2\partial_b\mathcal{Q}_1\right] + \cdots\,,$$

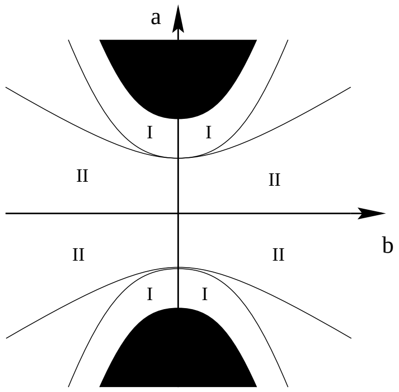

Fig. 12.2. Regions in the (a, b) plane where solutions of the type (12.7)-(12.8) with $\alpha = -1$ exist and regions where they are modulation unstable (Region I) and spatially-uniform unstable (Region II).

and so a modulational instability occurs when

$$\det \begin{bmatrix} \partial_a \mathcal{P}_1 & \partial_b \mathcal{P}_1 \\ \partial_a \mathcal{Q}_1 & \partial_b \mathcal{Q}_1 \end{bmatrix} = \frac{\alpha - 3b^2 + 3a^2}{\alpha} > 0\,, \tag{12.18}$$

just as derived in §4. There are open regions in the (a, b)-plane for which this condition is also satisfied. In Figure 12.2 we show the (a, b) plane for $\alpha < 0$ showing the two regions where the asymptotic state is linearly unstable. The solid region corresponds to the values of (a, b) where no solutions of the form (12.7)-(12.8) exist.

The momentum for this example can be explicitly constructed and the theory of §4 shows that $\partial\mathcal{D} = 1$ for this example. Because of the bifurcation and stability structure of the relative equilibria at infinity, it is clear that the family of oscillating fronts will also have a rich bifurcation and stability structure. There are many cases to consider! The details will be given elsewhere.

References

1. R. Abraham and J.E. Marsden. *Foundations of Mechanics, Second Edition* Benjamin-Cummings Publ: Boston, 1978.
2. J.C. Alexander, R. Gardner, and C.K.R.T. Jones. A topological invariant arising in the stability analysis of travelling waves. *J. Reine Angew. Math.*, 410:167–212, 1990.
3. J.C. Alexander and C.K.R.T. Jones. Existence and stability of asymptotically oscillatory double pulses. *J. Reine Angew. Math.*, 446:49–79, 1994.
4. J.C. Alexander and R. Sachs. Linear instability of solitary waves of a Boussinesq–type equation: A computer assisted computation. *Nonl. World*, 2:471–507, 1995.
5. J.M. Arnold. Stability theory for periodic pulse train solutions of the nonlinear Schrödinger equation. *IMA J. Appl. Math.* 52:123–140, 1994.
6. T.B. Benjamin. The stability of solitary waves. *Proc. Roy. Soc. Lond. A*, 328:153–183, 1972.
7. T.B. Benjamin and T.J. Bridges. Reappraisal of the Kelvin-Helmholtz problem. Part 2. Interaction of the Kelvin-Helmholtz, superharmonic and Benjamin-Feir instabilities. *J. Fluid Mech.*, 333: 327–373, 1997.
8. D.J. Benney and A.C. Newell. The propagation of nonlinear wave envelopes. *J. Math. and Phys.*, 46: 133–139, 1967.
9. E. Binz, J. Śniatycki and H. Fischer. *Geometry of Classical Fields*, North Holland Elsevier, Amsterdam, 1988.
10. J.L. Bona and R.L. Sachs. Global existence of smooth solutions and stability of solitary waves for a generalized Boussinesq equation. *Comm. Math. Phys.*, 118:15–29, 1988.
11. J.L. Bona. On the stability of solitary waves. *Proc. Roy. Soc. Lond. A*, 344:363–374, 1975.
12. T.J. Bridges. Hamiltonian bifurcations of the spatial structure for coupled nonlinear Schrödinger equations. *PhysicaD* 57:375–394, 1992.

13. T.J. Bridges. Multi-symplectic structures and wave propagation. *Math. Proc. Camb. Phil. Soc.*, 121: 147–190, 1997.
14. T.J. Bridges. A geometric formulation of the conservation of wave action and its implications for signature and the classification of instabilities. *Proc. Roy. Soc. Lond. A*, 453:1365–1395, 1997.
15. T.J. Bridges. Toral-equivariant partial differential equations and quasiperiodic patterns. *Nonlinearity*, 11:467–500, 1998.
16. T.J. Bridges and G. Derks. Unstable eigenvalues, and the linearisation about solitary waves and fronts with symmetry. *Proc. Roy. Soc. Lond. A*, 455:2427–2469, 1999.
17. T.J. Bridges and G. Derks. Hodge duality and the Evans function. *Phys. Lett. A*, 251:363–372, 1999.
18. W.A. Coppel. *Stability and asymptotic behavour of differential equations.* D.C. Heath and Company, 1965.
19. R.W.R. Darling. *Differential Forms and Connections.* Cambridge University Press, 1994.
20. M.S.P. Eastham. *The Asymptotic Solution of Linear Differential Systems.* Clarendon Press: Oxford, 1989.
21. J.W. Evans. Nerve axon equations, IV: The stable and unstable impulse. *Indiana Univ. Math. J.*, 24:1169-1190, 1975.
22. J.W. Evans and J. Feroe. Local stability theory of the nerve impulse. *Biosci.*, 37:23–50, 1977.
23. R. A. Gardner & K. Zumbrun. The gap lemma and geometric criteria for instability of viscous shock profiles. *Comm. Pure Appl. Math.* 51:797–855, 1998.
24. M. Grillakis. Linearized stability for nonlinear Schrödinger and Klein-Gordon equations. *Comm. Pure Appl. Math.* 41:747-774, 1988.
25. M. Grillakis, J. Shatah & W. Strauss. Stability theory of solitary waves in the presence of symmetry, I. *J. Funct. Anal.*, 74:160–197, 1987.
26. M. Grillakis, J. Shatah & W. Strauss. Stability theory of solitary waves in the presence of symmetry, II. *J. Funct. Anal.*, 94:308–348, 1990.
27. D. D. Holm, J. E. Marsden, T. S. Ratiu & A. Weinstein. Nonlinear stability of fluid and plasma equilibria. *Phys. Rep.* 123:1–116, 1985.
28. C. K. R. T. Jones. Instability of standing waves for nonlinear Schrödinger type equations. *Ergodic Theory and Dyn. Sys.* 8:119–128, 1988.
29. T. Kapitula and J. Rubin. Existence and stability of standing hole solutions to complex Ginzburg-Landau equations. *Nonlinearity* 13:77–112, 2000.
30. T. Kapitula and B. Sandstede. Stability of bright solitary-wave solutions to perturbed nonlinear Schrödinger equations. *PhysicaD* 124:58–103, 1998.
31. T. Kato. *Perturbation theory for linear operators*, volume 132 of *Grundlehren de mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2nd (2nd corrected printing) edition, 1984.
32. J. Kijowski and W.M. Tulczyjew. *A Symplectic Framework for Field Theories*, Lecture Notes in Physics **107**, Springer-Verlag, Heidelberg, 1979.
33. E.A. Kuznetsov and J.J. Rasmussen. Instability of two-dimensional solitons and vortices in defocusing media *Physical Review E*, **51**:4479–4484, 1995.
34. Y. A. Li & K. Promislow. Structural stability of non-ground state traveling waves of coupled nonlinear Schrödinger equations. *PhysicaD* 124:137–165, 1998.
35. J. H. Maddocks & R. L. Sachs, On the stability of KdV multi-solitons. *Comm. Pure & Appl. Math.* 45:867–901, 1993.
36. J. E. Marsden and S. Shkoller. Multisymplectic geometry, covariant Hamiltonians and water waves. *Math. Proc. Camb. Phil. Soc.* 125:553–575, 1999.
37. C.R. Menyuk, R. Schiek & L. Torner. Solitary waves due to $\chi^{(2)} : \chi^{(2)}$ cascading. *J. Opt. Soc. Am. B* 11:2434–2443, 1994.
38. S. Nii. An extension of the stability index for traveling-wave solutions and its application to bifurcations. *SIAM J. Math. Anal.* 28:402–433, 1997.
39. D.F. Parker. Exact representations for coupled bright and dark solitary waves of quadratically nonlinear systems. *J. Opt. Soc. Am. B* 15:1061–1068, 1998.
40. D.F. Parker and G.K. Newboult. Coupled nonlinear Schrödinger equations arising in fibre optics. *J. Phys. Colloq.* 50:137–146. 1989.
41. J. Pedlosky. Finite-amplitude baroclinic wave packets. *J. Atmos. Sci.*, 29:680, 1972.
42. R.L. Pego and M.L. Weinstein. Eigenvalues, and instabilities of solitary waves. *Phil. Trans. Roy. Soc. Lond. A*, 340:47–94, 1992.
43. M. Weinstein, Modulational stability of ground states of nonlinear Schrödinger equations. *SIAM J. Math. Anal.* 16:472–491, 1985.
44. M.A. Weissman. Nonlinear wave packets in the Kelvin-Helmholtz instability. *Phil. Trans. R. Soc. Lond. A*, 290: 639–685, 1979.
45. V. E. Zakharov. Instability of self-focussing of light. Sov. Phys. JETP 26:994–998, 1968.

46. V.E. Zakharov and A.B. Shabat. Exact theory of two-dimensional self-focusing and one-dimensional self-modulating waves in nonlinear media. *Sov. Phys. JETP*, 34: 62–69, 1973. [trs. of *Zh. Eksp. Teor. Fiz.*, 61: 118-134, 1972].
47. V. E. Zakharov & E. I. Schulman. To the integrability of the system of two coupled nonlinear Schrödinger equations. *PhysicaD* 4:270–274, 1982.

Department of Mathematics and Statistics, University of Surrey, Guilford, Surrey GU2 7XH, England