

# Unstable eigenvalues and the linearization about solitary waves and fronts with symmetry

BY THOMAS J. BRIDGES AND GIANNE DERKS

*Department of Mathematics and Statistics, University of Surrey,  
Guildford GU2 5XH, UK*

*Received 13 November 1997; revised 25 September 1998; accepted 4 December 1998*

The linear stability of solitary-wave or front solutions of Hamiltonian evolutionary equations, which are equivariant with respect to a Lie group, is studied. The organizing centre for the analysis is a multisymplectic formulation of Hamiltonian partial differential equations (PDEs) where a distinct symplectic operator is assigned for time and space. This separation of symplectic structures leads to new characterizations of the following components of the analysis. The states at infinity are characterized as manifolds of relative equilibria associated with the spatial symplectic structure. The momentum of the connecting orbit, or shape of the solitary wave, considered as a heteroclinic orbit in a phase-space representation, is given a new characterization as a one-form on the tangent space to the heteroclinic manifold and this one-form is a restriction of the temporal symplectic structure. For the linear stability analysis, a new symplectic characterization of the Evans function and its derivatives are obtained, leading to an abstract geometric proof of instability for a large class of solitary-wave states of equivariant Hamiltonian evolutionary PDEs. The theory sheds new light on several well-known models: the gKdV equation, a Boussinesq system and a nonlinear wave equation. The generalization to solitary waves associated with multidimensional heteroclinic manifolds and the implications for solitary waves or fronts which are biasymptotic to invariant manifolds such as periodic states are also discussed.

**Keywords:** stability; Evans function; multisymplectic structure; wave equations; Hamiltonian structure; solitary waves

## 1. Introduction

The calculus of variations, critical-point theory and Hamiltonian structure have played an important role in the analysis of the stability and instability of solitary-wave solutions of translation-invariant Hamiltonian evolution equations. For example, the characterization of solitary-wave states 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 (see, for example, Benjamin 1972; Bona 1975; Holm *et al.* 1985; Grillakis *et al.* 1987, 1990; Maddocks & Sachs 1993, and references therein). However, for many Hamiltonian evolution equations, particularly coupled systems of partial differential equations (PDEs), even though the characterization of the 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 framework for analysing spectral problems associated with the linearization about a solitary-wave or front solution was introduced by Evans (1975) and extended by Jones (1984), Alexander *et al.* (1990) and Pego & Weinstein (1992). 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  $\lambda_0$  is an unstable eigenvalue associated with the linearization about a solitary wave. Pego & Weinstein (1992) adapted the Evans-function theory for Hamiltonian evolutionary equations and proved, for three families of Hamiltonian evolutionary equations, that  $D(\lambda)$  satisfies

$$D(0) = D'(0) = 0, \quad \text{sgn } D''(0) = \text{sgn } \frac{dI}{dc}, \quad (1.1)$$

and that  $D(\lambda) \rightarrow 1$  as  $\lambda \rightarrow +\infty$  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 two-fold: it uses 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,  $dI/dc < 0$  is a sufficient condition for the existence of at least one unstable eigenvalue. However, in the theory of Pego & Weinstein (1992), the system of ordinary differential equations (ODEs) associated with the spectral problem had no special structure, and explicit calculations for each particular Hamiltonian evolution equation were therefore necessary in parts of the proof of (1.1). This limits applicability of (1.1) to the particular PDEs studied and to cases where the solitary wave is known explicitly.

In this paper we give an abstract proof of (1.1) and generalize it in several directions. The cornerstone of the theory is a multisymplectic formulation of Hamiltonian PDEs, where a distinct symplectic operator is assigned for time and space. With two symplectic structures we are able to give a geometric characterization of the ODE associated with the spectral problem. To be explicit, the Hamiltonian PDE will be written in the form

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

with  $\mathbf{M}$  and  $\mathbf{K}$  skew-symmetric constant matrices in  $\mathbb{R}^{2n \times 2n}$  and  $S$  a smooth functional in  $\mathbb{R}^{2n}$ .

To illustrate how such a multi-symplectic Hamiltonian PDE can be derived from a Hamiltonian evolution equation, consider the generalized Korteweg–de Vries (gKdV) equation, written in Hamiltonian form:

$$u_t = -\frac{1}{2} \frac{\partial}{\partial x} \frac{\delta H}{\delta u}, \quad \text{with } H(u) = \int_{-\infty}^{+\infty} [F(u) - \frac{1}{2}u_x^2] dx, \quad (1.3)$$

where  $F(u)$  is some smooth function. The gKdV equation can be reformulated as a Hamiltonian system on a multisymplectic structure as follows. With new variables,

$\phi$ ,  $v$  and  $w$ , defined below, the gKdV equation can be written as a first-order system of PDEs,

$$u = \phi_x, \quad v = u_x, \quad w = \phi_t + v_x + f(u), \quad u_t + w_x = 0, \tag{1.4}$$

or, with  $Z = (\phi, u, v, w) \in \mathbb{R}^4$ ,

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z), \quad Z \in \mathbb{R}^4, \tag{1.5}$$

where

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{K} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \tag{1.6}$$

and

$$S(Z) = \frac{1}{2}v^2 - uw + F(u), \quad f(u) = F'(u). \tag{1.7}$$

The skew-symmetric operators  $\mathbf{K}$  and  $\mathbf{M}$  define the two fundamental presymplectic structures on  $\mathbb{R}^4$ , with two-forms

$$\kappa = dv \wedge du + d\phi \wedge dw \quad \text{and} \quad \omega = d\phi \wedge du. \tag{1.8}$$

The two-form  $\kappa$  is in fact a symplectic form (closed and non-degenerate two-form) on the full four-dimensional phase space and  $\omega$  defines a rank-two symplectic structure associated with the time dependence. A third symplectic structure, with defining two-form  $\Omega = \kappa - c\omega$ , is associated with a restriction of (1.5) to a moving frame travelling with speed  $c$ . Further aspects of this reformulation of the gKdV equation are covered in §6. Other examples are given in §§5, 7 and 8. However, the proof of instability will not rely on properties of any particular equation; it is based on the abstract form of the equations (1.2).

As an example of how the multisymplectic formulation (1.5) enters the analysis and linear stability of solitary waves, consider the solitary-wave solution of the gKdV equation. For suitable functions  $f$ , it is well known that there exists a solution of (1.3) of the form

$$u(x, t) = \mathcal{T}_{ct}\hat{u}(x) \stackrel{\text{def}}{=} \hat{u}(x - ct) \tag{1.9}$$

which decays exponentially as  $x \rightarrow \pm\infty$  and satisfies the ODE

$$\frac{d}{dx}[\hat{u}_{xx} - 2c\hat{u} + f(\hat{u})] = 0. \tag{1.10}$$

Linearization of (1.3) about this solitary-wave state and a spectral ansatz  $u(x, t) = \hat{u}(x - ct) + v(x - ct)e^{\lambda t}$  leads to

$$-\frac{1}{2}[v_{xxx} + (f'(\hat{u})v)_x - 2cv_x] = \lambda v, \tag{1.11}$$

where  $\lambda \in \mathbb{C}$  is the spectral parameter. For the Evans-function theory, this spectral problem is formulated as a first-order system of ODEs; a natural way of doing this is

$$V_x = \mathbf{A}(x, \lambda)V, \quad V \in \mathbb{C}^3, \quad x \in \mathbb{R}, \tag{1.12}$$

where  $V = (v, v_x, v_{xx})$ . However, this equation, considered as a dynamical system in the  $x$ -variable, does not have a Hamiltonian structure, nor is it obvious that it might be Hamiltonian. On the other hand, using the multisymplectic formulation of the

equation, both the spatial and temporal symplectic structures are retained in the spectral problem. Linearizing (1.5) about the solitary wave results in the spectral problem

$$\mathbf{J}_c Z_x = [D^2 S(\hat{Z}) - \lambda \mathbf{M}] Z, \quad \text{with } \mathbf{J}_c = \mathbf{K} - c\mathbf{M}, \quad (1.13)$$

where  $\hat{Z}$  is the representation of the solitary wave in the  $Z$ -coordinates in (1.5). Since  $\mathbf{J}_c$  is skew-symmetric and non-degenerate, this system is Hamiltonian, with respect to the  $x$ -derivative, when  $\lambda = 0$ , and  $\mathbf{M}$  is associated with the temporal symplectic structure.

Equations (1.10) and (1.12) (or (1.13)) show that two important components of the analysis of solitary waves and their stability for (1.3) can be formulated as dynamical systems with respect to the spatial variable  $x$ . Indeed, existence of solitary-wave solutions proceeds most easily by looking for homoclinic or heteroclinic orbits of the ‘dynamical system’ (1.10), and the Evans-function theory can be interpreted as a transversal intersection criterion for the stable and unstable manifolds associated with (1.12). The system (1.10) is evidently Hamiltonian—the symplectic form is  $d\hat{v} \wedge d\hat{u}$  with  $\hat{v} = \hat{u}_x$ —but (1.12) is not, whereas the formulation (1.13) retains the geometric information from the two symplectic structures in (1.5).

The purpose of this paper is not to study the linear stability problem for particular PDEs. Our argument is that a multisymplectic formulation of the evolution equation and the spectral problem associated with the linearization about a solitary-wave state are general properties of a large class of Hamiltonian evolution equations. With this formulation, a new geometric characterization of the linearized stability equation is possible, leading to an abstract proof of a generalization of (1.1)—a proof that relies on the interplay between the two symplectic structures (temporal,  $\omega$ , and spatial,  $\kappa$  or  $\Omega$ ) but not on properties of a particular equation.

In this paper a multisymplectic manifold is a manifold with a collection of closed two-forms, each of which is non-degenerate on a submanifold. The concept of a multisymplectic manifold, in the sense used here was introduced in Bridges (1995, 1997*a, b*), and shown to be a natural dynamical systems framework for analysing multidimensional patterns. The term multisymplectic manifold is also used in classical field theory, such as the theory of relativity, with a different meaning. In field theory a ‘multisymplectic manifold’ is a manifold with a single higher-order differential form, also known as the Cartan form (cf. Binz *et al.* 1988, and references therein). In order to give an invariantive framework for the collection of presymplectic two-forms used here, an interesting approach is to concatenate them into a single three-form—a ‘metasymplectic form’—equivalent to the exterior derivative of the Cartan form in field theory (cf. Marsden & Shkoller 1999).

The concept of multisymplecticity used here is not to be confused with the concept of a bi-Hamiltonian or multi-Hamiltonian system (cf. Olver 1986, §7.3). Bi-Hamiltonian systems involve two distinct Hamiltonian structures for the time evolution only and are a precursor to integrability. Multisymplecticity assigns a distinct symplectic operator to each spatial direction as well as time and is independent of the question of integrability.

The importance of the abstract multisymplectic formulation (1.2) stems from the fact that it *a priori* decomposes each feature of the analysis. In §2, the abstract formulation (1.2) is taken as a starting point and implications of this formulation for encoding geometric information about an existing solitary wave or front are deduced.

Symmetry arises naturally in many model equations for solitary waves or fronts. Therefore we assume that the system in the form (1.2) is equivariant with respect to a  $q$ -dimensional Lie group,  $G$ . This Lie group is assumed to be abelian and the action is either compact (a subgroup of the orthogonal group on  $\mathbb{R}^{2n}$ ) or a group of affine translations. (This hypothesis is not essential but is sufficient for the present purposes and covers a wide range of applications.)  $G$ -equivariance of systems of the form (1.2) can be stated succinctly as  $G$ -invariance of  $\omega$ ,  $\kappa$  and the functional  $S$ .

An implication of symmetry for systems of the form (1.2) is that a multisymplectification of Noether theory can be used to decompose the effect of symmetry into an  $\omega$ -symplectic part and a  $\kappa$ -symplectic part. In other words, according to the classical Noether theory for symplectic systems, the symplectic flow of the group generates an invariant functional. However, in the multisymplectic setting, there is a multisymplectic flow of the group which generates a family of functionals (cf. Bridges 1997a, Appendix C). For example, for the system (1.2), we have, for each one-parameter subgroup of  $G$  with infinitesimal generator  $\mathfrak{g}$ , functionals  $P$  and  $Q$  such that

$$\mathbf{M}\mathfrak{g}(Z) = \nabla Q(Z) \quad \text{and} \quad \mathbf{K}\mathfrak{g}(Z) = \nabla P(Z). \quad (1.14)$$

A consequence is that there are two types of relative equilibria: the usual ‘temporal’ relative equilibria and the new ‘spatial’ relative equilibria. Explicitly, if  $\mathcal{G}_\theta$  is the action of a one-parameter subgroup of  $G$ , the temporal relative equilibrium is of the form  $u(x, t) = \mathcal{G}_{ct+\theta_0}\tilde{u}_1$  and a spatial relative equilibrium is of the form  $u(x, t) = \mathcal{G}_{ax+x_0}\tilde{u}_2$  (see Appendix A for elaboration).

This generalization of relative equilibria allows for a generalization of the possible asymptotic ( $x \rightarrow \pm\infty$ ) states of the solitary wave or front. Classically, a solitary wave is a state which is biasymptotic to the trivial state as  $x \rightarrow \pm\infty$ , and a front is asymptotic to distinct constant states at plus and minus infinity. In this paper we will consider a larger class of states at infinity: travelling waves which are biasymptotic to an invariant manifold,  $\mathcal{M}_\infty$ , of ‘spatial relative equilibria’ at infinity. For example, if the Lie group consists of a  $q$ -parameter affine group of translations in the phase space, the solitary wave can be chosen to be asymptotic to a non-compact invariant manifold at infinity consisting of a  $q$ -dimensional affine plane, or if the Lie group consists of a rotation, the solitary wave can be asymptotic to a spatially rotating wave.

A simple example of this generalization can already be seen with gKdV. Although it is usual to study solitary waves of gKdV which are biasymptotic to the trivial state, Gardner (1997) found that when the solitary wave is biasymptotic to a non-zero constant state at infinity, the stability is dependent on the value of this constant. In §6 this non-vanishing state at infinity will be characterized as a spatial relative equilibrium associated with a natural Lie group symmetry of gKdV: the state at infinity is an invariant manifold consisting of a line, and the  $x$ -direction flow along this invariant manifold is associated with the above non-zero constant state.

In the sequel, a solitary wave or front will be assumed to exist and the stability theory will use only natural abstract properties of the given wave. The characterization of the solitary wave or front will be divided into two parts: the asymptotic state at infinity which is given a new characterization as a spatial relative equilibrium associated with a spatial symplectic structure and, secondly, the ‘shape’ of the solitary wave or front which is the part connecting the states at plus and minus infinity. The characterization of the shape leads to a new geometric interpretation

of the momentum as a one-form on the homoclinic or heteroclinic manifold which represents the solitary wave or front in the overall  $2n$ -dimensional phase space.

In §§ 3 and 4 the linear stability problem is formulated and analysed. The linear stability problem is reduced to an ODE of the form

$$U_x = \mathbf{A}(x, \lambda)U, \quad U \in \mathbb{C}^{2n}, \quad \lambda \in \Lambda \subset \mathbb{C}, \tag{1.15}$$

where  $\Lambda$  is a suitably chosen subset of the complex  $\lambda$ -plane with  $\mathbf{A}(x, \lambda)$  satisfying

$$\lim_{x \rightarrow \pm\infty} \mathbf{A}(x, \lambda) = \mathbf{A}_{\pm\infty}(\lambda). \tag{1.16}$$

With suitable hypotheses on the spectrum of  $\mathbf{A}_{\pm\infty}(\lambda)$  and the way the limit (1.16) is approached, this system is in standard form for the application of the Evans-function theory. However, in this case, since (1.15) is obtained from (1.2), there is considerable geometric information encoded in  $\mathbf{A}(x, \lambda)$ ; in particular,  $\mathbf{A}(x, \lambda)$  has a multisymplectic decomposition,

$$\mathbf{A}(x, \lambda) = (\mathbf{K} - c\mathbf{M})^{-1}[\mathbf{B}(x, c) - \lambda\mathbf{M}],$$

where  $\mathbf{B}(x, c)$  is a symmetric  $(x, c)$ -dependent matrix and  $\mathbf{K}$  and  $\mathbf{M}$  are associated with the two-forms  $\omega$  and  $\kappa$ .

To fix hypotheses on the spectrum of  $\mathbf{A}_{\pm\infty}(\lambda)$ , define

$$\left. \begin{aligned} E_{\pm\infty}^s(\lambda) &= \{ \xi \in \mathbb{C}^{2n} : \lim_{x \rightarrow +\infty} e^{\mathbf{A}_{\pm\infty}(\lambda)x} \xi = 0, \lambda \in \Lambda \subset \mathbb{C} \} \\ E_{\pm\infty}^u(\lambda) &= \{ \xi \in \mathbb{C}^{2n} : \lim_{x \rightarrow -\infty} e^{\mathbf{A}_{\pm\infty}(\lambda)x} \xi = 0, \lambda \in \Lambda \subset \mathbb{C} \} \end{aligned} \right\} \tag{1.17}$$

and let  $E_{\pm\infty}^c$  be the complement of  $E_{\pm\infty}^s(\lambda) + E_{\pm\infty}^u(\lambda)$  in  $\mathbb{C}^{2n}$ . We will prove that, when the state at infinity is an invariant manifold of relative equilibria associated with the group  $G$ , the matrices  $\mathbf{A}_{-\infty}(\lambda)$  and  $\mathbf{A}_{+\infty}(\lambda)$  have the same spectrum. Therefore, although  $\mathbf{A}_{-\infty}(\lambda) \neq \mathbf{A}_{+\infty}(\lambda)$  in general, the dimensions of  $E^u$ ,  $E^s$  and  $E^c$  at plus and minus infinity are equal.

We take the following hypotheses on the systems at infinity,

$$\dim E_{\pm\infty}^s(0) = \dim E_{\pm\infty}^u(0) = 1, \quad \text{thus } \dim E_{\pm\infty}^c(0) = 2(n - 1), \tag{1.18}$$

and the hypothesis when  $\lambda \neq 0$  is

$$\min\{\dim E_{\pm\infty}^s(\lambda), \dim E_{\pm\infty}^u(\lambda)\} = 1, \quad \forall \lambda \in \Lambda \subset \mathbb{C}. \tag{1.19}$$

The  $\Omega$ -symplectic structure forces the dimensions of  $E_{\pm\infty}^s(0)$  and  $E_{\pm\infty}^u(0)$  to be equal and the dimension of  $E_{\pm\infty}^c(0)$  to be even. Furthermore—if all the generators of the symmetry group act non-trivially on the state at infinity—it follows that  $\dim E_{\pm\infty}^c(0) \geq 2q$ . It is natural to assume that the centre subspace consists entirely of zero eigenvalues due to symmetry: hence,  $n = q + 1$ ; that is, there exists only one pair of hyperbolic eigenvalues of  $\mathbf{A}_{\pm\infty}(0)$ . The case when  $\dim E_{\pm\infty}^s(0) > 1$  and when the minimum in (1.19) is greater than one can also be treated but requires a generalization of the theory of § 4 (cf. § 8 and Bridges & Derks (1999)). (With minor modification it is also possible to let  $q < n - 1$ , but this case isn't considered here.)

The set  $\Lambda$  is some simply connected subset of the right-half  $\mathbb{C}$ -plane including at least a subset of the positive real axis near the origin. The hypothesis (1.19) ensures that the spectra associated with  $E_{\pm\infty}^c(0)$  are perturbed in one direction when  $\text{Re}(\lambda) > 0$ . Examples of the hypotheses (1.18) and (1.19) are shown in figures 1*a*, 2*a* and figure 4*a, b* (for the case  $q = 1$  and  $n = 2$ ), figure 6*a, b* (for the case  $q = 2$  and

$n = 3$ ) and figure 7*a, b* (for the case  $q = 2$ ,  $n = 4$  but the minimum in (1.19) for this case is 2).

Section 4 presents a new symplectic characterization of the Evans function. Our main results here are two-fold: the Evans function corresponds precisely to a restriction of the  $\Omega$ -symplectic form and the Pego–Weinstein derivative formula is characterized as a transformation from the spatial (travelling wave)  $\Omega$ -symplectic structure to the temporal  $\omega$ -symplectic structure. These two results are then combined with the geometric characterization of momentum to give an abstract proof of a generalization of (1.1); in particular, the second derivative in (1.1) takes the form

$$\operatorname{sgn} D''(0) = \Pi \operatorname{sgn} \left( \frac{dI}{dc} - \mathcal{B}(c) \right), \quad (1.20)$$

where  $\Pi = \pm 1$  is a geometric sign associated with the shape of the wave, and  $\mathcal{B}(c)$  is associated with the properties of the state at infinity. For example, for a classical solitary wave (i.e. biasymptotic to the trivial state), the term  $\mathcal{B}(c)$  vanishes identically; see lemma 2.1 for a precise definition of this term.

In §§ 5–7 the implications of the theory for three well-known Hamiltonian PDEs are presented. In § 5 we consider a classical semilinear wave equation  $u_{tt} - u_{xx} + V'(u) = 0$ . This example has an interesting multisymplectic formulation and provides an example where both  $dI/dc < 0$  and  $dI/dc > 0$  can imply the existence of an unstable eigenvalue ( $\Pi$  can take either sign in (1.20) depending on the value of  $c$ ). For the gKdV equation, considered in § 6,  $q = 1$  and  $n = 2$  and for the Boussinesq system in § 7,  $q = 2$  and  $n = 3$ . The theory sheds new light on these two PDEs: both examples have natural affine symmetry groups which capture important features of the solitary-wave solutions and, in the case of the gKdV equation, affect the linear stability. And, for the gKdV equation and the Boussinesq system, the theory shows that an instability criterion based on  $dI/dc$  follows directly from abstract properties of these PDEs. In § 8 we indicate how the theory applies to other Hamiltonian PDEs, particularly where the linearization about a solitary wave or front has multi-dimensional stable or unstable manifolds (i.e. the minimum in (1.19) is greater than one).

## 2. Geometric characterization of solitary waves and fronts

In this section, the implications of the abstract formulation of Hamiltonian evolution equations on the real line,

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

for encoding geometric information about solitary-wave states, are studied. The basic hypotheses are that  $\mathbf{M}$  and  $\mathbf{K}$  are constant skew-symmetric operators on  $\mathbb{R}^{2n}$ ,  $S(Z)$  is a smooth function (at least twice continuously differentiable) and  $\mathbf{J}_c = \mathbf{K} - c\mathbf{M}$  with  $\det \mathbf{J}_c \neq 0$  when  $c \in \mathcal{C}$ , where  $\mathcal{C}$  is an open subset of  $\mathbb{R}$ . 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.

The system (2.1) is taken to be equivariant with respect to a  $q$ -parameter abelian Lie group,  $G$ . The action on  $\mathbb{R}^{2n}$  associated with this group is denoted by  $\mathcal{G}_\theta$ , with  $\theta = (\theta_1, \dots, \theta_q)$ . The group is assumed to be either compact (a subgroup of the orthogonal group on  $\mathbb{R}^{2n}$ ) or a subgroup of affine translations of the phase space.

Associated with (2.1) are the three differential two-forms, defined by

$$\left. \begin{aligned} \omega(\xi_1, \xi_2) &= \langle \mathbf{M}\xi_1, \xi_2 \rangle \\ \kappa(\xi_1, \xi_2) &= \langle \mathbf{K}\xi_1, \xi_2 \rangle \\ \Omega(\xi_1, \xi_2) &= \langle \mathbf{J}_c\xi_1, \xi_2 \rangle \end{aligned} \right\} \text{ for any } \xi_1, \xi_2 \in \mathbb{R}^{2n}. \quad (2.2)$$

The two-forms  $\omega$ ,  $\kappa$  and  $\Omega$  are constant and therefore closed. By hypothesis,  $\det \mathbf{J}_c \neq 0$ , when  $c \in \mathcal{C}$ , and therefore  $\Omega$  is a symplectic form and  $(\Omega, \mathbb{R}^{2n})$  is a symplectic manifold. It is the natural symplectic structure associated with travelling-wave states of (2.1).

There are two other symplectic structures associated with (2.1). Restriction of  $\omega$  to the complement of the kernel of  $\mathbf{M}$  in  $\mathbb{R}^{2n}$  defines a second symplectic structure  $(\omega, \mathbb{R}^{2m})$ ,  $m \leq n$ . The integral over  $x$  of  $\omega$  is associated with the classical Hamiltonian structure of (2.1) in the  $t$ -direction. Restriction of  $\kappa$  to the complement of the kernel of  $\mathbf{K}$  defines a third symplectic structure  $(\kappa, \mathbb{R}^{2k})$ ,  $k \leq n$ , which is the natural symplectic structure for time-independent states of (2.1).

The existence of a solitary-wave state or front will be assumed, and we will deduce its abstract properties needed for the proof of instability in §4. The geometric characterization of solitary-wave states will consist of two parts. First, the structure of the solitary wave for  $|x|$  near infinity is characterized: the asymptotic states. Then we characterize the ‘shape’ of the solitary wave: the main part of the wave which connects the asymptotic states at plus and minus infinity.

The simplest class of states at infinity is the uniform (constant, homogeneous) states, obtained by setting  $Z_t = Z_x = 0$  in (2.1). Such states are critical points of  $S(Z)$  in  $\mathbb{R}^{2n}$ . When the system is equivariant, there is a larger class of states at infinity which we will call *relative uniform states*, because they have a natural characterization as spatial relative equilibria (cf. Appendix A). To be explicit, a relative uniform state is a time-independent solution of (2.1) of the form

$$Z(x, t) = \mathcal{G}_{\theta(x)}Z_0, \quad \text{with } \theta(x) = (a_1x + \theta_1^0, \dots, a_qx + \theta_q^0), \quad Z_0 \in \mathbb{R}^{2n}, \quad (2.3)$$

where  $\mathcal{G}_\theta$  is the action of  $G$  and  $a_1, \dots, a_q$  are real parameters. This class of states at infinity arises naturally in examples (cf. §§5–7). The  $x$ -independent uniform states are included in this representation by taking  $\mathcal{G}_\theta$  to be the identity.

The nature of these relative uniform states can be seen as follows. Time-independent states of (2.1) satisfy

$$\mathbf{K}Z_x = \nabla S(Z), \quad Z \in \mathbb{R}^{2n}, \quad x \in \mathbb{R}. \quad (2.4)$$

Let  $\mathcal{G}_{\theta(x)}Z_0$  be a relative uniform state as described in (2.3). The parameters  $a_j$ ,  $j = 1, \dots, q$  and vector  $Z_0$  are determined by a constrained variational problem: substituting (2.3) into (2.4) and using the equivariance of (2.4) results in

$$\nabla S(Z_0) = \sum_{j=1}^q a_j \nabla P_j(Z_0), \quad \text{where } \nabla P_j(Z_0) = \mathbf{K} \frac{d}{d\theta_j} \mathcal{G}_\theta|_{\theta=0}(Z_0) \text{ and } P_j(Z_0) = \mathcal{P}_j. \quad (2.5)$$

In other words,  $Z_0$  can be characterized as a critical point of  $S$  restricted to level sets of the functionals  $P_j$  with Lagrange multipliers  $a_j \in \mathbb{R}$  and the Euler–Lagrange equation (2.5). More details of this characterization and its implications for spatially hyperbolic eigenvalues in the linearization of (2.4) about a relative uniform state are



given in Appendix A. The orbit of a relative uniform state is in an invariant manifold in the phase space. Let  $Z_0$  be a non-degenerate solution of (2.5) and

$$\mathcal{M}_\infty = \text{closure}\{\mathcal{G}_\theta Z_0 : \mathcal{G}_\theta \in G\} \subset \mathbb{R}^{2n} = \begin{cases} \mathbb{T}^q, & \text{if } G \text{ is compact,} \\ \mathbb{R}^q, & \text{if } G \text{ consists of affine translations,} \end{cases}$$

where  $\mathbb{T}^q$  is a  $q$ -dimensional torus. The second equality follows from the hypothesis that  $G$  is abelian and either compact or consists of affine translations. The shape of the solitary wave or front is biasymptotic to points in this invariant manifold. When  $\mathcal{M}_\infty$  is non-trivial, the shape is, in general, asymptotic to distinct points on this manifold at plus and minus infinity.

To characterize the travelling part of the solitary wave, let  $\mathcal{T}_\tau$  be an action of the  $x$ -translation group on vector functions on  $\mathbb{R}^{2n}$ , i.e.

$$\mathcal{T}_\tau Z(x) = Z(x - \tau) \quad \text{for any } \tau \in \mathbb{R}, \quad Z(x) \in \mathbb{R}^{2n}. \quad (2.6)$$

A solitary wave, travelling at speed  $c$ , which is asymptotic to  $\mathcal{M}_\infty$  as  $x \rightarrow \pm\infty$  is characterized as follows:

$$Z(x, t) = \mathcal{G}_{\theta(x)} \mathcal{T}_{\tau(t)} \tilde{Z}(x, c), \quad \text{where } \tau = ct + \tau_0, \quad (2.7)$$

and

$$\|\tilde{Z}(x, c) - Z_0^\pm\| \rightarrow 0 \quad \text{as } x \rightarrow \pm\infty, \quad (2.8)$$

where  $Z_0^\pm$  are two points in  $\mathcal{M}_\infty$ , and  $\|\cdot\|$  is a norm on  $\mathbb{R}^{2n}$ . In general,  $\tilde{Z}$  depends on  $a_j$ ,  $j = 1, \dots, q$  and also  $Z_0^\pm$ , but since they are of secondary importance in the stability theory in §§ 3, 4, their explicit dependence is suppressed. Since the equation for the state at infinity (2.5) is independent of  $c$ , we can choose  $Z_0^-$  to be independent of  $c$  (by varying the constants  $\theta_j^0$  in the definition of  $\theta(x)$ ). This fixes  $Z_0^+$  and, since  $Z_0^+ \in \mathcal{M}_\infty$ , it has to be of the form  $\mathcal{G}_{\gamma(c)} Z_0^-$ , and so  $Z_0^+(c)$  is, in general,  $c$ -dependent.

An alternative to the form (2.7) is  $Z(x, t) = \mathcal{T}_{\tau(t)} \mathcal{G}_{\theta(x)} \hat{Z}(x, c)$ . In this case, the relative uniform state at infinity is also moving at speed  $c$ . Although this case will not be considered here, the present analysis goes through for this case as well with minor changes.

A solitary wave biasymptotic to  $\mathcal{M}_\infty$  is characterized as follows:

$$Z(x, t) = \mathcal{G}_{\theta(x)} \mathcal{T}_{\tau(t)} [Z_0^- + \hat{Z}(x, c)] = \mathcal{G}_{\theta(x)} [Z_0^- + \mathcal{T}_{\tau(t)} \hat{Z}(x, c)], \quad (2.9 a)$$

with

$$\|\hat{Z}(x, c)\| \rightarrow 0 \text{ as } x \rightarrow -\infty \quad \text{and} \quad \|\hat{Z}(x, c) + Z_0^- - Z_0^+(c)\| \rightarrow 0 \text{ as } x \rightarrow +\infty. \quad (2.9 b)$$

We will call  $Z_0^\pm$  the *asymptotic wave states* and  $\hat{Z}$  the *shape of the wave*. A traditional solitary wave asymptotic to the trivial state at infinity is represented by taking  $Z_0^\pm = 0$  and  $\mathcal{G}_{\theta(x)}$  to be the identity.

To obtain the governing equation for the shape  $\hat{Z}(x, c)$ , substitute (2.9 a) into (2.1). Then the equation for  $\hat{Z}$  is

$$-c\mathbf{M}\hat{Z}_x + \mathbf{K}\hat{Z}_x = \nabla S(Z_0^- + \hat{Z}) - \sum_{j=1}^q a_j \nabla P_j(Z_0^- + \hat{Z}), \quad (2.10)$$

where  $q$  is the dimension of the Lie group and  $P_1, \dots, P_q$  are the functionals associated with the  $\kappa$ -symplectic flow of the group  $G$  as introduced in (2.5). The  $x$ -dependent

group action  $\mathcal{G}_{\theta(x)}$  factors out and the only  $x$ -dependence is in  $\hat{Z}$ . The equation can be simplified by renaming the right-hand side,

$$\mathbf{J}_c \hat{Z}_x = \nabla V(\hat{Z}), \quad \hat{Z} \in \mathbb{R}^{2n}, \quad (2.11)$$

where

$$V(\hat{Z}) = S(Z_0^- + \hat{Z}) - S(Z_0^-) - \sum_{j=1}^q a_j [P_j(Z_0^- + \hat{Z}) - P_j(Z_0^-)]. \quad (2.12)$$

The definition of  $V$  carries information about the states at infinity by virtue of  $Z_0^\pm$  and  $a_1, \dots, a_q$ . The function  $V(\hat{Z})$  is defined so that  $\lim_{x \rightarrow \pm\infty} V = 0$ . Since  $\hat{Z}(x, c) \rightarrow 0$  as  $x \rightarrow -\infty$  it is immediate from (2.12) that  $V \rightarrow 0$  as  $x \rightarrow -\infty$ . As  $x \rightarrow +\infty$ ,

$$\lim_{x \rightarrow +\infty} V(\hat{Z}) = V(Z_0^+ - Z_0^-) = S(Z_0^+) - S(Z_0^-) - \sum_{j=1}^q a_j [P_j(Z_0^+) - P_j(Z_0^-)].$$

But  $Z_0^+ = \mathcal{G}_\gamma Z_0^-$  and  $S, P_1, \dots, P_q$  are  $G$ -invariant; hence,  $\lim_{x \rightarrow +\infty} V(\hat{Z}) = 0$ .

There are two interpretations of the shape of the solitary wave as described by (2.11) that will be used in the sequel. First, they can be viewed as heteroclinic orbits, biasymptotic to  $\mathcal{M}_\infty$ , of the spatial Hamiltonian system  $(\mathbb{R}^{2n}, \Omega, V)$  with governing equation (2.11).

The second viewpoint, the classical or temporal viewpoint, is to consider the shape of a solitary-wave state as a critical point of an energy functional restricted to level sets of the momentum. Rewrite (2.11) as

$$\nabla V(\hat{Z}) - \mathbf{K} \hat{Z}_x = -c \mathbf{M} \hat{Z}_x, \quad (2.13)$$

and define the functionals

$$H(\hat{Z}) = \int_{-\infty}^{+\infty} [\frac{1}{2} \kappa(\hat{Z}, \hat{Z}_x) + V(\hat{Z})] dx \quad \text{and} \quad I(\hat{Z}) = \int_{-\infty}^{+\infty} \frac{1}{2} \omega(\hat{Z}, \hat{Z}_x) dx. \quad (2.14)$$

Then, (2.13) is formally recovered by

$$\frac{\delta}{\delta \hat{Z}} H(\hat{Z}) = c \frac{\delta}{\delta \hat{Z}} I(\hat{Z}).$$

That is, the *shape* of the solitary wave corresponds to a critical point of  $H$  restricted to level sets of the functional  $I$  with  $c$  as a Lagrange multiplier, in which case the solitary wave is said to be non-degenerate when  $dI(\hat{Z})/dc \neq 0$ . For the convergence of the integrals in (2.14), it is necessary to subtract off the asymptotic part at minus infinity, and that the asymptotic part at plus infinity is obtained from  $Z_0^-$  via a group action.

However, we will not use this variational principle as a basis for the stability analysis. The most important feature of the energy-momentum characterization, for the present purposes, is that *the momentum density in (2.14) is a one-form on the tangent space to the solitary wave, or heteroclinic manifold satisfying (2.11)*. To see this let  $m(Z) = \frac{1}{2} \mathbf{M} Z$ . Then

$$\chi = \sum_{j=1}^{2n} m_j(Z) dZ_j \quad (2.15)$$

is a one-form. Let  $\mathcal{N}_h(c)$  be the homoclinic manifold:  $\mathcal{N}_h(c) = \{\text{image}(\hat{Z}(x, c)) \subset \mathbb{R}^{2n} : x \in \mathbb{R}\}$ . Then the pullback of the one-form  $\chi$  on  $\mathcal{N}_h(c)$  is precisely the momentum density:

$$\begin{aligned} \int_{\mathcal{N}_h(c)} \chi &= \int_{\mathbb{R}} \sum_{j=1}^{2n} m_j(\hat{Z}) \frac{\partial \hat{Z}_j}{\partial x} dx = \int_{-\infty}^{+\infty} \frac{1}{2} \langle \mathbf{M} \hat{Z}, \hat{Z}_x \rangle dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{2} \omega(\hat{Z}, \hat{Z}_x) dx = I(\hat{Z}), \end{aligned} \tag{2.16}$$

recovering the expression in (2.14).

To summarize, the solitary wave is defined by (2.9 a). It consists of a shape,  $\hat{Z}(x, c)$ , which is biasymptotic to the invariant manifold  $\mathcal{M}_\infty$ . The shape is characterized as a heteroclinic connection in  $\mathbb{R}^{2n}$  satisfying (2.11); it is an orbit of the Hamiltonian system (2.11) with symplectic form  $\Omega$  and it satisfies the asymptotic conditions (2.9 b) with  $Z_0^+ = \mathcal{G}_\gamma Z_0^-$  for some  $\gamma \in G$ . In this setting the momentum associated with the temporal evolution takes the geometric form (2.16). An interesting consequence of this geometric characterization of the momentum is the following result, which provides the abstract connection between  $dI/dc$  and the  $\omega$ -symplectic structure needed for the proof of a generalization of (1.1) in § 4.

**Lemma 2.1.** *Let  $\hat{Z}(x, c)$  be the shape of a solitary-wave state of (2.1), depending smoothly on  $(x, c)$  (at least continuously differentiable) and suppose  $\omega(\hat{Z}_c, \hat{Z}_x) \in L_1(\mathbb{R})$ ; then*

$$\frac{d}{dc} I(\hat{Z}) = \int_{-\infty}^{+\infty} \omega(\hat{Z}_c, \hat{Z}_x) dx + \frac{1}{2} \omega(Z_0^+ - Z_0^-, \partial_c Z_0^+). \tag{2.17}$$

*Proof.* Formal differentiation of the momentum integral over the interval  $(-R, S)$  for  $R, S > 0$ , leads to

$$\begin{aligned} \frac{d}{dc} \int_{-R}^S \frac{1}{2} \omega(\hat{Z}, \hat{Z}_x) dx &= \int_{-R}^S [\frac{1}{2} \omega(\hat{Z}_c, \hat{Z}_x) + \frac{1}{2} \omega(\hat{Z}, \hat{Z}_{xc})] dx \\ &= \int_{-R}^S \omega(\hat{Z}_c, \hat{Z}_x) dx + \frac{1}{2} \omega(\hat{Z}, \hat{Z}_c)|_{-R}^S. \end{aligned}$$

Since  $\omega(\hat{Z}_c, \hat{Z}_x) \in L_1(\mathbb{R})$ , the proof is completed by taking the limit as  $R, S \rightarrow \infty$  and using the fact that  $\hat{Z} \rightarrow Z_0^+(c) - Z_0^-$  as  $x \rightarrow +\infty$ , and that  $Z_0^-$  is independent of  $c$ . ■

There are two useful corollaries of this lemma; the proof of the first is immediate from the definition.

**Corollary 2.2.** *Suppose  $Z_0^+$  is independent of  $c$  or  $Z_0^+ = Z_0^-$ . Then the boundary terms vanish and (2.17) reduces to*

$$\frac{d}{dc} I(\hat{Z}) = \int_{-\infty}^{+\infty} \omega(\hat{Z}_c, \hat{Z}_x) dx. \tag{2.18}$$

**Corollary 2.3.** *Suppose  $G$  consists only of affine translations. Then the boundary terms vanish and (2.17) is reduced to (2.18).*

*Proof.* The action of  $G$  can be written

$$\mathcal{G}_\theta Z = Z + \sum_{j=1}^q \theta_j V_j$$

for some  $V_1, \dots, V_q \in \mathbb{R}^{2n}$ , since  $G$  consists of a  $q$ -dimensional group of affine translations. Therefore

$$Z_0^+(c) = \mathcal{G}_{\gamma(c)} Z_0^- = Z_0^- + \sum_{j=1}^q \gamma_j(c) V_j,$$

and

$$\frac{1}{2} \omega(Z_0^+ - Z_0^-, \partial_c Z_0^+) = \frac{1}{2} \omega \left( \sum_{j=1}^q \gamma_j V_j, \sum_{k=1}^q \gamma_k'(c) V_k \right) = \frac{1}{2} \sum_{j,k=1}^q \gamma_j(c) \gamma_k'(c) \omega(V_j, V_k).$$

But, symplectic Noether theory implies that there are functions  $Q_1, \dots, Q_q$  such that  $\mathbf{M}V_j = \nabla Q_j(Z)$  (see Appendix A). Hence,

$$\omega(V_j, V_k) = \langle \mathbf{M}V_j, V_k \rangle = \langle \nabla Q_j(Z), V_k \rangle = \frac{d}{d\theta_k} Q_j(\mathcal{G}_\theta Z)|_{\theta=0} = 0,$$

with the last equality following from the fact that the group is abelian and hence each  $Q_j(Z)$  is  $G$ -invariant. ■

### 3. Asymptotic properties of the linearization about a wave

In this section we will linearize (2.1) about a wave  $\mathcal{G}_{\theta(x)} T_{\tau(t)} [Z_0^- + \hat{Z}(x, c)]$  and analyse the behaviour of the linearized equation for  $|x| \rightarrow \infty$ . Write a solution  $Z(x, t)$  of (2.1) as

$$Z(x, t) = \mathcal{G}_{\theta(x)} T_{\tau(t)} [Z_0^- + \hat{Z}(x, c) + \hat{U}(x, t)]. \tag{3.1}$$

Then, substitution of (3.1) into (2.1) and linearization about the shape of the wave combined with the state at infinity lead to

$$\mathbf{M}\hat{U}_t + \mathbf{J}_c \hat{U}_x = \mathbf{B}(x, c) \hat{U}, \quad \hat{U}(x, t) \in \mathbb{R}^{2n}, \quad x \in \mathbb{R}, \quad t > 0, \tag{3.2}$$

where

$$\mathbf{B}(x, c) = D^2 V(\hat{Z}(x, c)) = D^2 S(Z_0^- + \hat{Z}(x, c)) - \sum_{j=1}^q a_j D^2 P(Z_0^- + \hat{Z}(x, c)). \tag{3.3}$$

The ansatz  $\hat{U}(x, t) = e^{\lambda t} U(x, \lambda)$  reduces (3.2) to the spectral problem:

$$\mathbf{J}_c U_x = [\mathbf{B}(x, c) - \lambda \mathbf{M}] U, \tag{3.4}$$

or

$$U_x = \mathbf{A}(x, \lambda) U, \quad U \in \mathbb{C}^{2n}, \quad \lambda \in \mathbb{C}, \quad x \in \mathbb{R}, \tag{3.5}$$

where

$$\mathbf{A}(x, \lambda) = \mathbf{J}_c^{-1} [\mathbf{B}(x, c) - \lambda \mathbf{M}]. \tag{3.6}$$

The matrix  $\mathbf{A}(x, \lambda)$  also depends on  $c$ , but this dependence is secondary and is suppressed to simplify notation. In the next section, the spectral problem (3.4)–(3.6)

is studied. In the remainder of this section, the asymptotic ( $|x| \rightarrow +\infty$ ) properties of the system (3.5) are considered.

In the limit as  $x \rightarrow \pm\infty$ , the systems (3.5) at infinity reduce to the linearization about the relative uniform state, i.e.

$$\mathbf{J}_c U_x - \mathbf{B}_0^\pm U = -\lambda \mathbf{M} U \quad \text{or} \quad U_x = \mathbf{A}_{\pm\infty}(\lambda) U, \tag{3.7}$$

where

$$\mathbf{B}_0^\pm = D^2 S(Z_0^\pm) - \sum_{j=1}^q a_j D^2 P_j(Z_0^\pm) \quad \text{and} \quad \mathbf{A}_{\pm\infty}(\lambda) = \mathbf{J}_c^{-1}(\mathbf{B}_0^\pm - \lambda \mathbf{M}). \tag{3.8}$$

Associated with the systems at infinity, (3.7) is the eigenvalue problem for the spatial exponent. Let  $U(x) = e^{\mu x} \tilde{U}$  and define

$$\Delta^\pm(\mu, \lambda) = \det[\mathbf{B}_0^\pm - \mu \mathbf{J}_c - \lambda \mathbf{M}], \quad \mu \in \mathbb{C}, \quad \lambda \in \Lambda. \tag{3.9}$$

The spectrum of matrices  $\mathbf{A}_{\pm\infty}(\lambda)$  is then

$$\Sigma^\pm(\lambda) = \{\mu \in \mathbb{C} : \Delta^\pm(\mu, \lambda) = 0\}, \quad \lambda \in \Lambda \subset \mathbb{C}. \tag{3.10}$$

Although the explicit dependence in the notation is suppressed,  $\Sigma$  and  $\Delta$  both depend on  $c$ . Although  $\mathbf{A}_{-\infty}(\lambda)$  and  $\mathbf{A}_{+\infty}(\lambda)$  are not equal in general, their spectrum is identical since  $Z_0^+$  and  $Z_0^-$  are related by a group element.

**Proposition 3.1.** *Suppose  $Z_0^+ = \mathcal{G}_\gamma Z_0^-$  for some  $\gamma \in G$ . Then  $\Delta^+(\mu, \lambda) = \Delta^-(\mu, \lambda)$ . Moreover, for any  $(\mu, \lambda) \in \mathbb{C}^2$ , the functions  $\Delta^\pm(\mu, \lambda)$  satisfy the multisymplectic reflection property*

$$\Delta^\pm(\mu, \lambda) = \Delta^\pm(-\mu, -\lambda). \tag{3.11}$$

*Proof.* By hypothesis, the group is abelian and either compact or consists of affine translations. When the group consists of affine translations, the invariance of  $S, P_1, \dots, P_q$  implies

$$D^2 S(\mathcal{G}_\theta Z_0^-) - \sum_{j=1}^q a_j D^2 P_j(\mathcal{G}_\theta Z_0^-) = D^2 S(Z_0^-) - \sum_{j=1}^q a_j D^2 P_j(Z_0^-), \quad \forall \theta \in G.$$

Choosing  $\theta = \gamma$ , it follows that  $\mathbf{B}_0^+ = \mathbf{B}_0^-$ .

When the group is compact and abelian, the invariance of  $S, P_1, \dots, P_q$  implies

$$\begin{aligned} D^2 S(\mathcal{G}_\theta Z_0^-) - \sum_{j=1}^q a_j D^2 P_j(\mathcal{G}_\theta Z_0^-) \\ = \mathcal{G}_\theta^{-T} \left[ D^2 S(Z_0^-) - \sum_{j=1}^q a_j D^2 P_j(Z_0^-) \right] \mathcal{G}_\theta^{-1}, \quad \forall \theta \in G. \end{aligned}$$

Hence, with  $\theta = \gamma$ ,  $\mathbf{B}_0^+ = \mathcal{G}_\gamma^{-T} \mathbf{B}_0^- \mathcal{G}_\gamma^{-1}$ . By hypothesis  $\omega$  and  $\kappa$  are also  $G$ -invariant, therefore  $\mathbf{J}_c$  and  $\mathbf{M}$  commute with the action of the group, and so

$$\begin{aligned} \Delta^+(\mu, \lambda) &= \det[\mathbf{B}_0^+ - \mu \mathbf{J}_c - \lambda \mathbf{M}] = \det[\mathcal{G}_\gamma^{-T} \mathbf{B}_0^- \mathcal{G}_\gamma^{-1} - \mu \mathbf{J}_c - \lambda \mathbf{M}] \\ &= \det \mathcal{G}_\gamma^{-T} \det[\mathbf{B}_0^- - \mu \mathbf{J}_c - \lambda \mathbf{M}] \det \mathcal{G}_\gamma^{-1} \\ &= \det[\mathbf{B}_0^- - \mu \mathbf{J}_c - \lambda \mathbf{M}] = \Delta^-(\mu, \lambda). \end{aligned}$$

The multisymplectic reflection property follows from the invariance of the determinant under transposition:

$$\begin{aligned}\Delta^\pm(\mu, \lambda) &\stackrel{\text{def}}{=} \det[\mathbf{B}_0^\pm - \mu\mathbf{J}_c - \lambda\mathbf{M}] \\ &= \det[(\mathbf{B}_0^\pm)^\text{T} - \mu\mathbf{J}_c^\text{T} - \lambda\mathbf{M}^\text{T}] \\ &= \det[(\mathbf{B}_0^\pm) + \mu\mathbf{J}_c + \lambda\mathbf{M}] = \Delta^\pm(-\mu, -\lambda),\end{aligned}$$

since the matrices  $\mathbf{B}_0^\pm$ ,  $\mathbf{J}_c$  and  $\mathbf{M}$  are, respectively, symmetric, skew-symmetric and skew-symmetric. ■

Since  $\Delta^+ = \Delta^-$  and  $\Sigma^+ = \Sigma^-$ , we will henceforth drop the  $\pm$  superscript. Another property of  $\Delta(\mu, \lambda)$ , which follows since the matrices  $\mathbf{B}_0^\pm$ ,  $\mathbf{J}_c$  and  $\mathbf{M}$  are real-valued, is

$$\overline{\Delta(\mu, \lambda)} = \Delta(\bar{\mu}, \bar{\lambda}). \quad (3.12)$$

A consequence of this property combined with proposition 3.1 is that, if  $\mu \in i\mathbb{R}$  and  $\lambda \in i\mathbb{R}$ , then  $\Delta(\mu, \lambda) \in \mathbb{R}$ . In other words, when  $\mu, \lambda$  are either purely real or purely imaginary,  $\Delta(\mu, \lambda)$  is purely real.

When  $\lambda = 0$ , (3.7) is a linear autonomous Hamiltonian system with symplectic form  $\Omega$ . The reflection in proposition 3.1 then reduces to  $\Delta(\mu, 0) = \Delta(-\mu, 0)$ , which is the symplectic reflection property associated with the  $\Omega$ -symplectic structure. A consequence of the  $G$ -symmetry is that  $\Delta(\mu, 0)$  has zero as a root with multiplicity  $2q$ .

**Proposition 3.2.** *Suppose that all the elements in the Lie algebra of  $G$  act non-trivially on  $Z_0^\pm$ . Then the function  $\Delta(\mu, 0)$  is an even function of  $\mu$  and satisfies*

$$\Delta(\mu, 0) = \mu^{2q}(\det[\mathbf{J}_c]\mu^{2(n-q)} + \delta_1(\mu)), \quad (3.13)$$

where  $\delta_1(\mu)$  is an even polynomial of degree less than  $2(n - q)$ . In particular, zero is a root of algebraic multiplicity  $2q$ .

*Proof.* The  $G$ -equivariance of  $S$  and the functionals  $P_1, \dots, P_q$  imply that

$$\nabla S(\mathcal{G}_\theta Z_0^\pm) = \sum_{j=1}^q a_j \nabla P_j(\mathcal{G}_\theta Z_0^\pm).$$

Differentiating with respect to each group parameter, setting  $\theta = 0$  and writing  $\partial_{\theta_j} \mathcal{G}_\theta|_{\theta=0} = \mathfrak{g}_j(\cdot)$  leads to

$$[D^2 S(Z_0^\pm) - \sum_{j=1}^q a_j D^2 P_j(Z_0^\pm)] \mathfrak{g}_j(Z_0^\pm) = 0, \quad j = 1, \dots, q,$$

that is,  $\mathbf{B}_0^\pm \mathfrak{g}_j(Z_0^\pm) = 0$  for  $j = 1, \dots, q$ , and by hypothesis each  $\mathfrak{g}_j(Z_0^\pm)$  is non-trivial. The fact that each zero has multiplicity two follows from Jordan chain theory and the symmetry of  $\mathbf{B}_0^\pm$ , which implies that  $\mathbf{B}_0^\pm U = \mathbf{J}_c \mathfrak{g}_j(Z_0^\pm)$  is solvable. ■

With the hypotheses  $n = q + 1$  and  $\mathfrak{g}_j(Z_0^\pm) \neq 0$ , for  $j = 1, \dots, q$ , the spectrum of  $\mathbf{A}_{\pm\infty}(0)$  reduces to

$$\Sigma(0) = \{0\} \cup \{\beta(c, 0), -\beta(c, 0)\}, \quad \forall c \in \mathcal{C}, \quad (3.14)$$

where  $\beta(c, 0)$  is a positive real number and it is assumed to depend differentiably on  $c$  for  $c \in \mathcal{C}$ . A geometric sufficient condition for the existence of such a pair of hyperbolic real eigenvalues in the linearization about a *relative* uniform state is given in Appendix A. The hypothesis that there is only one pair of eigenvalues with non-zero real part is not essential but is taken for definiteness. The case of higher-dimensional stable or unstable manifolds is considered in Bridges & Derks (1999). If  $q < n - 1$  and the spectrum  $\Sigma(0)$  is of the form as described above, the analysis goes through as well. (Although this case is of less interest, as the higher multiplicity without symmetry is not generic, and it would then be natural to add additional parameters.)

When  $\lambda$  is non-zero the set  $\Sigma(0)$  will perturb. When  $\lambda \in \Lambda$  is non-zero, with positive real part, we will assume that  $\Sigma(\lambda)$  perturbs to

$$\Sigma(\lambda) = \{-\beta(c, \lambda)\} \cup \Sigma_+ \tag{3.15}$$

(cf. equations (1.17)–(1.19)). The exponent  $-\beta(c, \lambda)$  is a simple eigenvalue of  $\mathbf{A}_{\pm\infty}(\lambda)$  with strictly negative real part and  $\Sigma_+$  consists of  $2n - 1$  eigenvalues with non-negative real part (see figures 4 and 6 for examples). (Equivalently,  $\Sigma(\lambda)$  could be taken as  $\Sigma_- \cup \{\beta(c, \lambda)\}$  with  $\Sigma_-$  consisting of non-positive eigenvalues.)

The multisymplectic reflection property ensures the existence of an eigenvalue dual to  $-\beta(c, \lambda)$  and a natural normalization for the associated eigenvectors.

**Proposition 3.3.** *Let  $\zeta_+(c, \lambda)$  be the eigenvector associated with the eigenvalue  $-\beta(c, \lambda)$  satisfying  $\mathbf{A}_{+\infty}(\lambda)\zeta_+ = -\beta\zeta_+$  or*

$$[\mathbf{B}_0^+ - \lambda\mathbf{M} + \beta\mathbf{J}_c]\zeta_+(c, \lambda) = 0. \tag{3.16}$$

*There exists a vector  $\zeta_-(c, \lambda)$  satisfying the dual eigenvalue problem*

$$[\mathbf{B}_0^- + \lambda\mathbf{M} - \beta\mathbf{J}_c]\zeta_-(c, \lambda) = 0. \tag{3.17}$$

*The natural normalization for  $\zeta_+(c, \lambda)$  and  $\zeta_-(c, \lambda)$  is*

$$\Omega(\Gamma_\gamma\zeta_-, \zeta_+) = 1, \quad \text{where } \Gamma_\gamma = \begin{cases} \mathcal{G}_\gamma, & \text{if } G \text{ is compact,} \\ \mathbf{I}, & \text{if } G \text{ is affine,} \end{cases} \tag{3.18}$$

*and  $\gamma$  is defined by the rotation  $Z_0^+ = \mathcal{G}_\gamma Z_0^-$ .*

*Proof.* The fact that (3.17) has a solution follows from proposition 3.1. The natural normalization for the eigenvector  $\zeta_+(c, \lambda)$  and its adjoint eigenvector, denoted  $\Upsilon_+$ , is

$$\langle \Upsilon_+, \zeta_+ \rangle_{\mathbb{C}} = 1, \tag{3.19}$$

where  $\Upsilon_+$  is the usual adjoint eigenvector satisfying  $[\mathbf{A}_{+\infty}(\lambda)^* + \bar{\beta}\mathbf{I}]\Upsilon_+ = 0$ . But

$$\mathbf{A}_{+\infty}(\lambda)^* = [\mathbf{B}_0^+ - \lambda\mathbf{M}]^*(\mathbf{J}_c^{-1})^* = -[\mathbf{B}_0^+ + \bar{\lambda}\mathbf{M}]\mathbf{J}_c^{-1}$$

and so  $\Upsilon_+$  satisfies

$$-[\mathbf{B}_0^+ + \bar{\lambda}\mathbf{M} - \bar{\beta}\mathbf{J}_c]\mathbf{J}_c^{-1}\Upsilon_+ = 0.$$

If the group  $G$  consists of affine translations, then  $\mathbf{B}_0^+ = \mathbf{B}_0^-$ , in which case  $\mathbf{J}_c^{-1}\Upsilon_+ = \bar{\zeta}_-$ . If the group is compact, then  $\mathbf{B}_0^+ = \mathcal{G}_\gamma^{-T}\mathbf{B}_0^-\mathcal{G}_\gamma^{-1}$  and

$$\mathcal{G}_\gamma^{-T}[-\mathbf{B}_0^- - \bar{\lambda}\mathbf{M} + \bar{\beta}\mathbf{J}_c]\mathcal{G}_\gamma^{-1}\mathbf{J}_c^{-1}\Upsilon_+ = 0.$$

Therefore  $\mathcal{G}_\gamma^{-1}\mathbf{J}_c^{-1}\Upsilon_+ = \bar{\zeta}_-$  or with  $\Gamma_\gamma$  as defined in (3.17),  $\mathbf{J}_c^{-1}\Upsilon_+ = \Gamma_\gamma\bar{\zeta}_-$ ; hence

$$1 = \langle \Upsilon_+, \zeta_+ \rangle_{\mathbb{C}} = \langle \mathbf{J}_c\Gamma_\gamma\zeta_-, \zeta_+ \rangle_{\mathbb{R}} = \Omega(\Gamma_\gamma\zeta_-, \zeta_+).$$

■

When  $\lambda = 0$ , the derivative (with respect to  $x$ ) of the shape of the solitary wave or front is a solution of the spectral problem (3.5): differentiating (2.10) with respect to  $x$ ,

$$\mathbf{J}_c(\hat{Z}_x)_x = D^2S(Z_0^- + \hat{Z})\hat{Z}_x - \sum_{j=1}^q a_j D^2P_j(Z_0^- + \hat{Z})\hat{Z}_x = \mathbf{B}(x, c)\hat{Z}_x.$$

With one pair of hyperbolic real eigenvalues in  $\Sigma(0)$  when  $\lambda = 0$ , it is clear that the derivative of the shape of the solitary wave or front,  $\hat{Z}_x$ , is asymptotically tangent to the eigenspace associated with the eigenvalues  $\mp\beta(c, 0)$  of  $\mathbf{A}_{\pm\infty}(0)$ . Therefore,  $\hat{Z}_x$  satisfies the natural asymptotic conditions

$$\lim_{x \rightarrow \pm\infty} e^{\pm\beta(c,0)x} \hat{Z}_x(x, c) = C_{\pm} \zeta_{\pm}(c, 0), \tag{3.20}$$

where  $C_{\pm}$  are real constants defined by

$$C_- = - \lim_{x \rightarrow -\infty} e^{-\beta(c,0)x} \Omega(\Gamma_{\gamma}^T \zeta_+, \hat{Z}_x(x, c)),$$

$$C_+ = \lim_{x \rightarrow +\infty} e^{+\beta(c,0)x} \Omega(\Gamma_{\gamma} \zeta_-, \hat{Z}_x(x, c)).$$

Combining these two asymptotic conditions in (3.20),

$$C_- C_+ = \lim_{x \rightarrow +\infty} e^{2\beta(c,0)x} \Omega(\Gamma_{\gamma} \hat{Z}_x(-x, c), \hat{Z}_x(x, c)). \tag{3.21 a}$$

We therefore define the following geometric sign of the heteroclinic or homoclinic connection:

$$\Pi = \text{sgn}(C_- C_+). \tag{3.21 b}$$

As far as we are aware, this geometric sign of a heteroclinic connection has not previously been recognized. It is important here because it arises naturally in the analysis of the linear stability equation in §4. A useful simplification of the formula for  $\Pi$  occurs when the shape corresponds to a reversible homoclinic or heteroclinic orbit; that is, there exists an involution  $\mathbf{S}$  such that  $\mathbf{S}\hat{Z}(x, c) = \hat{Z}(-x, c)$  (cf. the examples in §§5–7).

The exponential decay of the shape  $\hat{Z}$  to the states at infinity  $Z_0^{\pm}$  implies that

$$\int_{-\infty}^{+\infty} \|\mathbf{B}(x, c) - \tilde{\mathbf{B}}_0(x)\| dx < +\infty, \quad \text{where } \tilde{\mathbf{B}}_0(x) = \begin{cases} \mathbf{B}_0^-, & x < 0, \\ \mathbf{B}_0^+, & x \geq 0. \end{cases} \tag{3.22}$$

Note that  $\mathbf{A}(x, \lambda) - \mathbf{A}_{\pm\infty}(\lambda) = \mathbf{J}_c^{-1}[\mathbf{B}(x, c) - \mathbf{B}_0^{\pm}]$ . Therefore, a consequence of (3.22) and the decomposition of  $\mathbf{A}(x, \lambda)$  is

$$\int_{-\infty}^{+\infty} \|\mathbf{R}(x, \lambda)\| dx < +\infty, \quad \text{independent of } \lambda, \tag{3.23 a}$$

where

$$\mathbf{R}(x, \lambda) = \begin{cases} \mathbf{A}(x, \lambda) - \mathbf{A}_{-\infty}(\lambda), & x < 0, \\ \mathbf{A}(x, \lambda) - \mathbf{A}_{+\infty}(\lambda), & x \geq 0. \end{cases} \tag{3.23 b}$$

This implies that the behaviour of the solutions of (3.7) is relevant for the asymptotic behaviour of the solutions of (3.5).



#### 4. Symplectifying the Evans function and its derivatives

In this section the Evans function theory is applied to (3.5), i.e.

$$U_x = \mathbf{A}(x, \lambda)U, \quad U \in \mathbb{C}^{2n}, \quad x \in \mathbb{R}, \quad \lambda \in \Lambda \subset \mathbb{C}, \tag{4.1}$$

with particular attention to the implications of the multisymplectic decomposition of  $\mathbf{A}(x, \lambda)$  in (3.6), under the hypotheses that  $\lim_{x \rightarrow \pm\infty} \mathbf{A}(x, \lambda) = \mathbf{A}_{\pm\infty}(\lambda)$  with the spectrum of  $\mathbf{A}_{\pm\infty}(\lambda)$  taking the form (3.15), for  $\lambda \in \Lambda$ . In this section, the set  $\Lambda$  is some open simply connected subset of the right-half complex  $\lambda$ -plane which includes all or at least the part of the positive real axis near the origin and is such that (3.15) holds for all  $\lambda \in \Lambda$ .

Our main results are that the Evans function can be characterized as a restriction of the *spatial* symplectic form  $\Omega$  and the Pego–Weinstein derivative formula can be characterized as a transfer of information from the *spatial* symplectic structure to the *temporal* symplectic structure. These two results, when combined with the  $\omega$ -symplectic characterization of momentum in lemma 2.1, will lead to an abstract proof of a generalization of (1.1).

With the hypothesis (3.15) on the spectrum  $\Sigma(c, \lambda)$  and the estimate (3.23), it follows from the Levinson theory (cf. Coppel 1965; Eastham 1989; Pego & Weinstein 1992, proposition 1.2) that there exists a solution of (4.1) analytic in  $\lambda$  for  $\lambda \in \Lambda$  and satisfying

$$\lim_{x \rightarrow +\infty} e^{\beta(c, \lambda)x} \xi(x, \lambda) = \zeta_+(c, \lambda), \tag{4.2}$$

with  $\zeta_+$  defined in (3.16). To simplify notation the explicit dependence of  $\xi(x, \lambda)$  on  $c$  is suppressed. Associated with the  $2n - 1$  non-negative eigenvalues of  $\mathbf{A}_{-\infty}(\lambda)$ , there exist linearly independent solutions  $\xi_2(x, \lambda), \dots, \xi_{2n}(x, \lambda)$ , analytic in  $\lambda$  for  $\lambda \in \Lambda$ , which grow at most algebraically as  $x \rightarrow -\infty$ .

Let

$$W(x, \lambda) = \xi_2(x, \lambda) \wedge \dots \wedge \xi_{2n}(x, \lambda) \in \wedge^{2n-1}(\mathbb{C}^{2n}); \tag{4.3}$$

then the Evans function is defined to be

$$\tilde{D}(\lambda) = \exp\left(-\int_0^x \text{Tr}(\mathbf{A}(s, \lambda)) ds\right) W(x, \lambda) \wedge \xi(x, \lambda), \quad \lambda \in \Lambda,$$

(cf. Alexander *et al.* 1990, p. 184). With the special form for  $\mathbf{A}(x, \lambda)$ , the  $\text{Tr}(\mathbf{A}(x, \lambda))$  can be simplified. Using (3.6),

$$\text{Tr}(\mathbf{A}(x, \lambda)) = \text{Tr}(\mathbf{J}_c^{-1}(\mathbf{B}(x, c) - \lambda \mathbf{M})) = \text{Tr}(\mathbf{J}_c^{-1} \mathbf{B}(x, c)) - \lambda \text{Tr}(\mathbf{J}_c^{-1} \mathbf{M}).$$

But  $\text{Tr}(\mathbf{J}_c^{-1} \mathbf{B}(x, c))$  vanishes identically for all  $x$  since  $\mathbf{J}_c$  is skew-symmetric and  $\mathbf{B}(x, c)$  is symmetric. Hence

$$\tilde{D}(\lambda) = e^{\lambda \tau x} W(x, \lambda) \wedge \xi(x, \lambda), \quad \lambda \in \Lambda, \quad \text{with } \tau = \text{Tr}(\mathbf{J}_c^{-1} \mathbf{M}). \tag{4.4}$$

The  $x$ -independence of  $\tilde{D}(\lambda)$  follows from the fact that  $W \wedge \xi$  is the Wronskian of the  $2n$  solutions of (4.1) times a volume form and therefore satisfies the Abel–Liouville theorem:

$$\frac{d}{dx} W(x, \lambda) \wedge \xi(x, \lambda) = \text{Tr}(\mathbf{A}(x, \lambda)) W(x, \lambda) \wedge \xi(x, \lambda). \tag{4.5}$$

Multiplication of  $W \wedge \xi$  by the exponential term, combined with the Abel–Liouville theorem, ensures that  $\tilde{D}(\lambda)$  is independent of  $x$ . The above hypotheses also ensure that  $\tilde{D}(\lambda)$  is an analytic function of  $\lambda$  for  $\lambda \in \Lambda$ .

If  $\tilde{D}(\lambda_0) = 0$  for some  $\lambda_0 \in \Lambda$  with  $\operatorname{Re}(\lambda_0) > 0$  then the function  $\xi(x, \lambda)$  is an eigenfunction of the spectral problem (3.5), and  $\lambda_0$  is an unstable eigenvalue. Such an eigenfunction decays exponentially as  $x \rightarrow +\infty$  and grows at most algebraically as  $x \rightarrow -\infty$ , with the algebraic growth associated with the zero eigenvalues in  $\Sigma_+$  in (3.15).

In coordinates,  $\tilde{D}(\lambda)$  is a complex function times a volume form. We find that this complex function has a natural characterization in terms of the symplectic form  $\Omega$ .

**Lemma 4.1 (Symplectic Evans function).** *Let  $\xi(x, \lambda)$ ,  $\lambda \in \Lambda$ , be a solution of (4.1) satisfying (4.2) and suppose  $\eta(x, \lambda)$  satisfies*

$$\eta_x = \mathbf{A}(x, -\lambda)\eta \quad \text{and} \quad \lim_{x \rightarrow -\infty} e^{-\beta(c, \lambda)x} \eta(x, \lambda) = \zeta_-(c, \lambda), \tag{4.6}$$

with  $\zeta_-(c, \lambda)$  defined in (3.17). Then with a suitable scaling of  $\xi_2(x, \lambda), \dots, \xi_{2n}(x, \lambda)$ ,

$$\tilde{D}(\lambda) = D(\lambda)\mathcal{V}, \quad \text{with } D(\lambda) = \Omega(\eta, \xi). \tag{4.7}$$

In this expression,  $\mathcal{V}$  is a volume form on  $\mathbb{R}^{2n}$  and  $\Omega$  is the symplectic form defined in (2.2).

*Proof.* The characterization (4.6) is a consequence of Hodge duality, the decomposition of  $\mathbf{A}(x, \lambda)$  and the multisymplectic reflection property. In Appendix B, it is proved that  $\star W(x, \lambda)$ , with  $W$  in (4.3) and  $\star$  the Hodge star operator, satisfies the differential equation

$$(\star W)_x = [-\bar{\lambda}\tau - \mathbf{A}(x, \bar{\lambda})^T](\star W), \quad \star W \in \wedge^1(\mathbb{C}^{2n}),$$

where the overbar denotes complex conjugation. Define  $\mathcal{W}(x, \lambda) = e^{\lambda\tau x}W(x, \lambda)$ ; then

$$\star \mathcal{W} = e^{\bar{\lambda}\tau x}(\star W(x, \lambda)) \quad \text{and} \quad (\star \mathcal{W})_x = -\mathbf{A}(x, \bar{\lambda})^T(\star \mathcal{W}). \tag{4.8}$$

Using the above expression and Hodge duality (see equation (B 5) in Appendix B),

$$\tilde{D}(\lambda) = e^{\lambda\tau x}W(x, \lambda) \wedge \xi(x, \lambda) = \mathcal{W}(x, \lambda) \wedge \xi(x, \lambda) = \langle \star \mathcal{W}, \xi \rangle_{\mathbb{C}} \mathcal{V} = \langle \overline{\star \mathcal{W}}, \xi \rangle_{\mathbb{R}} \mathcal{V}, \tag{4.9}$$

where  $\mathcal{V}$  is a volume form. Let  $\eta = \mathbf{J}_c^{-1}(\overline{\star \mathcal{W}})$ . Then  $\eta$  satisfies

$$\mathbf{J}_c \eta_x = -\mathbf{A}(x, \lambda)^T(\mathbf{J}_c \eta) = [\mathbf{B}(x, c) + \lambda \mathbf{M}]\eta,$$

or  $\eta_x = \mathbf{A}(x, -\lambda)\eta$ . Substitution into (4.9) results in

$$\tilde{D}(\lambda) = \langle \overline{\star \mathcal{W}}, \xi \rangle_{\mathbb{R}} \mathcal{V} = \langle \mathbf{J}_c \eta, \xi \rangle_{\mathbb{R}} \mathcal{V} = \Omega(\eta, \xi)\mathcal{V}.$$

It remains to establish the asymptotics of  $\eta(x, \lambda)$  as  $x \rightarrow -\infty$  in (4.6). As  $x \rightarrow -\infty$ , the solutions  $\xi_2(x, \lambda), \dots, \xi_{2n}(x, \lambda)$  are asymptotic to the eigenspace associated with the eigenvalues of  $\mathbf{A}_{-\infty}(\lambda)$  with non-negative real part. Now  $\operatorname{Tr}(\mathbf{A}_{-\infty}(\lambda)) = -\lambda\tau$ ; hence, the sum of the eigenvalues of  $\mathbf{A}_{-\infty}(\lambda)$  with non-negative real part is  $-\lambda\tau + \beta(c, \lambda)$ . Therefore, as  $x \rightarrow -\infty$  the function  $W(x, \lambda)$  in (4.3) has the following asymptotics:

$$\lim_{x \rightarrow -\infty} e^{(-\lambda\tau + \beta(c, \lambda))x} \xi_2(x, \lambda) \wedge \dots \wedge \xi_{2n}(x, \lambda) = W_\infty \in \wedge^{2n-1}(\mathbb{C}^{2n}),$$

where  $W_\infty$  is some constant vector in  $\wedge^{2n-1}(\mathbb{C}^{2n})$ . Since the trace of  $\mathbf{A}(x, \lambda)$  is independent of  $x$ , it follows, from the definition of  $\mathcal{W}$  and the above expression, that

$$\lim_{x \rightarrow -\infty} e^{-\beta(c, \lambda)x} \mathcal{W}(x, \lambda) = W_\infty.$$

A similar argument for  $\star W$  shows that

$$\lim_{x \rightarrow -\infty} e^{-\overline{\beta(c, \lambda)x}} (\star \mathcal{W}(x, \lambda)) = \star W_\infty \in \wedge^1(\mathbb{C}^{2n}).$$

However, by the Levinson theory, since  $\beta(c, \lambda)$  is a simple eigenvalue of  $-\mathbf{A}_{-\infty}(\lambda)^T = \mathbf{J}_c \mathbf{A}_{-\infty}(-\lambda) \mathbf{J}_c^{-1}$ ,  $\mathbf{J}_c^{-1} \star W_\infty$  must be a multiple of the eigenvector of  $\mathbf{A}_{-\infty}(-\lambda)$  associated with  $\beta(c, \lambda)$  which is  $\zeta_-$ . Since the equation is linear, the functions  $\xi_2(x, \lambda), \dots, \xi_{2n}(x, \lambda)$  can be scaled so that the constant multiple can be taken to be unity, resulting in the asymptotic condition (4.6). ■

In Pego & Weinstein (1992, definition 1.8), the Evans function is defined by

$$D(\lambda) = \Upsilon(x, \lambda) \cdot \xi(x, \lambda), \tag{4.10}$$

where  $\cdot$  is a scalar product,  $\xi(x, \lambda)$  satisfies the asymptotic estimate (4.2) and  $\Upsilon(x, \lambda)$  satisfies the adjoint system,  $\Upsilon_x = -\mathbf{A}(x, \lambda)^T \Upsilon$ , and an asymptotic estimate like that in (4.6). The main result of lemma 4.1 is the characterization of  $D(\lambda)$  as a restricted symplectic form. In the present notation

$$\Upsilon(x, \lambda) = \overline{\star \mathcal{W}(x, \lambda)} = \mathbf{J}_c \eta(x, \lambda). \tag{4.11}$$

An important consequence of this characterization is a multisymplectification of the Pego–Weinstein derivative formula.

**Lemma 4.2.** *Suppose  $D(\lambda_0) = 0$ ,  $\lambda_0 \in \Lambda$  and that  $\omega(\eta, \xi)|_{\lambda=\lambda_0} \in L_1(\mathbb{R})$ . Then*

$$D'(\lambda_0) \stackrel{\text{def}}{=} \frac{d}{d\lambda} \Omega(\eta, \xi)|_{\lambda=\lambda_0} = \int_{-\infty}^{+\infty} \omega(\eta, \xi)|_{\lambda=\lambda_0} dx.$$

*Suppose in addition that  $D'(\lambda_0) = 0$  and  $\partial_\lambda \omega(\eta, \xi)|_{\lambda=\lambda_0} \in L_1(\mathbb{R})$ . Then*

$$D''(\lambda_0) = \int_{-\infty}^{+\infty} \frac{d}{d\lambda} \omega(\eta, \xi)|_{\lambda=\lambda_0} dx.$$

*Proof.* The proof of each expression is obtained by substituting the characterization (4.11) and the multisymplectic decomposition of  $\mathbf{A}(x, \lambda)$  into theorem 1.11 of Pego & Weinstein (1992). If  $D(\lambda_0) = 0$ , Pego & Weinstein prove that (adapting the notation to the present setting)

$$D'(\lambda_0) = - \int_{-\infty}^{+\infty} \langle \Upsilon(x, \lambda), \frac{\partial}{\partial \lambda} \mathbf{A}(x, \lambda) \xi(x, \lambda) \rangle_{\mathbb{R}} |_{\lambda=\lambda_0} dx. \tag{4.12}$$

Using the definition of  $\mathbf{A}(x, \lambda)$  we have

$$\frac{\partial}{\partial \lambda} \mathbf{A}(x, \lambda) = -\mathbf{J}_c^{-1} \mathbf{M}.$$

Substituting this expression and (4.11) into (4.12),

$$\begin{aligned} D'(\lambda_0) &= \int_{-\infty}^{+\infty} \langle \mathbf{J}_c \eta(x, \lambda), \mathbf{J}_c^{-1} \mathbf{M} \xi(x, \lambda) \rangle_{\mathbb{R}} |_{\lambda=\lambda_0} dx \\ &= - \int_{-\infty}^{+\infty} \langle \eta(x, \lambda), \mathbf{M} \xi(x, \lambda) \rangle_{\mathbb{R}} |_{\lambda=\lambda_0} dx = \int_{-\infty}^{+\infty} \omega(\eta, \xi)|_{\lambda=\lambda_0} dx, \end{aligned}$$

using the definition of  $\omega$  in (2.2). A similar argument proves the expression for  $D''(\lambda_0)$ . ■

The asymptotics of the shape of the solitary wave (cf. equation (3.20)) show that  $\xi(x, 0)$  and  $\eta(x, 0)$  are multiples of  $\hat{Z}_x$ . A similar argument gives an abstract relation between the solitary-wave shape and the  $\lambda$ -derivatives of  $\xi(x, \lambda)$  and  $\eta(x, \lambda)$  at  $\lambda = 0$ .

**Proposition 4.3.** *The functions  $\xi(x, \lambda)$  and  $\eta(x, \lambda)$  satisfy*

$$\begin{aligned} \xi^0 &\stackrel{\text{def}}{=} C_+ \xi(x, \lambda)|_{\lambda=0} = \hat{Z}_x, & \xi_\lambda^0 &\stackrel{\text{def}}{=} C_+ \frac{\partial}{\partial \lambda} \xi(x, \lambda)|_{\lambda=0} = -\hat{Z}_c + C_1 \hat{Z}_x, \\ \eta^0 &\stackrel{\text{def}}{=} C_- \eta(x, \lambda)|_{\lambda=0} = \hat{Z}_x, & \eta_\lambda^0 &\stackrel{\text{def}}{=} C_- \frac{\partial}{\partial \lambda} \eta(x, \lambda)|_{\lambda=0} = \hat{Z}_c + C_2 \hat{Z}_x, \end{aligned}$$

for some constants  $C_1$  and  $C_2$ , where  $C_-$  and  $C_+$  are defined in (3.20).

*Proof.* The functions  $\xi^0$  and  $\eta^0$  satisfy

$$\mathbf{J}_c(\xi^0)_x = \mathbf{B}(x, c)(\xi^0) \quad \text{and} \quad \mathbf{J}_c(\eta^0)_x = \mathbf{B}(x, c)(\eta^0).$$

Differentiation of (2.11) with respect to  $x$  results in  $\mathbf{J}_c(\hat{Z}_x)_x = \mathbf{B}(x, c)\hat{Z}_x$ . Together with the asymptotics of  $\hat{Z}_x$  in (3.20), this shows that  $\xi^0 = \eta^0 = \hat{Z}_x$ .

The equations governing  $\xi_\lambda^0$  and  $\eta_\lambda^0$  are obtained by differentiating (4.1) and using the above results for  $\xi^0$  and  $\eta^0$ :

$$\begin{aligned} \mathbf{J}_c(\xi_\lambda^0)_x &= \mathbf{B}(x, c)(\xi_\lambda^0) - \mathbf{M}\xi^0 = \mathbf{B}(x, c)(\xi_\lambda^0) - \mathbf{M}\hat{Z}_x, \\ \mathbf{J}_c(\eta_\lambda^0)_x &= \mathbf{B}(x, c)(\eta_\lambda^0) + \mathbf{M}\eta^0 = \mathbf{B}(x, c)(\eta_\lambda^0) + \mathbf{M}\hat{Z}_x. \end{aligned}$$

Differentiation of (2.11) with respect to  $c$  results in

$$\mathbf{J}_c(\hat{Z}_c)_x = \mathbf{B}(x, c)\hat{Z}_c + \mathbf{M}\hat{Z}_x,$$

from which it follows that  $\xi_\lambda^0$  and  $\eta_\lambda^0$  can be written as a constant multiple of  $\hat{Z}_c$  plus a vector from the kernel of  $(\mathbf{J}_c D_x - \mathbf{B}(x, c))$ . Since  $\xi^0$  and  $\eta^0$  decay like  $e^{-\beta(c,0)|x|}$  for  $x \rightarrow +\infty$ , respectively  $x \rightarrow -\infty$ , the functions  $\xi_\lambda^0$  and  $\eta_\lambda^0$  decay like  $x e^{-\beta(c,0)|x|}$  or faster. The only vector in the kernel of  $(\mathbf{J}_c D_x - \mathbf{B}(x, c))$  with exponential decay is  $\hat{Z}_x$ . Hence this gives the expressions for  $\xi_\lambda^0$  and  $\eta_\lambda^0$ . ■

**Lemma 4.4.** *Assume the hypotheses of lemma 2.1 and let  $\Pi$  be as defined in (3.21b). Under the above hypotheses on the symplectic Evans function, it has the following properties at  $\lambda = 0$ :*

$$D(0) = D'(0) = 0 \quad \text{and} \quad \text{sgn } D''(0) = \Pi \text{sgn} \left( \frac{dI(\hat{Z})}{dc} - \frac{1}{2} \omega(Z_0^+ - Z_0^-, \partial_c Z_0^+) \right).$$

**Remark 4.5.** See corollaries 2.2 and 2.3 for special cases when the boundary term  $\frac{1}{2} \omega(Z_0^+ - Z_0^-, \partial_c Z_0^+)$  vanishes.

*Proof.* Using proposition 4.3 we have

$$D(\lambda)|_{\lambda=0} = \Omega(\eta, \xi)|_{\lambda=0} = \frac{1}{C_- C_+} \Omega(\eta^0, \xi^0) = \frac{1}{C_+ C_-} \Omega(\hat{Z}_x, \hat{Z}_x) = 0$$

by skew-symmetry of  $\Omega$ . With  $D(0) = 0$  we can apply lemma 4.2:

$$\begin{aligned} D'(\lambda)|_{\lambda=0} &= \int_{-\infty}^{+\infty} \omega(\eta, \xi)|_{\lambda=0} dx = \frac{1}{C_- C_+} \int_{-\infty}^{+\infty} \omega(\eta^0, \xi^0) dx \\ &= \frac{1}{C_+ C_-} \int_{-\infty}^{+\infty} \omega(\hat{Z}_x, \hat{Z}_x) dx = 0, \end{aligned}$$

which vanishes due to the skew-symmetry of  $\omega$ . Using proposition 4.3,

$$\begin{aligned} C_- C_+ \frac{\partial}{\partial \lambda} \omega(\eta, \xi)|_{\lambda=0} &= \omega(\eta_\lambda^0, \xi^0) + \omega(\eta^0, \xi_\lambda^0) \\ &= \omega(\hat{Z}_c + C_2 \hat{Z}_x, \hat{Z}_x) + \omega(\hat{Z}_x, -\hat{Z}_c + C_1 \hat{Z}_x) = 2\omega(\hat{Z}_c, \hat{Z}_x), \end{aligned}$$

which by hypothesis is integrable. Therefore, using lemmas 2.1 and 4.2,

$$\begin{aligned} D''(0) &= \int_{-\infty}^{+\infty} \frac{\partial}{\partial \lambda} \omega(\eta, \xi)|_{\lambda=0} dx \\ &= \int_{-\infty}^{+\infty} \frac{2}{C_+ C_-} \omega(\hat{Z}_c, \hat{Z}_x) dx = \frac{2}{C_+ C_-} \left( \frac{dI(\hat{Z})}{dc} - \frac{1}{2} \omega(Z_0^+ - Z_0^-, \partial_c Z_0^+) \right) \end{aligned}$$

Hence, using the definition of  $\Pi$  in (3.21 b),

$$D''(0) = \frac{2}{|C_- C_+|} \Pi \left( \frac{dI(\hat{Z})}{dc} - \frac{1}{2} \omega(Z_0^+ - Z_0^-, \partial_c Z_0^+) \right),$$

completing the proof. ■

When  $\lambda \in \mathbb{R}$ ,  $D(\lambda) \in \mathbb{R}$ . Define  $d_\infty$  to be the sign of  $D(\lambda_\infty)$  where  $\lambda_\infty \in \mathbb{R}$  is some positive value of  $\lambda$ . We do not rule out the possibility that  $D(\lambda)$  may oscillate for  $\lambda \in \mathbb{R}$  and  $\lambda \rightarrow \infty$  or that  $D(\lambda)$  may only be defined for a part of the real  $\lambda$ -axis (for example in § 5). On the other hand, if the initial-value problem is well posed and  $\mathbb{C}_+$ , the right-half complex  $\lambda$ -plane, does not contain any branch of the continuous spectrum, it is often the case that  $D(\lambda)$  will be of one sign for all real  $\lambda > \lambda_\infty$ .

Lemma 4.4, the hypothesis on  $D(\lambda)$  for  $\lambda$  real and large, and the intermediate-value theorem combine to prove the main result of this section.

**Theorem 4.6.** *Let  $\lambda_\infty$  be a real positive value of  $\lambda$  and let  $d_\infty = D(\lambda_\infty)$ . Let  $\Pi$  be as defined in (3.21) and assume the hypotheses of lemma 2.1. If*

$$d_\infty \Pi \left( \frac{dI(\hat{Z})}{dc} - \frac{1}{2} \omega(Z_0^+ - Z_0^-, \partial_c Z_0^+) \right) < 0,$$

*then the solitary wave or front is linearly spectrally unstable.*

The sign of  $d_\infty$  can be deduced in various ways. When the Evans function is analytic in an open set containing the positive real  $\lambda$ -axis and the initial-value problem is well posed, the asymptotics of the Evans function for large (real and positive)  $\lambda$  can be studied. Results of this type for the Evans function are proved in Alexander *et al.* (1990, proposition 2.2) and Pego & Weinstein (1992, proposition 1.17, corollary 1.18).

### 5. Instability for a nonlinear wave equation

Consider the semilinear wave equation  $u_{tt} - u_{xx} + f(u) = 0$  where  $f(u)$  is some smooth function. With natural hypotheses on  $f(u)$  it is straightforward to establish the existence of a solitary-wave solution using phase-plane techniques, and it is also well known that this state is unstable (cf. Grillakis *et al.* 1987, § 6A; Grillakis *et al.* 1990, § 11). The purpose of this section is two-fold. This Hamiltonian PDE fits into the framework of §§ 2–4 in a novel way: it has a multisymplectic formulation on  $\mathbb{R}^4$  with both two-forms  $\omega$  and  $\kappa$  non-degenerate on the full phase space. Secondly, it provides an elementary example where  $dI/dc$  can be either positive or negative, but in both cases  $D''(0) < 0$  and there exists an unstable eigenvalue—because  $\Pi$  also takes a different sign for the two cases. Since the points we intend to make about this equation are specific, we will work with the specific potential:  $f(u) = \alpha(u - u^2)$  where  $\alpha$  is a non-zero real parameter.

Introduce new variables  $v$ ,  $w$  and  $p$  defined by

$$v = u_t + p_x, \quad w = -u_x - p_t \quad \text{and} \quad w_t + v_x = 0. \quad (5.1 a)$$

Then the wave equation takes the form

$$v_t + w_x + f(u) = 0, \quad (5.1 b)$$

and it has the multisymplectic formulation

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z), \quad Z \in \mathbb{R}^4, \quad (5.2 a)$$

with

$$Z = \begin{pmatrix} u \\ v \\ w \\ p \end{pmatrix} \in \mathbb{R}^4, \quad \mathbf{M} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{K} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}. \quad (5.2 b)$$

The operators  $\mathbf{M}$  and  $\mathbf{K}$  define the two forms

$$\omega = dv \wedge du + dw \wedge dp, \quad \kappa = dw \wedge du + dv \wedge dp \quad (5.3)$$

and the function  $S(Z)$  is defined by

$$S(Z) = \frac{1}{2}(v^2 - w^2) + F(u), \quad \text{where} \quad F(u) = \int_0^u f(s) ds. \quad (5.4)$$

The matrices  $\mathbf{M}$  and  $\mathbf{K}$  are non-degenerate on  $\mathbb{R}^4$ . In fact,  $\mathbf{K} = \mathbf{PMP}$  where  $\mathbf{P}$  is a permutation matrix. The variable  $p$  is an artificial variable which acts to enforce the divergence condition  $w_t + v_x = 0$ , and, since only derivatives of  $p$  appear and  $S(Z)$  is independent of  $p$ , there is an affine translation symmetry associated with  $p$ . Let

$$V = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathcal{G}_\theta Z = Z + \theta V, \quad \forall \theta \in \mathbb{R}. \quad (5.5)$$

Then the system (5.2) is equivariant with respect to this one-parameter affine group. However, the invariant manifold of relative uniform states at infinity associated with

this group is not particularly interesting (the manifold of relative uniform states corresponds to a line parallel to the  $p$ -axis). Therefore we take the basic solitary wave to be

$$Z(x, t) = \mathcal{T}_{\tau(t)} \hat{Z}(x, c), \quad \tau(t) = ct + \tau_0,$$

with  $\hat{Z}(x, c)$  a homoclinic orbit of the Hamiltonian system

$$\mathbf{J}_c \hat{Z}_x = \nabla S(\hat{Z}), \quad \hat{Z} \in \mathbb{R}^4.$$

With  $f(u) = \alpha(u - u^2)$ , this system has the following orbit homoclinic to the origin:

$$\hat{u}(x, c) = \frac{3}{2} \operatorname{sech}^2\left(\frac{1}{2}\vartheta x + x_0\right), \quad \text{with } \hat{v}(x, c) = c\hat{w}(x, c) = -c\hat{u}_x(x, c), \quad \hat{p}(x, c) = 0, \tag{5.6}$$

where

$$\vartheta = \left(\frac{\alpha}{1 - c^2}\right)^{1/2} \quad \text{with the hypotheses } c^2 \neq 1 \quad \text{and} \quad \frac{\alpha}{1 - c^2} > 0.$$

By choosing  $x_0 = 0$  this state is reversible with respect to a standard reversor:

$$\hat{Z}(-x, c) = \mathbf{S}\hat{Z}(x, c) \quad \text{and} \quad \hat{Z}_x(-x, c) = -\mathbf{S}\hat{Z}_x(x, c) \quad \text{with } \mathbf{S} = \operatorname{diag}(1, -1, -1, 1).$$

Therefore, noting that  $\beta(c, 0) = \vartheta$  and  $Z_0^- = Z_0^+ = 0$ , the computation of  $C_+C_-$  in (3.20), (3.21) is reduced to

$$C_+C_- = - \lim_{x \rightarrow \infty} e^{2\vartheta x} \Omega(\mathbf{S}\hat{Z}_x(x, c), \hat{Z}_x(x, c)).$$

But

$$\Omega(\mathbf{S}\hat{Z}_x(x, c), \hat{Z}_x(x, c)) = \langle \mathbf{J}_c \mathbf{S}\hat{Z}_x(x, c), \hat{Z}_x(x, c) \rangle = 2\hat{u}_x \hat{w}_x (1 - c^2) = -(1 - c^2) \frac{d}{dx} \hat{u}_x^2.$$

Since  $d\hat{u}_x^2/dx < 0$  for  $x > 0$  and  $e^{2\vartheta x} \hat{u}_x \hat{w}_x$  is bounded as  $x \rightarrow \infty$ , it follows that

$$\Pi = \operatorname{sgn}(c^2 - 1).$$

To compute  $dI/dc$  we use the expression for the momentum in (2.14). The momentum density is

$$\frac{1}{2}\omega(\hat{Z}, \hat{Z}_x) = \frac{1}{2}\langle \mathbf{M}\hat{Z}, \hat{Z}_x \rangle = \frac{1}{2}(\hat{u}\hat{v}_x - \hat{v}\hat{u}_x).$$

Hence,

$$\begin{aligned} I(\hat{Z}) &= \int_{-\infty}^{+\infty} \frac{1}{2}\omega(\hat{Z}, \hat{Z}_x) dx = - \int_{-\infty}^{+\infty} \hat{v}\hat{u}_x dx = c \int_{-\infty}^{+\infty} \hat{u}_x^2 dx \\ &= \frac{9}{2}c\vartheta \int_{-\infty}^{+\infty} \tanh^2 s \operatorname{sech}^4 s ds \quad (s = \frac{1}{2}\vartheta x). \end{aligned}$$

Since  $\vartheta^2 = \alpha/(1 - c^2)$  and the integral is independent of  $c$ ,

$$\frac{d}{dc} I(\hat{Z}) = \frac{9}{2} \frac{\vartheta}{1 - c^2} \int_{-\infty}^{+\infty} \tanh^2 s \operatorname{sech}^4 s ds.$$

The product of  $\Pi$  and  $dI/dc$  is therefore negative for both cases:  $c^2 < 1$  and  $c^2 > 1$ .

But

$$\frac{dI}{dc} > 0 \text{ and } \Pi < 0 \text{ if } c^2 < 1 \quad \text{and} \quad \frac{dI}{dc} < 0 \text{ and } \Pi > 0 \text{ if } c^2 > 1.$$

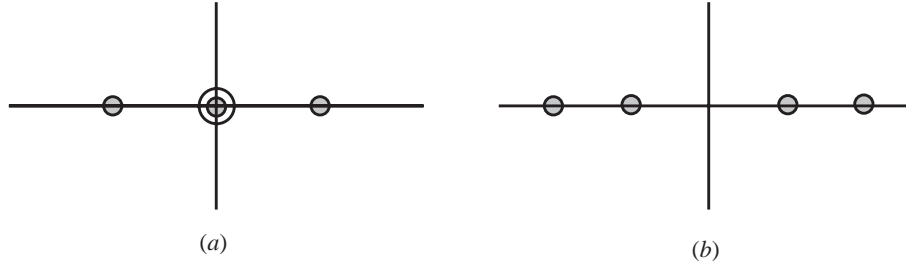


Figure 1. Qualitative position of the eigenvalues of  $\mathbf{A}_\infty(\lambda)$  in the complex  $\mu$ -plane for (a)  $\lambda = 0$  and (b)  $c^2 < 1$  and  $\lambda > 0$  is real.

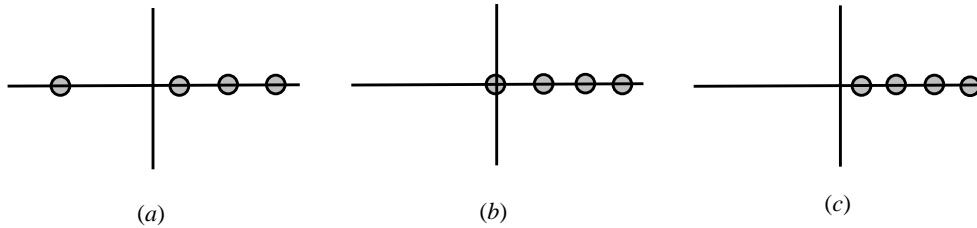


Figure 2. Qualitative position of the eigenvalues of  $\mathbf{A}_\infty(\lambda)$  in the complex  $\mu$ -plane when  $\lambda \in \mathbb{R}$  and  $\lambda > 0$  for the case  $c^2 > 1$ : (a)  $0 < \lambda < \sqrt{|\alpha|}$ ; (b)  $\lambda = \sqrt{|\alpha|}$ ; (c)  $\lambda > \sqrt{|\alpha|}$ .

Since  $Z_0^+ = Z_0^- = 0$ , the boundary terms in lemma 2.1 vanish (cf. corollary 2.2). By lemma 4.4 this suggests that  $D''(0) < 0$  in both cases; however, we have to check the hypotheses on the linearized stability equation: the asymptotic behaviour of  $\mathbf{B}(x, c)$  and the spectrum of  $\mathbf{A}_{\pm\infty}(\lambda)$ .

The linearized stability equation is given in equations (3.4)–(3.6) with  $U \in \mathbb{C}^4$  and

$$\mathbf{B}(x, c) = D^2S(\hat{Z}(x, c)) = \text{diag}[\alpha(1 - 2\hat{u}(x, c)), 1, -1, 0]$$

with

$$\lim_{x \rightarrow \pm\infty} \mathbf{B}(x, c) = \mathbf{B}_0 = \text{diag}[\alpha, 1, -1, 0].$$

The spectrum of  $\mathbf{A}_\infty(\lambda)$  consists of all  $\mu \in \mathbb{C}$  such that  $\Delta(\mu, \lambda) = 0$ , where

$$\Delta(\mu, \lambda) = \det[\mathbf{B}_0 - \mu\mathbf{J}_c - \lambda\mathbf{M}] = (\mu^2 - (\lambda - c\mu)^2)(\mu^2 - \alpha - (\lambda - c\mu)^2).$$

In figures 1 and 2 the spectrum for  $\lambda$  along the non-negative real axis is shown. Figure 1a shows the case  $\lambda = 0$ : the two zero eigenvalues are due to the affine symmetry (5.5) and the two hyperbolic eigenvalues lie at  $\mu = \pm\vartheta$ . For  $c^2 < 1$  and  $\text{Re}(\lambda) > 0$  the dimension of  $E_\infty^u(\lambda)$  and  $E_\infty^s(\lambda)$  is two and the qualitative position is shown in figure 1b for  $\lambda$  along the positive real axis. Therefore, when  $c^2 < 1$ , the hypothesis (1.19) is not satisfied. (However, for the present purposes the existence of an unstable ( $\lambda$ ) eigenvalue is easily deduced; see remarks in Appendix C.)

When  $c^2 > 1$ , the spectrum of  $\mathbf{A}_\infty(\lambda)$  changes qualitatively. Figure 2 shows the qualitative position of the eigenvalues of  $\mathbf{A}_\infty(\lambda)$  for  $\lambda$  along the positive real axis. The zero eigenvalue that arises when  $\lambda = \sqrt{|\alpha|}$  in figure 2b is associated with a branch of the continuous spectrum in the  $\lambda$ -plane.

The region of interest for the theory of §§ 2–4 is that associated with figure 2a, where the hypotheses (1.18) and (1.19) are satisfied. Let  $\mathbb{C}_+ = \{\lambda \in \mathbb{C} : \text{Re}(\lambda) > 0\}$ .



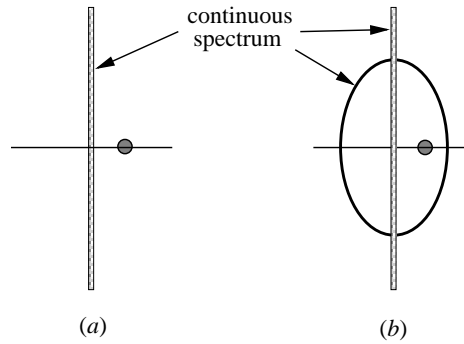


Figure 3. Qualitative properties of the continuous spectrum and the unstable eigenvalue in the complex  $\lambda$ -plane for the cases (a)  $c^2 < 1$  and (b)  $1 < c^2 < 1 + \frac{4}{5}$ .

Then the region  $\Lambda$  is defined by

$$\Lambda = \mathbb{C}_+ \cap \mathbb{E} \quad \text{where } \mathbb{E} = \{\lambda = \lambda_r + i\lambda_i : \lambda_r^2 + \lambda_i^2/c < |\alpha|\}. \quad (5.7)$$

For  $\lambda \in \Lambda$  all the hypotheses for theorem 4.6 are satisfied. Therefore, if  $d_\infty = D(\lambda_\infty) > 0$  for some  $\lambda_\infty \in \Lambda \cap \mathbb{R}$ , there exists an unstable eigenvalue. A positive value of  $d_\infty$  can be deduced in various ways; in fact an explicit expression for the Evans function can be computed in this case. In Appendix C, we show that the value of the unstable eigenvalue is

$$\lambda_u = \frac{1}{2}\sqrt{5}(|\alpha|/\theta), \quad 1 < c^2 < 1 + \frac{4}{5}.$$

The restriction that  $1 < c^2 < 1 + \frac{4}{5}$  follows from the requirement that  $\lambda_u$  must lie in  $\Lambda$  and so  $\lambda_u < \sqrt{|\alpha|}$ . Although, when  $c^2 > 1 + \frac{4}{5}$  there does not exist an unstable eigenvalue, the solitary wave is still unstable because of the branch of unstable continuous spectrum. A qualitative picture of the unstable eigenvalue and the continuous spectrum in the complex  $\lambda$ -plane is shown in figure 3.

### 6. Instability for an equivariant variant of the gKdV equation

Solitary-wave states of the gKdV equation with  $f'(u) = u^p$ , which are biasymptotic to the trivial state at plus and minus infinity, have the property that there is a change of stability at  $p = 4$  and these solitary waves are unstable for all  $p > 4$  (see, for example, Bona *et al.* 1987; Pego & Weinstein 1992, § 2a, and references therein). When the solitary-wave state is biasymptotic to a non-zero constant at infinity, Gardner (1997) has proved that the critical exponent is no longer  $p = 4$ , but is dependent on the value of the constant state at infinity. In this section, we use the multisymplectic formulation of the gKdV equation introduced in (1.5)–(1.8) and the theory of §§ 2–4 to extend the above results. The non-zero constant states at infinity are characterized as points on a one-dimensional manifold,  $\mathcal{M}_\infty$  of relative uniform states. We show that  $\Pi = +1$  for solitary waves which are biasymptotic to  $\mathcal{M}_\infty$  and that  $dI/dc < 0$  is a sufficient condition for the existence of an unstable eigenvalue, for any continuously differentiable  $f(u)$  for which a solitary wave exists.

The structure of the gKdV equation needed for the analysis is defined by the pair of two forms  $\omega$  and  $\kappa$  given in (1.8) and the functional  $S(Z)$  given in (1.7) with coordinates  $Z = (\phi, u, v, w) \in \mathbb{R}^4$ . This system has a natural one-parameter

symmetry group  $G$  with action

$$\mathcal{G}_\theta Z = Z + \theta V, \quad \text{for any } \theta \in \mathbb{R}, \quad \text{where } V = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (6.1)$$

The manifold of relative uniform states associated with the  $\kappa$ -system is constructed as follows. Let

$$P(Z) = \langle \mathbf{K}V, Z \rangle \quad \text{then } \nabla P(Z) = \mathbf{K}V;$$

that is,  $P(Z)$  is the (spatial) invariant associated with the symmetry group  $G$ . Therefore, using the theory of Appendix A, there is a one-parameter family of relative uniform states:

$$Z(x) = \mathcal{G}_{\theta(x)} Z_0, \quad \text{where } \theta(x) = ax + \theta_0 \quad (6.2)$$

and  $Z_0 \in \mathbb{R}^4$  satisfies

$$\nabla S(Z_0) = a \nabla P(Z_0) \quad \text{and} \quad P(Z_0) = \mathcal{P}. \quad (6.3)$$

Solving this equation for  $Z_0$ ,

$$Z_0 = \begin{pmatrix} \phi_0 \\ a \\ 0 \\ f(a) \end{pmatrix}, \quad \text{with } P(Z_0) = -f(a) = \mathcal{P}, \quad (6.4)$$

and  $\phi_0$  is an arbitrary constant. This state is non-degenerate when  $d\mathcal{P}/da = -f'(a) \neq 0$ . Suppose the state is non-degenerate and let  $Z_0$  be as in (6.4) with  $\phi_0 = 0$ . Then the manifold  $\mathcal{M}_\infty$  is a line through the point  $Z_0$ :

$$\mathcal{M}_\infty = \{Z_0 + \theta V : \theta \in \mathbb{R}\}. \quad (6.5)$$

This line passes through the point on the  $u$ -axis at  $u = a$ , which is the non-constant state at infinity associated with the scalar form of the gKdV equation.

Since  $\mathbf{K}$  is non-degenerate, this example also satisfies the hypotheses of proposition A 1 in Appendix A. The operator  $\mathbf{B}_0$  takes the form

$$\mathbf{B}_0 = D^2 S(Z_0) - a D^2 P(Z_0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & f'(a) & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \quad (6.6)$$

The spectrum of  $\mathbf{B}_0$  consists of a zero eigenvalue due to symmetry (the eigenvector is  $V$  in (6.1)), two positive eigenvalues and one negative eigenvalue. Therefore  $\varrho = 1$  in proposition A 1 and

$$\text{sgn}(-1)^\varrho j(a) = -\text{sgn} \frac{d\mathcal{P}}{da} = \text{sgn} f'(a).$$

Hence, by proposition A 1 and corollary A 2,  $f'(a) < 0$  is a sufficient condition for the existence of a pair of hyperbolic real eigenvalues of the spectral problem  $\mathbf{B}_0 U = \mu \mathbf{K}U$ . This result could have easily been deduced explicitly, but provides an illustration of proposition A 1.

However, we are interested in the spectrum of the linearization about the state (6.2), relative to a moving frame. First we note that a solitary-wave state which is biasymptotic to  $\mathcal{M}_\infty$  is of the form

$$Z(x, t) = \mathcal{G}_{\theta(x)}[Z_0^- + \mathcal{T}_{\tau(t)}\hat{Z}(x, c)], \tag{6.7}$$

where  $Z_0^-$  satisfies (6.3),  $\tau(t) = ct + \tau_0$  and  $\hat{Z}(x, c)$  is a heteroclinic orbit of the Hamiltonian system

$$\mathbf{J}_c \hat{Z}_x = \nabla V(\hat{Z}) = \nabla S(Z_0^- + \hat{Z}) - a \nabla P(Z_0^- + \hat{Z}). \tag{6.8}$$

This Hamiltonian system can be reduced to a planar system. With  $\hat{u}(x, c)$  satisfying

$$\hat{u}_{xx} = 2c\hat{u} + f(a + \hat{u}) - f(a), \tag{6.9 a}$$

the other components of  $\hat{Z}$  are defined by

$$\hat{\phi}_x = \hat{u}, \quad \hat{v} = \hat{u}_x, \quad \hat{w} = c\hat{u}. \tag{6.9 b}$$

Note that, although the orbit  $\hat{Z}$  of (6.8) is a heteroclinic orbit in  $\mathbb{R}^4$ , the orbit of (6.9 a), in the reduced phase space  $\mathbb{R}^2$ , is taken to be a homoclinic orbit. We assume that there exists an even solution of (6.9 a) satisfying  $\hat{u} \rightarrow 0$  as  $x \rightarrow \pm\infty$ . Such a state is reversible:

$$\mathbf{S}\hat{Z}(x, c) = \hat{Z}(-x, c) \quad \text{and} \quad \mathbf{S}\hat{Z}_x(x, c) = -\hat{Z}_x(-x, c) \quad \text{with} \quad \mathbf{S} = \text{diag}(-1, 1, -1, 1).$$

Using these abstract properties of the solitary wave, it is easily deduced that  $\Pi = +1$ . First note that

$$\mathbf{J}_c \mathbf{S} \hat{Z}_x(x, c) = \begin{pmatrix} \hat{w}_x - c\hat{u}_x \\ \hat{v}_x - c\hat{\phi}_x \\ \hat{u}_x \\ \hat{\phi}_x \end{pmatrix}.$$

Then, using the definitions in (3.20), (3.21),

$$\begin{aligned} C_+ C_- &= - \lim_{x \rightarrow +\infty} e^{2\beta(c,0)x} \Omega(\mathbf{S}\hat{Z}_x(x, c), \hat{Z}_x(x, c)) \\ &= -2 \lim_{x \rightarrow +\infty} e^{2\beta(c,0)x} \hat{u}_x(x, c) \hat{v}_x(x, c). \end{aligned}$$

However,  $\hat{v} = \hat{u}_x$ ,  $d\hat{u}_x^2/dx < 0$  for  $x > 0$  and  $e^{2\beta(c,0)x} \hat{v} \hat{u}_x$  is bounded as  $x \rightarrow \infty$ ; hence,  $\Pi = +1$ .

The phase shift  $\gamma$  between the asymptotic points on  $\mathcal{M}_\infty$  at  $\pm\infty$ , defined by  $Z_0^+ = \mathcal{G}_\gamma Z_0^- = Z_0^- + \gamma V$  is computed as follows. Fix  $Z_0^-$  to be as in (6.4); this implies that  $\lim_{x \rightarrow -\infty} \hat{\phi}(x, c) = \phi_0$ , where  $\phi_0$  is arbitrary. Then

$$\hat{\phi}(x, c) = \phi_0 + \int_{-\infty}^x \hat{u}(s, c) ds \quad \text{and} \quad \lim_{x \rightarrow +\infty} \hat{\phi}(x, c) = \phi_0 + \int_{-\infty}^{+\infty} \hat{u}(s, c) ds.$$

Hence

$$\lim_{x \rightarrow +\infty} \hat{Z}(x, c) = Z_0^+ - Z_0^- = \gamma V \quad \text{with} \quad \gamma = \int_{-\infty}^{+\infty} \hat{u}(s, c) ds.$$

In order to apply theorem 4.6, it is necessary to check the spectrum of  $\mathbf{A}_{\pm\infty}(\lambda)$ . When  $\lambda = 0$ , the spectrum takes the form shown in figure 4a with

$$\beta(c, 0) = [2c - f'(a)]^{1/2}, \quad \text{where } 2c - f'(a) > 0.$$

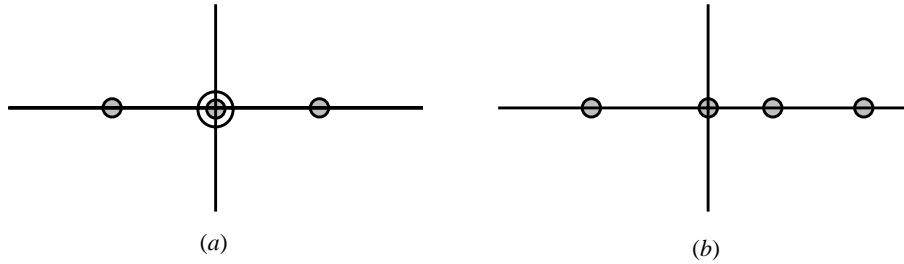


Figure 4. Eigenvalues of  $\mathbf{A}_\infty(\lambda)$  in the complex  $\mu$ -plane for (a)  $\lambda = 0$  and (b)  $\lambda \in \mathbb{R}$ , small and positive.

The set  $\Lambda$  can be taken to be  $\mathbb{C}_+$ , the right-half  $\lambda$ -plane. For  $\lambda \in \Lambda$  there is one root with negative real part, a zero eigenvalue and two with positive real part. Figure 4b shows the qualitative position of the spectrum of  $\mathbf{A}_{\pm\infty}(\lambda)$  when  $\lambda \ll 1$  is real and positive.

The asymptotics of the Evans function for the gKdV equation have been studied by Pego & Weinstein (1992, § 2a) and with minor modification of those results it can be deduced that  $D(\lambda)$  is positive for  $\lambda \in \mathbb{R}$  as  $\lambda \rightarrow +\infty$ . Moreover, since the symmetry group consists only of affine translations, the boundary term in lemma 2.1 vanishes (cf. corollary 2.3). Therefore, by theorem 4.6, for any solitary-wave state biasymptotic to  $\mathcal{M}_\infty$  (for the scalar gKdV equation this corresponds to  $u \rightarrow a$  as  $x \rightarrow \pm\infty$ ), there exists an unstable eigenvalue when  $dI/dc < 0$ . In general, an explicit expression for such a solitary wave is not known for general  $f$ . The two most well-known cases of special interest are when  $f'(u) = u^p$  (and  $a = 0$ ), in which case  $dI/dc < 0$  when  $p > 4$ , and the case when  $0 < a \ll 1$  and  $f'(u) = u^p$ , which corresponds to the case studied by Gardner (1997). Gardner (1997, proposition 3.2) proves that the critical exponent is no longer  $p = 4$ , but is increased when  $a$  is positive and small.

When the PDE has an affine translation symmetry it is possible to create artificial instabilities along the group orbit, whose implication is vacuous for the dynamics of the system. This point is also related to the correspondence between solutions of the multisymplectic formulation of the gKdV equation and its classical scalar formulation.

Consider the linear operator

$$\mathcal{D} = \begin{pmatrix} \partial_x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \partial_x \end{pmatrix}$$

acting on vector-valued functions whose first and fourth components are differentiable. The kernel of  $\mathcal{D}$  consists of vector-valued functions whose first and fourth components are independent of  $x$ . Comparing the multisymplectic formulation (1.5) with the original scalar gKdV equation, it is clear that the implication of the kernel of  $\mathcal{D}$  for the scalar gKdV equation is vacuous. Therefore, in order for a one-to-one correspondence between solutions of (1.5) and the scalar gKdV equation, (1.5) should be studied on a subspace of functions associated with the complement of the kernel of  $\mathcal{D}$ . We will not consider this question here as it is outside the main scope of the paper. However, there is one implication which produces an interesting spectral anomaly.

The linearization of (1.5) about a solitary wave takes the form

$$\mathbf{M}\hat{U}_t + \mathbf{J}_c\hat{U}_x = \mathbf{B}(x, c)\hat{U}, \tag{6.10}$$

where

$$\mathbf{B}(x, c) = D^2(Z_0^- + \hat{Z}) - aD^2P(Z_0^- + \hat{Z}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & f'(a + \hat{u}) & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

with associated spectral problem

$$\mathbf{J}_cU_x = [\mathbf{B}(x, c) - \lambda\mathbf{M}]U. \tag{6.11}$$

If we consider the spectral problem (6.11) on a space of functions which are bounded and uniformly continuous, every  $\lambda \in \mathbb{C}$  is in the spectrum: for every  $\lambda \in \mathbb{C}$  there exists a solution of (6.11), bounded for all  $x \in \mathbb{R}$ . Moreover, equation (6.10) has solutions which are bounded for all  $x \in \mathbb{R}$  and grow faster than exponential in time. To see this, let  $\nu(t)$  be any differentiable function of time. Then, an exact solution of (6.10) is

$$\hat{U} = \begin{pmatrix} \nu(t) \\ 0 \\ 0 \\ \nu'(t) \end{pmatrix}. \tag{6.12}$$

This solution is clearly a bounded function of  $x \in \mathbb{R}$ . Letting  $\nu(t) = \nu_0 e^{\lambda t}$  shows that  $U = \nu_0(1, 0, 0, \lambda)$  is a solution of the spectral problem (6.11) for any  $\lambda \in \mathbb{C}$ . Taking  $\nu(t) = \exp(e^{\lambda t})$  (or any such function) in (6.12) shows that super-exponential growth is also possible.

While such an anomaly is admissible from a spectral theory point of view, it is a curiosity associated with the temporal flow along the orbit of the affine translation group  $G$ . Another view is that the solution  $\hat{U}$  in (6.12) is in fact in the kernel of  $\mathcal{D}$  which is ‘orthogonal’ to the natural solution space for the gKdV equation, and, in an analysis of the initial-value problem, should be factored out.

### 7. Instability for an equivariant Boussinesq system

The Boussinesq equations are a class of model equations for dispersive shallow-water waves. They consist of an equivalence class of model PDEs with many variants. In this section, we will consider one of the most well-known variants of the Boussinesq systems, one of the systems first proposed by Boussinesq (cf. Whitham 1974, p. 462):

$$\left. \begin{aligned} h_t + (uh)_x &= 0, \\ u_t + (h + \frac{1}{2}u^2)_x + \frac{1}{3}h_{xtt} &= 0, \end{aligned} \right\} x \in \mathbb{R}, \quad t > 0, \tag{7.1}$$

where  $(h, u)$  are scalar-valued functions of  $(x, t)$  with  $h(x, t)$  required to be positive for all  $(x, t)$ . For shallow-water flow in a one-dimensional channel, aligned in the  $x$ -direction,  $h(x, t)$  is the elevation of the free surface and  $u(x, t)$  is the average velocity in the  $x$ -direction.

Classical solitary-wave solutions of (7.1) take the form

$$(h(x, t), u(x, t)) = \mathcal{T}_{\tau(t)}(\tilde{h}(x, c), \tilde{u}(x, c)), \quad \tau(t) = ct + \tau_0, \tag{7.2}$$

with asymptotic conditions

$$\lim_{x \rightarrow \pm\infty} \tilde{h}(x, c) = h_0 \quad \text{and} \quad \lim_{x \rightarrow \pm\infty} \tilde{u}(x, c) = u_0, \quad (7.3)$$

where  $h_0$  and  $u_0$  are constant values of elevation and velocity, respectively, at infinity. In this section, the Boussinesq system will be formulated as an equivariant Hamiltonian system on a multisymplectic structure, and the theory of §§ 2–4 is applied to deduce an instability criterion for solitary-wave states. The asymptotic states  $(h_0, u_0)$  will be related to points on a two-dimensional manifold,  $\mathcal{M}_\infty$ , of relative uniform states. We will show that solitary-wave states which are biasymptotic to  $\mathcal{M}_\infty$  have the property  $\mathcal{H} = +1$ . The linearization about this family of solitary-wave states has the property that  $n = 3$  and  $q = 2$  and the spectrum of  $\mathbf{A}_{\pm\infty}(\lambda)$  satisfies the hypothesis (1.19) when  $\Lambda = \mathbb{C}_+$ . The symmetry group  $G$  is a two-parameter group of affine translations (cf. equation (7.9)). By corollary 2.3, the boundary term vanishes and therefore theorem 4.6 leads to a criterion based on  $dI/dc$  for the existence of an unstable eigenvalue.

The multisymplectic formulation of (7.1) is constructed as follows (cf. Bridges (1995), where a different multisymplectic formulation is presented, and the stability of periodic travelling waves is studied). Introduce new variables,

$$u = \phi_x, \quad q = uh, \quad \eta = h_t, \quad h = r_x \quad \text{and} \quad p = h + \frac{1}{2}u^2 + \phi_t + \frac{1}{3}\eta t; \quad (7.4)$$

then with

$$Z = \begin{pmatrix} h \\ q \\ \eta \\ \phi \\ r \\ p \end{pmatrix}, \quad \mathbf{M} = \begin{bmatrix} 0 & 0 & \frac{1}{3} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (7.5)$$

and

$$S(Z) = ph - \frac{1}{2}h^2 + \frac{1}{2}q^2/h - \frac{1}{6}\eta^2, \quad (7.6)$$

the system (7.1) takes the multisymplectic form

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z), \quad Z \in \mathbb{R}^6, \quad x \in \mathbb{R}. \quad (7.7)$$

The three differential two-forms associated with (7.7) are

$$\left. \begin{aligned} \omega &= dh \wedge d\phi + \frac{1}{3}dh \wedge d\eta, \\ \kappa &= dq \wedge d\phi + dp \wedge dr, \\ \Omega &= \kappa - c\omega, \quad c \neq 0, \end{aligned} \right\} \quad (7.8)$$

and through the identities (2.2), they define the matrices  $\mathbf{M}$ ,  $\mathbf{K}$  and  $\mathbf{J}_c$ . The two-forms in (7.8) are symplectic on three different spaces:

$$\text{rank}(\omega) = 2, \quad \text{rank}(\kappa) = 4 \quad \text{rank}(\Omega) = 6.$$

The system (7.7) is equivariant with respect to the two-parameter affine group  $G$  with action on  $\mathbb{R}^6$ :

$$\mathcal{G}_\theta Z = Z + \theta_1 V_1 + \theta_2 V_2, \quad \text{for any } (\theta_1, \theta_2) \in \mathbb{R}^2, \quad V_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad V_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \quad (7.9)$$

The  $\kappa$ -symplectic flow of this group generates two invariants, denoted by  $P_1(Z)$  and  $P_2(Z)$ , which satisfy

$$\mathbf{K}V_1 = \nabla P_1(Z) \quad \text{and} \quad \mathbf{K}V_2 = \nabla P_2(Z), \quad (7.10)$$

or

$$P_1(Z) = \langle \mathbf{K}V_1, Z \rangle = p \quad \text{and} \quad P_2(Z) = \langle \mathbf{K}V_2, Z \rangle = q. \quad (7.11)$$

Using the theory of Appendix A, we define a two-parameter family of relative uniform states by

$$Z(x) = \mathcal{G}_{\theta(x)} Z_0 \quad \text{with } \theta(x) = (\theta_1(x), \theta_2(x)) \quad \text{and} \quad \theta_j(x) = a_j x + \theta_j^0, \quad j = 1, 2. \quad (7.12)$$

Let  $\mathbf{a} = (a_1, a_2)$ ; the connection with the classical asymptotic conditions in (7.3) is that  $a_1 = h_0$  and  $a_2 = u_0$ .

According to the theory of Appendix A, the point  $Z_0$  and the parameters  $\mathbf{a}$  are determined by the constrained variational principle:  $Z_0 \in \mathbb{R}^6$  is a critical point of  $S$  restricted to level sets of the functions  $P_1(Z) = \mathcal{P}_1$ ,  $P_2(Z) = \mathcal{P}_2$ . With  $a_1$  and  $a_2$  as Lagrange multipliers, the Lagrange necessary condition for this constrained variational principle is

$$\nabla S(Z_0) = a_1 \nabla P_1(Z_0) + a_2 \nabla P_2(Z_0) \quad \text{with } P_j(Z_0) = \mathcal{P}_j, \quad j = 1, 2. \quad (7.13)$$

This constrained variational principle is easily solved to find

$$Z_0 = \begin{pmatrix} a_1 \\ a_1 a_2 \\ 0 \\ \phi_0 \\ r_0 \\ a_1 + \frac{1}{2} a_2^2 \end{pmatrix} \quad \text{with} \quad \left\{ \begin{array}{l} P_1(Z_0) = a_1 + \frac{1}{2} a_2^2 = \mathcal{P}_1, \\ P_2(Z_0) = a_1 a_2 = \mathcal{P}_2, \end{array} \right\} \quad (7.14)$$

and  $\phi_0$  and  $r_0$  are arbitrary real parameters. The relative uniform state is non-degenerate when

$$j(\mathbf{a}) \stackrel{\text{def}}{=} \det \begin{bmatrix} \frac{\partial \mathcal{P}_1}{\partial a_1} & \frac{\partial \mathcal{P}_2}{\partial a_2} \\ \frac{\partial \mathcal{P}_2}{\partial a_1} & \frac{\partial \mathcal{P}_1}{\partial a_2} \end{bmatrix} = \det \begin{pmatrix} 1 & a_2 \\ a_2 & a_1 \end{pmatrix} = a_1 - a_2^2 \neq 0. \quad (7.15)$$

Using the identification  $a_1 = h_0$  and  $a_2 = u_0$ , the non-degeneracy condition is equivalent to

$$j(\mathbf{a}) = h_0(1 - \mathcal{F}^2) \neq 0 \quad \text{where } \mathcal{F} = \frac{u_0}{\sqrt{h_0}}.$$

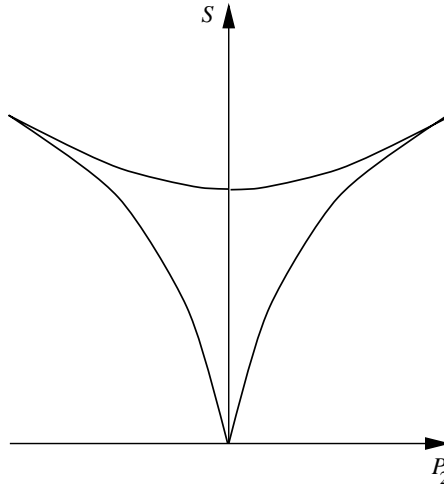


Figure 5. Swallowtail for uniform states of the Boussinesq system, with  $\mathcal{P}_1 > 0$  fixed.

The parameter  $\mathcal{F}$  is the *Froude number* and degeneracy corresponds to a Froude number of unity. The above theory gives a geometric characterization of classical uniform flow theory. It also recovers the swallowtail formulation of uniform flow of Sewell & Porter (1980, § 4). The Sewell–Porter swallowtail is constructed as follows. Evaluating the function  $S$  at  $Z_0$ , we have a map from  $\mathbf{a} \in \mathbb{R}^2$  to  $\mathbb{R}^3$  with coordinates  $(S, P_1, P_2)$ . This map defines a surface in  $\mathbb{R}^3$  parametrized by  $(a_1, a_2)$ , and a section of this surface, with the value of  $P_1$  fixed, is shown in figure 5 (cf. Sewell & Porter 1980, fig. 4).

Every point on the swallowtail in figure 5 corresponds to a particular uniform state in the two-parameter family (7.12) and the two cusp points in the figure correspond to the points where the determinant in (7.15) vanishes.

For any non-degenerate point  $Z_0$  satisfying the constrained variational principle (7.13), the manifold  $\mathcal{M}_\infty$  is a two-dimensional plane passing through  $Z_0$ :

$$\mathcal{M}_\infty = \{Z_0 + \theta_1 V_1 + \theta_2 V_2 : (\theta_1, \theta_2) \in \mathbb{R}^2\}. \tag{7.16}$$

This plane passes through points with  $h = a_1 = h_0$  and  $u = a_2 = u_0$ .

A solitary wave which is biasymptotic to  $\mathcal{M}_\infty$  takes the form

$$Z(x, t) = \mathcal{G}_{\theta(x)}[Z_0^- + \mathcal{T}_{\tau(t)} \hat{Z}(x, c)], \tag{7.17}$$

where  $Z_0^-$  is any non-degenerate solution of (7.13) and  $\hat{Z}(x, c)$  is a heteroclinic orbit of the Hamiltonian system

$$\mathbf{J}_c \hat{Z}_x = \nabla V(\hat{Z}) = \nabla S(Z_0^- + \hat{Z}) - a_1 \nabla P_1(Z_0^- + \hat{Z}) - a_2 \nabla P_2(Z_0^- + \hat{Z}), \quad \hat{Z} \in \mathbb{R}^6. \tag{7.18}$$

An explicit expression for a heteroclinic orbit of this system is not, in general, known, even though it is an integrable Hamiltonian system and can be reduced to a planar system. However, we will deduce that such orbits have the property that  $\mathcal{I} = +1$ .

Writing out (7.18),

$$\hat{r}_x = \hat{h}, \quad \hat{q} = c\hat{h}, \quad \hat{\eta} = -c\hat{h}_x, \quad \hat{\phi}_x = \frac{(c - a_2)\hat{h}}{a_1 + \hat{h}}, \quad \hat{p} = 0,$$



with  $\hat{h}$  satisfying

$$\hat{h}_{xx} + f(\hat{h}) = 0, \tag{7.19 a}$$

where

$$f(\hat{h}) = -\frac{3}{c^2} \left[ \frac{c(c - a_2)\hat{h}}{a_1 + \hat{h}} - \hat{h} + \frac{1}{2}a_2^2 - \frac{1}{2} \left( \frac{a_1 a_2 + c\hat{h}}{a_1 + \hat{h}} \right)^2 \right]. \tag{7.19 b}$$

Even though  $\hat{Z}(x, c)$  is a heteroclinic orbit, the function  $\hat{h}(x, c)$  is required to be a homoclinic orbit in the reduced two-dimensional phase space associated with (7.19 a). Suppose there exists a homoclinic orbit of (7.19 a) with  $\hat{h}(x, c) \rightarrow 0$  as  $x \rightarrow \pm\infty$  which is an even function of  $x$ . Then the heteroclinic orbit  $\hat{Z}(x, c)$  is reversible,

$$\mathbf{S}\hat{Z}(x, c) = \hat{Z}(-x, c) \quad \text{and} \quad \mathbf{S}\hat{Z}_x(x, c) = -\hat{Z}_x(-x, c)$$

with

$$\mathbf{S} = \text{diag}[1, 1, -1, -1, -1, 1].$$

Therefore,

$$\Omega(\mathbf{S}\hat{Z}_x(x, c), \hat{Z}_x(x, c)) = -\frac{1}{3}c^2 \frac{d}{dx} \hat{h}^2.$$

By an argument similar to that for  $C_-C_+$  in § 6, we deduce that the sign of  $C_-C_+$  is positive and hence  $\Pi = +1$  for any such heteroclinic orbit of (7.18).

The phase shift on  $\mathcal{M}_\infty$  between the asymptotic points  $Z_0^-$  and  $Z_0^+$  (with  $\phi_0 = r_0 = 0$ ) is computed as follows. By definition  $Z_0^+ = \mathcal{G}_\gamma Z_0^- = Z_0^- + \gamma_1 V_1 + \gamma_2 V_2$ . Fix  $Z_0^-$  to be as in (7.14). Then

$$\lim_{x \rightarrow -\infty} \hat{r}(x, c) = r_0 \quad \text{and} \quad \lim_{x \rightarrow -\infty} \hat{\phi}(x, c) = \phi_0,$$

where  $(r_0, \phi_0)$  are arbitrary parameters, and

$$\hat{r}(x, c) = r_0 + \int_{-\infty}^x \hat{h}(s, c) ds \quad \text{and} \quad \hat{\phi}(x, c) = \phi_0 + \int_{-\infty}^x \frac{(c - a_2)\hat{h}(s, c)}{a_1 + \hat{h}(s, c)} ds.$$

Hence  $\lim_{x \rightarrow +\infty} \hat{Z}(x, c) = Z_0^+ - Z_0^- = \gamma_1 V_1 + \gamma_2 V_2$  with

$$\gamma_1 = \int_{-\infty}^{+\infty} \hat{h}(s, c) ds \quad \text{and} \quad \gamma_2 = (c - a_2) \int_{-\infty}^{+\infty} \frac{\hat{h}(s, c)}{a_1 + \hat{h}(s, c)} ds.$$

To verify that the hypotheses (1.18) and (1.19) are met, it is necessary to study the spectrum of  $\mathbf{A}_{\pm\infty}(\lambda)$ , and it is shown in figure 6. In figure 6a, the spectrum when  $\lambda = 0$  is shown. There are four zero eigenvalues due to symmetry and one pair of hyperbolic real eigenvalues with values

$$\mu = \pm\sqrt{-f'(0)}, \quad \text{where} \quad f'(0) = -\frac{3}{a_1 c^2} \left[ -a_1 \left( 1 - \frac{a_2^2}{a_1} \right) + c^2 - 2ca_2 \right],$$

and  $a_1, a_2, c$  are such that  $f'(0) < 0$ . When  $\lambda \neq 0$ , two zero eigenvalues remain at the origin: one has strictly negative real part and the other three have strictly positive real part for all  $\lambda \in \Lambda = \mathbb{C}^+$ . A qualitative picture of the spectrum of  $\mathbf{A}_{\pm\infty}(\lambda)$  when  $\lambda \ll 1$  is real and positive is shown in figure 6b.

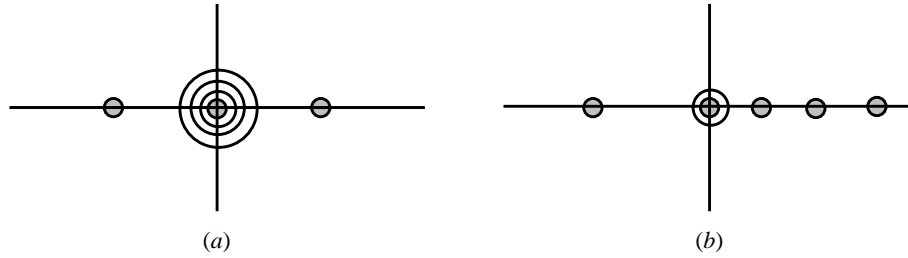


Figure 6. Eigenvalues of  $\mathbf{A}_\infty(\lambda)$  in the complex  $\mu$ -plane for (a)  $\lambda = 0$  and (b)  $\lambda \in \mathbb{R}$ , small and positive.

In order to deduce the sign of  $D(\lambda)$  as  $\lambda \rightarrow \infty$  along the real axis we can appeal to the recent results of Pego & Weinstein (1997). For the system (7.1), linearized about a solitary wave asymptotic to the trivial state, Pego & Weinstein prove that the Evans function is positive for  $\lambda \rightarrow +\infty$  along the real axis and that the classical solitary wave with  $u_0 = 0$  is stable in the KdV limit (i.e. Froude number near unity; in a neighbourhood of the cusp in figure 5a). With minor modification of their analysis it can be deduced that there exists  $d_\infty > 0$  and therefore, by theorem 4.6,  $dI/dc < 0$  is a sufficient condition for the existence of an unstable eigenvalue in the linearization about a solitary wave of the form (7.17).

## 8. Concluding remarks

In this section we sketch some of the implications of the framework presented here for other Hamiltonian PDEs. One of the important features of the result is that an explicit expression for the solitary wave is not needed, and therefore the theory applies to PDEs where a solitary wave is known to exist only abstractly. An example is the solitary-wave state of the Boussinesq-type model proposed by Bona & Smith (1976):

$$\left. \begin{aligned} \eta_t + u_x + (u\eta)_x - \frac{1}{3}\eta_{xxt} &= 0, \\ u_t + \eta_x + uu_x - \frac{1}{3}(u_t + \eta_x)_{xx} &= 0. \end{aligned} \right\} \quad (8.1)$$

This system is formally equivalent to the Boussinesq system as a model for shallow-water waves but it has several important differences, one of which is that the restriction of (8.1) to a moving frame results in a sixth-order ODE which has a first integral but is not in general integrable and therefore may have very complicated solutions. However, an abstract proof of the existence of a solitary-wave state, biasymptotic to the trivial state, of (8.1) has been given by Toland (1981). The stability of this class of solitary waves has never been studied, partly because most existing methods require an explicit solution. The framework of this paper applies to (8.1) as follows.

The system (8.1) can be reformulated as a Hamiltonian system on a multisymplectic structure by introducing new variables

$$\phi_x = u, \quad \gamma_x = \eta + 1, \quad h = \eta_x, \quad v = h_t \quad \text{and} \quad p = \frac{1}{3}(u_t + \eta_x).$$

Integrating the second equation of (8.1) leads to a Bernoulli-type equation:

$$\phi_t + \eta + 1 + \frac{1}{2}u^2 - \frac{1}{3}(u_t + \eta_x)_x = r.$$

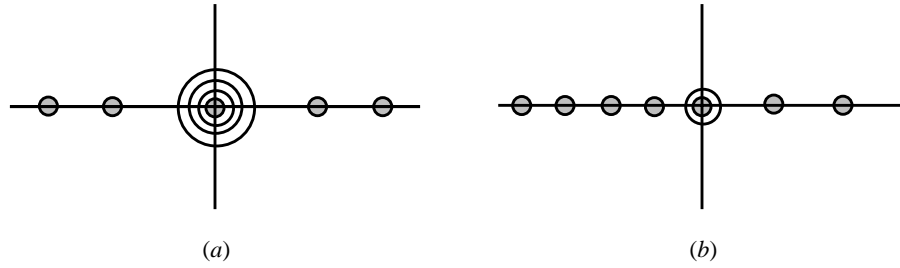


Figure 7. Eigenvalues of  $\mathbf{A}_{\pm\infty}(\lambda)$  for (8.1) in the complex  $\mu$ -plane for (a)  $\lambda = 0$  and (b)  $\lambda \in \mathbb{R}$ , small and positive.

In terms of the variables  $Z = (\eta, h, v, p, u, \phi, \gamma, r)$ , the system (8.1) takes the form

$$\mathbf{M}Z_t + \mathbf{K}(Z)Z_x = \nabla S(Z), \quad Z \in \mathbb{R}^8, \tag{8.2}$$

where the matrices  $\mathbf{M}$  and  $\mathbf{K}(Z)$  are defined by the two-forms

$$\left. \begin{aligned} \omega &= \frac{1}{3}dh \wedge du + d\eta \wedge d\phi, \\ \kappa &= ud\eta \wedge d\phi + (1 + \eta)du \wedge d\phi - \frac{1}{3}dv \wedge d\eta + dr \wedge d\gamma, \end{aligned} \right\} \tag{8.3}$$

and the functional  $S(Z)$  is

$$S(Z) = r(1 + \eta) + ph - \frac{1}{2}(1 + \eta)^2 - \frac{1}{6}h^2 - \frac{1}{3}uv + \frac{1}{2}(1 + \eta)u^2. \tag{8.4}$$

The operator  $\mathbf{K}(Z)$  can be transformed to a constant skew-symmetric operator by introducing the new variable  $q = (1 + \eta)u$  instead of  $u$ . This system is in fact equivariant with respect to a two-parameter affine Lie group (similar to that for the Boussinesq system in § 7). Linearization of (8.2) about the two-parameter family of relative uniform states associated with this group leads to a spectrum of the asymptotic systems at infinity (cf. § 3) as shown in figure 7a, b.

However, since the  $\min\{\dim E_{\pm\infty}^u(\lambda), \dim E_{\pm\infty}^s(\lambda)\} = 2$ , a matrix-based symplectic Evans function is needed for proving an instability criterion and this theory is developed in Bridges & Derks (1999). The full details of the instability theory for solitary-wave solutions of (8.1) will be given elsewhere.

Another intriguing consequence of the framework presented here is the application to the linear stability of solitary waves or fronts which can be characterized as heteroclinic orbits connecting periodic or toral invariant manifolds at infinity. For example, a sketch of how the theory extends to the case where the solitary-wave state is asymptotic to a (spatially) periodic state at infinity is as follows. First consider the case where the system is  $SO(2)$ -equivariant, in which case the periodic state at infinity can be associated with a manifold of spatial relative equilibria. An example is the generalized nonlinear Schrödinger equation:

$$i\psi_t = \psi_{xx} + F'(|\psi|^2)\psi, \tag{8.5}$$

where  $\psi(x, t)$  is complex valued and, for the present purposes,  $F : \mathbb{R} \rightarrow \mathbb{R}$  can be any smooth function. With  $\psi = q_1 + iq_2$  and  $\psi_x = p_1 + ip_2$ , (8.5) has the following multisymplectic representation:

$$\mathbf{M}Z_t + \mathbf{K}Z_x = \nabla S(Z), \quad Z \in \mathbb{R}^4, \quad x \in \mathbb{R}, \tag{8.6}$$

with

$$M = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad K = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad Z = \begin{pmatrix} q_1 \\ q_2 \\ p_1 \\ p_2 \end{pmatrix} \in \mathbb{R}^4, \quad (8.7)$$

and

$$S(Z) = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}F(q_1^2 + q_2^2). \quad (8.8)$$

The system (8.2) is  $SO(2)$ -equivariant with an action for  $SO(2)$  on  $\mathbb{R}^4$  defined by

$$\mathcal{G}_\theta = R_\theta \oplus R_\theta \quad \text{with } R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (8.9)$$

The theory of Appendix A applies to show the existence of a family of relative uniform states at infinity,  $Z(x) = \mathcal{G}_{\theta(x)}Z_0$ , which are in fact (spatially) periodic states at infinity. Solitary-wave states that are travelling and are asymptotic to this family of states can be characterized as in (2.9), with appropriate conditions on  $\hat{Z}(x, c)$  as  $x \rightarrow \pm\infty$ . For particular choices of the function  $F$  in (8.5), such heteroclinic orbits can be explicitly constructed. A straightforward analysis shows that  $\Delta(\mu, 0)$ , as defined in (3.9), has  $\mu$ -roots precisely as in figure 4a. However, the four  $\mu$ -roots of

$$\Delta(\mu, \lambda) = \det[B_0^\pm - \mu J_c - \lambda M] = 0 \quad \text{with } \text{Re}(\lambda) > 0$$

consist of two with positive real part and two with negative real part; that is,  $\dim E_{\pm\infty}^u(\lambda) = \dim E_{\pm\infty}^s(\lambda) = 2, \forall \lambda \in \mathbb{C}$  with  $\text{Re}(\lambda) > 0$ . Therefore a generalization of the symplectic Evans function to a matrix-based symplectic Evans function is needed and this theory is developed in Bridges & Derks (1999).

In the example (8.5), the periodic state at infinity is aligned with the group action. However, the structure of the problem is the same for general classes of (spatially) periodic states at infinity. Weinstein (1978) has shown that, on any symplectic manifold on which the loop space is well defined, periodic orbits correspond precisely to relative equilibria on the loop space, and this theory certainly applies when the phase space is  $\mathbb{R}^{2n}$ . Therefore, the framework presented here should also extend to analyse the linear stability of the large, and largely uninvestigated, class of solitary-wave states and travelling fronts that are asymptotic to spatially periodic states at plus and minus infinity.

### Appendix A. Uniform states and spatial relative equilibria

For a system of PDEs in the form

$$MZ_t + KZ_x = \nabla S(Z), \quad Z \in \mathbb{R}^{2n}, \quad (A 1)$$

a uniform state—a state which is independent of  $x$  and  $t$ —is a critical point of the functional  $S$  on  $\mathbb{R}^{2n}$ . When the system is equivariant there is a more general class of such states: relative equilibria. However, for systems of the form (A 1), relative equilibria can be defined with respect to either (or both) of the symplectic structures  $\omega$  and  $\kappa$ . For illustration, consider a one-parameter Lie group, with action  $\mathcal{G}_\theta$  on  $\mathbb{R}^{2n}$ , and suppose that the two symplectic flows generate distinct functions:

$$Mg(Z) = \nabla Q(Z) \quad \text{and} \quad Kg(Z) = \nabla P(Z) \quad \text{where } g(Z) = \frac{d}{d\theta} \mathcal{G}_\theta Z|_{\theta=0}. \quad (A 2)$$

Temporal and spatial relative equilibria take the following forms, respectively:

$$Z_1(t) = \mathcal{G}_{\theta_1(t)}U, \quad \text{with } \theta_1(t) = bt + \theta_1^0, \quad U \in \mathbb{R}^{2n}, \quad (\text{A } 3 \text{ a})$$

$$Z_2(x) = \mathcal{G}_{\theta_2(x)}V, \quad \text{with } \theta_2(x) = ax + \theta_2^0, \quad V \in \mathbb{R}^{2n}. \quad (\text{A } 3 \text{ b})$$

Substitution of these expressions into (A 1) shows that  $(U, b)$  and  $(V, a)$  satisfy

$$\nabla S(U) - b\nabla Q(U) = 0 \quad \text{with } Q(U) = Q_0, \quad (\text{A } 4)$$

$$\nabla S(V) - a\nabla P(V) = 0 \quad \text{with } P(V) = P_0, \quad (\text{A } 5)$$

where  $Q_0$  and  $P_0$  are assigned real numbers.

In fact, a combination of the two types can also be considered, where the flow along the group is of the form  $\theta(x, y) = bt + ax + \theta_0$ . However, this can also be seen as a translation action after a group action, i.e. when  $a \neq 0$  the action is  $\mathcal{T}_{bt/a}\mathcal{G}_{ax+\theta_0}$ , where  $\mathcal{T}_\tau$  is defined in (2.6). Each type of relative equilibrium will have a different implication for the system (A 1). In this appendix, the characterization of relative equilibria associated with the  $\kappa$ -symplectic structure is developed. We will call this type of relative equilibrium a *spatial relative equilibrium*.

The spatial relative equilibria provide a natural setting for characterizing multiparameter uniform states and the case where the states at infinity are invariant manifolds (for example, periodic orbits). Since a spatial relative equilibrium is time independent, the starting point is the time-independent part of (A 1):

$$\mathbf{K}Z_x = \nabla S(Z), \quad Z \in \mathbb{R}^{2n}, \quad (\text{A } 6)$$

with the assumption that (A 6) is equivariant with respect to a  $q$ -parameter abelian Lie group ( $q < n$ ). For characterizing relative equilibria, it is not necessary to assume non-degeneracy of  $\mathbf{K}$ , but non-degeneracy of  $\mathbf{K}$  will be required for the geometric analysis of the spectrum of the linearization about a spatial relative equilibrium (cf. proposition A 1).

Characterization of relative equilibria of (A 6) now follows the standard theory for Hamiltonian ODEs (cf. Abraham & Marsden 1978, ch. 4) with minor modification for the case when  $\mathbf{K}$  is degenerate. By symplectic Noether theory, there exist functionals  $P_1(Z), \dots, P_q(Z)$  satisfying

$$\mathbf{K}\mathfrak{g}_j(Z) = \nabla P_j(Z), \quad j = 1, \dots, q, \quad \text{where } \mathfrak{g}_j(Z) = \frac{\partial}{\partial \theta_j}\mathcal{G}_\theta Z|_{\theta=0}. \quad (\text{A } 7)$$

A  $q$ -parameter relative equilibrium of (A 6) is a state of the form

$$Z(x) = \mathcal{G}_{\theta(x)}Z_0, \quad \text{with } \theta(x) = (\theta_1(x), \dots, \theta_q(x)), \quad (\text{A } 8 \text{ a})$$

where

$$\theta_j(x) = a_jx + \theta_j^0, \quad \mathbf{a} = (a_1, \dots, a_q) \in \mathbb{R}^q \quad \text{and} \quad Z_0 \in \mathbb{R}^{2n}. \quad (\text{A } 8 \text{ b})$$

Substitution of (A 8) into (A 6), and use of the equivariance of (A 6) and the identities (A 7), leads to

$$\nabla S(Z_0) - \sum_{j=1}^q a_j \nabla P_j(Z_0) = 0, \quad P_j(Z_0) = \mathcal{P}_j, \quad j = 1, \dots, q, \quad (\text{A } 9)$$

where each  $\mathcal{P}_j$  is an assigned real number. In other words,  $Z_0 \in \mathbb{R}^{2n}$  can be characterized as a critical point of  $S$  restricted to level sets of the  $q$  functionals  $P_1, \dots, P_q$

with  $\mathbf{a} \in \mathbb{R}^q$  as Lagrange multipliers and (A 9) the Lagrange necessary condition. It follows from the standard theory of Lagrange multipliers that such a critical point is non-degenerate precisely when

$$j(\mathbf{a}) = \det \begin{bmatrix} \frac{\partial \mathcal{P}_1}{\partial a_1} & \cdots & \frac{\partial \mathcal{P}_q}{\partial a_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathcal{P}_1}{\partial a_q} & \cdots & \frac{\partial \mathcal{P}_q}{\partial a_q} \end{bmatrix} \neq 0. \quad (\text{A } 10)$$

An interesting consequence of the above characterization of uniform states is that it carries information about the existence of spatially hyperbolic real eigenvalues of (A 6), when linearized about a spatial relative equilibrium.

**Proposition A 1.** *Let*

$$\mathbf{B}_0 = D^2 S(Z_0) - \sum_{j=1}^q a_j D^2 P_j(Z_0), \quad (\text{A } 11)$$

and suppose that  $\mathbf{B}_0$  has exactly  $q$  zero eigenvalues and  $\mathbf{K}$  is non-degenerate. Let  $\varrho$  be the number of negative eigenvalues of  $\mathbf{B}_0$  (counting multiplicities). If  $j(\mathbf{a}) \neq 0$  and

$$(-1)^\varrho j(\mathbf{a}) < 0, \quad (\text{A } 12)$$

then the linearization of (A 6) about the state (A 8) has a pair of hyperbolic real eigenvalues.

*Proof.* Let  $Z(x) = \mathcal{G}_{\theta(x)}(Z_0 + \hat{U}(x))$ . Then substitution into (A 6) and linearization about  $Z_0$  leads to

$$\mathbf{K}\hat{U}_x = \mathbf{B}_0\hat{U},$$

with  $\mathbf{B}_0$  defined in (A 11). The spectral ansatz  $\hat{U}(x) = e^{\mu x}U$  leads to the generalized eigenvalue problem:  $\mathbf{B}_0U = \mu\mathbf{K}U$  for the spatial eigenvalue  $\mu \in \mathbb{C}$ . Therefore define the spectral function

$$\delta(\mu) = \det[\mathbf{B}_0 - \mu\mathbf{K}]. \quad (\text{A } 13)$$

The claim is that the condition (A 12) implies the existence of a pair of real roots—one strictly positive and one strictly negative—of  $\delta(\mu) = 0$ . The remainder of the proof then follows from standard results on relative equilibria (cf. Grillakis *et al.* 1990; Maddocks & Sachs 1993, and references therein) adapted to the spatial setting—indeed, the proof here is possible using linear algebra, and so just a sketch is given. First note that  $\delta(-\mu) = \delta(\mu)$  and

$$\delta(\mu) = \det(\mathbf{K})\mu^{2n} + \text{terms of lower degree},$$

where  $\det(\mathbf{K}) > 0$ . Therefore  $\delta(\mu) > 0$  for  $\mu \in \mathbb{R}$  and sufficiently large. The remainder of the proof reduces to showing that  $\delta(\mu)$  has the form  $\delta(\mu) = \mu^{2\varrho}\hat{\delta}(\mu)$  with  $\text{sgn}(\hat{\delta}(0)) = (-1)^\varrho \text{sgn}(j(\mathbf{a}))$ . The result then follows from the intermediate-value theorem. ■

**Corollary A 2.** *Suppose  $q + 1 = n$ ,  $\mathbf{B}_0$  has exactly  $q$  zero eigenvalues,  $\mathbf{K}$  is non-degenerate and (A 12) is satisfied. Then there exists exactly one pair of non-zero real eigenvalues of opposite sign of the linearization of (A 6) about a spatial relative equilibrium  $\mathcal{G}_{\theta(x)}Z_0$ .*

*Proof.* When  $q+1 = n$  and the rank of  $\mathbf{K}$  is  $2n$ , it follows from the proposition A 1 that

$$\delta(\mu) = \mu^{2q}(\det(\mathbf{K})\mu^2 + \delta_1), \quad \text{with } \text{sgn}(\delta_1) = (-1)^q \text{sgn } j(\mathbf{a}). \quad (\text{A } 14)$$

Hence  $\delta(\mu)$  has  $2q$  zero roots and the pair of real non-zero roots  $\pm(-\delta_1/\det(\mathbf{K}))^{1/2}$ . ■

Note that the characterization of relative equilibria (A 6)–(A 9) does not require non-degeneracy of  $\mathbf{K}$  but it is required in proposition A 1. The reason for this is apparent from (A 14). If the rank of  $\mathbf{K}$  equals  $2q$ , and the dimension of the group  $G$  is  $q$ , then there are exactly  $2q$  zero roots of (A 13) and no others. An example of this phenomenon is the Boussinesq equation in § 7, where the hyperbolic real eigenvalues disappear when  $c = 0$ , since the rank( $\mathbf{K} - c\mathbf{M}$ ) is six if  $c \neq 0$  and four ( $= 2q$ ) if  $c = 0$ .

### Appendix B. Hodge duality and adjoint systems

Consider the system of ODEs

$$U_x = \mathbf{A}(x, \lambda)U, \quad U \in \mathbb{C}^\ell, \quad x \in \mathbb{R}, \quad (\text{B } 1)$$

where  $\lambda \in \Lambda \subset \mathbb{C}$  and  $\ell \geq 2$  is a natural number. In this appendix, the connection between the adjoint equation associated with (B 1) and the restriction of the equation to  $\wedge^{\ell-1}(\mathbb{C}^\ell)$  is established.

On  $\mathbb{R}^\ell$  with inner product  $\langle \cdot, \cdot \rangle$ , let  $e_1, \dots, e_\ell$  be an orthonormal basis and define an orientation on  $\mathbb{R}^\ell$  by choosing the natural volume form:  $\mathcal{V} = e_1 \wedge \dots \wedge e_\ell$ . Identify  $\wedge^1(\mathbb{R}^\ell)$  with  $\mathbb{R}^\ell$  with the same basis. A natural basis for  $\wedge^{\ell-1}(\mathbb{R}^\ell)$  can be defined using the Hodge star operator (cf. Wells 1979, ch. V). For any  $e_j \in \wedge^1(\mathbb{R}^\ell)$  the Hodge star of  $e_j$ , denoted  $\star e_j$ , satisfies  $e_j \wedge \star e_j = \mathcal{V}$ . Therefore, a natural basis for  $\wedge^{\ell-1}(\mathbb{R}^\ell)$  is  $\theta_1, \dots, \theta_\ell$  with

$$\theta_i = \star e_i = \begin{cases} e_2 \wedge \dots \wedge e_\ell, & i = 1, \\ \varepsilon_i e_{i-1} \wedge \dots \wedge e_\ell \wedge e_1 \wedge \dots \wedge e_{i-1}, & i = 2, \dots, \ell - 1, \\ (-1)^{\ell-1} e_1 \wedge \dots \wedge e_{\ell-1}, & i = \ell, \end{cases} \quad (\text{B } 2)$$

where  $\varepsilon_i$  is the sign of the permutation taking  $\{1, \dots, \ell\}$  to  $\{i+1, \dots, \ell, 1, \dots, i-1\}$ . By using the complex field, the basis  $e_1, \dots, e_\ell$  is also a basis for  $\mathbb{C}^\ell$  and  $\theta_1, \dots, \theta_\ell$  is a basis for  $\wedge^{\ell-1}(\mathbb{C}^\ell)$ .

Let  $\xi_1(x, \lambda), \dots, \xi_\ell(x, \lambda)$  be any  $\ell$  solutions of (B 1) and define

$$W(x, \lambda) = \xi_2 \wedge \dots \wedge \xi_\ell \in \wedge^{\ell-1}(\mathbb{C}^\ell). \quad (\text{B } 3)$$

Using the basis for  $\wedge^{\ell-1}(\mathbb{C}^\ell)$ , expand  $W$ :

$$W(x, \lambda) = \sum_{j=1}^{\ell} w_j(x, \lambda)\theta_j \quad \text{with } w_j \in \mathbb{C}. \quad (\text{B } 4)$$

Similarly,  $\xi_1(x, \lambda)$ , considered as an element of  $\wedge^1(\mathbb{C}^\ell)$ , can be expanded as

$$\xi_1(x, \lambda) = \sum_{j=1}^{\ell} a_j(x, \lambda)e_j$$

with each  $a_j$  complex valued. Using the above expansions,

$$\begin{aligned}\xi_1 \wedge W &= \left( \sum_{j=1}^{\ell} a_j e_j \right) \wedge \left( \sum_{k=1}^{\ell} w_k \theta_k \right) = \sum_{j=1}^{\ell} \sum_{k=1}^{\ell} a_j w_k e_j \wedge \theta_k \\ &= \left( \sum_{j=1}^{\ell} a_j w_j \right) e_1 \wedge \cdots \wedge e_{\ell},\end{aligned}$$

since  $e_j \wedge \theta_k = \mathcal{V}$  if  $j = k$  and zero when  $j \neq k$ . Now,  $\star W \in \wedge^1(\mathbb{C}^{\ell})$  and so

$$\begin{aligned}\xi_1 \wedge W &= (-1)^{\ell-1} W \wedge \xi_1 \quad \text{with } W \wedge \xi_1 = \langle \star W, \xi_1 \rangle_{\mathbb{C}} \mathcal{V}; \\ \text{hence } \star W &= (-1)^{\ell-1} \sum_{j=1}^{\ell} \bar{w}_j e_j,\end{aligned}\tag{B5}$$

noting that  $\star \theta_j = (-1)^{\ell-1} e_j$ .

**Proposition B 1.** *Let  $\xi_1, \dots, \xi_{\ell}$  be any  $\ell$  solutions of (B 1) and let  $W = \xi_2 \wedge \cdots \wedge \xi_{\ell} \in \wedge^{\ell-1}(\mathbb{C}^{\ell})$ . Then  $\star W \in \wedge^1(\mathbb{C}^{\ell})$  satisfies*

$$\frac{d}{dx}(\star W) = [\overline{\tau(x, \lambda)} - \mathbf{A}(x, \bar{\lambda})^T](\star W),$$

where  $\tau(x, \lambda) = \text{Tr}(\mathbf{A}(x, \lambda))$  and the superscript ‘T’ corresponds to matrix transpose.

*Proof.* Differentiate both sides of the expression for  $W \wedge \xi_1$  in (B 5) to obtain

$$\frac{d}{dx}(W \wedge \xi_1) = \left[ \left\langle \frac{d}{dx}(\star W), \xi_1 \right\rangle_{\mathbb{C}} + \left\langle \star W, \frac{d}{dx} \xi_1 \right\rangle_{\mathbb{C}} \right] \mathcal{V}.\tag{B6}$$

The left-hand side can be simplified using the Abel–Liouville theorem and (B 5):

$$\frac{d}{dx}(W \wedge \xi_1) = \tau(x, \lambda)(W \wedge \xi_1) = \tau(x, \lambda) \langle \star W, \xi_1 \rangle_{\mathbb{C}} \mathcal{V}.$$

Therefore (B 6) becomes

$$\left[ \left\langle \frac{d}{dx}(\star W), \xi_1 \right\rangle_{\mathbb{C}} + \left\langle \star W, \frac{d}{dx} \xi_1 - \tau(x, \lambda) \xi_1 \right\rangle_{\mathbb{C}} \right] \mathcal{V} = 0,$$

or, with  $(\xi_1)_x = \mathbf{A}(x, \lambda) \xi_1$ ,

$$\left\langle \xi_1, [\overline{\tau(x, \lambda)} \mathbf{I} + \mathbf{A}(x, \bar{\lambda})^T](\star W) + \frac{d}{dx}(\star W) \right\rangle_{\mathbb{C}} = 0.\tag{B7}$$

Since the Hodge star is independent of the orthonormal basis chosen in the orientation (cf. Marcus & Robinson 1975) and (B 7) must hold for all  $\xi_1$ , this gives the equation for  $\star W$  on  $\wedge^1(\mathbb{C}^{\ell})$  and the proof is complete. ■

### Appendix C. Evans function for the wave equation in § 5

For the nonlinear wave equation in § 5 with  $f(u) = \alpha(u - u^2)$ , the symplectic Evans function for the various regions can be computed explicitly, by decoupling the spectral problem into two simpler second-order ODEs each of which can be solved explicitly. To fix parameter values, take  $c^2 > 1$  and  $\Lambda = \mathbb{C}_+ \cap \mathbb{E}$  as defined in equation (5.7).



The vector-valued functions  $\xi(x, \lambda)$  in (4.2) and  $\eta(x, \lambda)$  in (4.6) are constructed as follows.

Let  $\text{th}(x) = \tanh \frac{1}{2}\vartheta x$  and define

$$h_{\pm}(x, \lambda) = \mp \frac{1}{15}r(4 - r^2) - \frac{1}{5}(3 - 2r^2) \text{th}(x) \pm r \text{th}(x)^2 + \text{th}(x)^3,$$

where

$$r = 2 \left( 1 + \frac{\lambda^2 \vartheta^2}{\alpha^2} \right)^{1/2}.$$

Let

$$h_{\infty}(\lambda) = \lim_{x \rightarrow +\infty} h_+(x, \lambda) = - \lim_{x \rightarrow -\infty} h_-(x, \lambda) = \frac{2}{5}(1 + r^2) + \frac{1}{15}r(11 + r^2).$$

Then

$$\xi(x, \lambda) = \frac{2\vartheta}{|\alpha|r h_{\infty}(\lambda)} e^{-\beta(c, \lambda)x} \begin{pmatrix} h_+(x, \lambda) \\ (\lambda + c\beta)h_+(x, \lambda) - c \frac{d}{dx} h_+(x, \lambda) \\ \beta h_+(x, \lambda) - \frac{d}{dx} h_+(x, \lambda) \\ 0 \end{pmatrix},$$

$$\eta(x, \lambda) = \frac{1}{2h_{\infty}(\lambda)} e^{+\beta(c, \lambda)x} \begin{pmatrix} h_-(x, \lambda) \\ -(\lambda + c\beta)h_-(x, \lambda) - c \frac{d}{dx} h_-(x, \lambda) \\ -\beta h_-(x, \lambda) - \frac{d}{dx} h_-(x, \lambda) \\ 0 \end{pmatrix},$$

where

$$\beta(c, \lambda) = \frac{\vartheta^2}{\alpha} (c\lambda - \sqrt{\lambda^2 + \alpha(1 - c^2)}), \quad \text{for } \lambda \in \mathbb{C}^+ \cap \mathbb{E}, \quad \text{with } \lim_{\lambda \rightarrow 0} \beta(c, \lambda) = \vartheta.$$

The vector-valued functions  $\xi(x, \lambda)$  and  $\eta(x, \lambda)$  satisfy the asymptotic conditions

$$\lim_{x \rightarrow +\infty} e^{\beta(c, \lambda)x} \xi(x, \lambda) = \zeta_+(c, \lambda) = \frac{2\vartheta}{|\alpha|r} \begin{pmatrix} 1 \\ \lambda + \beta c \\ \beta \\ 0 \end{pmatrix},$$

$$\lim_{x \rightarrow -\infty} e^{-\beta(c, \lambda)x} \eta(x, \lambda) = \zeta_-(c, \lambda) = -\frac{1}{2} \begin{pmatrix} 1 \\ -\lambda - \beta c \\ -\beta \\ 0 \end{pmatrix}$$

and  $\Omega(\zeta_-, \zeta_+) = 1$ . The symplectic Evans function is

$$D(\lambda) = \Omega(\eta, \xi) = \left( \frac{8\vartheta^3}{15|\alpha|^3 h_{\infty}(\lambda)} \right)^2 \lambda^2 \left( \lambda^2 + \frac{3}{4} \frac{\alpha^2}{\vartheta^2} \right) \left( \lambda^2 - \frac{5}{4} \frac{\alpha^2}{\vartheta^2} \right).$$

The unstable eigenvalue in  $\mathbb{C}_+ \cap \mathbb{E}$  is then  $\lambda_u = \frac{1}{2}\sqrt{5}|\alpha|/\vartheta$ , with the condition that  $\lambda_u < |\alpha|^{1/2}$  which implies  $1 < c^2 < 1 + \frac{4}{5}$ . The existence of an unstable eigenvalue for the case  $c^2 < 1$  can be deduced using the theory in Bridges & Derks (1999).

However, for the form of the nonlinearity in §5, the Evans function for this case can also be constructed explicitly, by constructing a basis  $\xi_1(x, \lambda), \dots, \xi_4(x, \lambda)$  for  $U_x = \mathbf{A}(x, \lambda)U$  and using the definition  $\tilde{D}(\lambda) = e^{\tau\lambda x} \xi_1 \wedge \dots \wedge \xi_4$  for the Evans function. Such a construction shows the existence of an unstable eigenvalue for all  $c^2 < 1$ ; the details are not important for the argument in §5 and are therefore omitted.

## References

- Abraham, R. & Marsden, J. E. 1978 *Foundations of mechanics*, 2nd edn. New York: Benjamin.
- Alexander, J., Gardner, R. & Jones, C. K. R. T. 1990 A topological invariant arising in the stability analysis of traveling waves. *J. Reine Angew. Math.* **410**, 167–212.
- Benjamin, T. B. 1972 The stability of solitary waves. *Proc. R. Soc. Lond. A* **328**, 153–183.
- Binz, E., Śniatycki, J. & Fischer, H. 1988 *Geometry of classical fields*. Amsterdam: North Holland.
- Bona, J. L. 1975 On the stability of solitary waves. *Proc. R. Soc. Lond. A* **344**, 363–374.
- Bona, J. L. & Smith, R. 1976 A model for two-way propagation of water waves in a channel. *Math. Proc. Camb. Phil. Soc.* **79**, 167–182.
- Bona, J. L., Souganidis, P. E. & Strauss, W. A. 1987 Stability and instability of solitary waves of KdV type. *Proc. R. Soc. Lond. A* **411**, 395–411.
- Bridges, T. J. 1995 Multisymplectic structures, Boussinesq equations and periodic travelling waves. *Proc. IUTAM Symp. on Nonlinear Waves in Fluids* (ed. A. Mielke & K. Kirchgässner). Singapore: World Scientific.
- Bridges, T. J. 1997a Multisymplectic structures and wave propagation. *Math. Proc. Camb. Phil. Soc.* **121**, 147–190.
- Bridges, T. J. 1997b A geometric formulation of the conservation of wave action and its implications for signature and the classification of instabilities. *Proc. R. Soc. Lond. A* **453**, 1365–1395.
- Bridges, T. J. & Derks, G. 1999 The symplectic Evans matrix, and the instability of solitary waves and fronts with symmetry. Preprint.
- Coppel, W. A. 1965 *Stability and asymptotic behavior of differential equations*. Boston, MA: D. C. Heath & Co.
- Eastham, M. S. P. 1989 *The asymptotic solution of linear differential systems*. Oxford: Clarendon.
- Evans, J. W. 1975 Nerve axon equations. IV. The stable and unstable impulse. *Indiana Univ. Math. Jl* **24**, 1169–1190.
- Gardner, R. A. 1997 Spectral analysis of long wavelength periodic waves with applications to the generalized KdV equation. *J. Reine Angew. Math.* **491**, 149–181.
- Grillakis, M., Shatah, J. & Strauss, W. 1987 Stability theory of solitary waves in the presence of symmetry. I. *J. Func. Analyt.* **74**, 160–197.
- Grillakis, M., Shatah, J. & Strauss, W. 1990 Stability theory of solitary waves in the presence of symmetry. II. *J. Func. Analyt.* **94**, 308–348.
- Holm, D. D., Marsden, J. E., Ratiu, T. S. & Weinstein, A. 1985 Nonlinear stability of fluid and plasma equilibria. *Phys. Rep.* **123**, 1–116.
- Jones, C. K. R. T. 1984 Stability of the travelling wave solution of the Fitzhugh–Nagumo system. *Trans. Am. Math. Soc.* **286**, 431–469.
- Maddocks, J. H. & Sachs, R. L. 1993 On the stability of KdV multisolitons. *Commun. Pure Appl. Math.* **46**, 867–901.
- Marcus, M. & Robinson, H. 1975 A note on the Hodge star operator. *Lin. Algeb. Appl.* **10**, 85–87.

- Marsden, J. E. & Shkoller, S. 1999 Multisymplectic geometry, covariant Hamiltonians and water waves. *Math. Proc. Camb. Phil. Soc.* **125**, 553–575.
- Olver, P. J. 1986 Applications of Lie groups to differential equations. *Graduate texts in mathematics*, vol. 107. New York: Springer.
- Pego, R. L. & Weinstein, M. I. 1992 Eigenvalues, and instabilities of solitary waves. *Phil. Trans. R. Soc. Lond. A* **340**, 47–94.
- Pego, R. L. & Weinstein, M. I. 1997 Convective linear stability of solitary waves for Boussinesq equations. *Stud. Appl. Math.* **99**, 311–375.
- Sewell, M. J. & Porter, D. 1980 Constitutive surfaces in fluid mechanics. *Math. Proc. Camb. Phil. Soc.* **88**, 517–546.
- Toland, J. F. 1981 Solitary-wave solution for a model of the two-way propagation of water waves in a channel. *Math. Proc. Camb. Phil. Soc.* **90**, 343–360.
- Weinstein, A. 1978 Bifurcations and Hamilton's principle. *Math. Z.* **159**, 235–248.
- Wells, R. O. 1979 Differential analysis on complex manifolds. *Graduate texts in mathematics*, vol. 65. New York: Springer.
- Whitham, G. B. 1974 *Linear and nonlinear waves*. New York: Wiley-Interscience.

—

—

—

—

|

|