Symplectic transversality and the Pego-Weinstein theory

Timothy J. Burchell and Thomas J. Bridges

Abstract. This paper studies the linear stability problem for solitary wave solutions of Hamiltonian PDEs. The linear stability problem is formulated in terms of the Evans function, a complex analytic function denoted by $D(\lambda)$, where $\lambda$ is the spectral parameter. The main result is the introduction of a new factor, denoted $\Pi$, in the Pego & Weinstein (1992) derivative formula

$$D''(0) = \chi \Pi \frac{dI}{dc},$$

where $I$ is the momentum of the solitary wave and $c$ is the speed. Moreover this factor turns out to be related to transversality of the solitary wave, modelled as a homoclinic orbit: the homoclinic orbit is transversely constructed if and only if $\Pi \neq 0$. The sign of $\Pi$ is a symplectic invariant, an intrinsic property of the solitary wave, and is a key new factor affecting the linear stability. The factor $\chi$ was already introduced by Bridges & Derks (1999) and is based on the asymptotics of the solitary wave. A supporting result is the introduction of a new abstract class of Hamiltonian PDEs built on a nonlinear Dirac-type equation, which model a wide range of PDEs in applications. Examples where the theory applies, other than Dirac operators, are the coupled mode equation in fluid mechanics and optics, the massive Thirring model, and coupled nonlinear wave equations. A calculation of $D''(0)$ for solitary wave solutions of the latter class is included to illustrate the theory.

1 Introduction

The stability of solitary waves of Hamiltonian partial differential equations (PDEs) can be approached many ways. One strategy is to use the calculus of variations, since solitary waves can be characterized as critical points of the energy restricted to level sets of the momentum, and show that the solitary wave is a minimizer on the constraint set, concluding, with some additional analysis, Lyapunov (nonlinear, orbital) stability. For dispersive nonlinear wave equations, like the Korteweg de-Vries and nonlinear Schrödinger equations, this approach goes back to Benjamin [5], Bona [6], Weinstein [42] and many others (see the book by Pava [37] for history and many references). This strategy was developed into a general and powerful approach for a class of abstract Hamiltonian PDEs with one or more constraints by Grillakis, Shatah, & Strauss [25, 20] (hereafter GSS). There has been a vast amount of work in this direction since, too much to review here (e.g. see Chapter 5 in Kapitula & Promislov [30] and references therein).

From the perspective of this paper, a key part of the GSS theory is the connection between the sign of the derivative of a scalar-valued function and minimization. When there is a single constraint set, say the momentum denoted by $I$, and a single Lagrange multiplier, the speed of

1Email: T.Burchell@surrey.ac.uk
2Email: T.Bridges@surrey.ac.uk
the solitary wave denoted by $c$, the condition is
\[ \frac{dI}{dc} > 0 \quad \Rightarrow \quad \text{solitary wave is a minimizer.} \quad (1.1) \]

This condition is useful since $I(c)$ is a property of the basic state and so is, in principle, available and easy to calculate.

On the other hand a central hypothesis in the GSS theory, required for $(1.1)$, is that the second variation of the functional, in the case of one constraint, should have at most one negative eigenvalue, one zero eigenvalue, and the remainder of the spectrum strictly positive. It is this GSS spectral hypothesis that is most difficult to satisfy, and indeed may not be satisfied, especially for coupled PDEs.

Another approach is to study the spectral problem associated with the linearization about solitary waves while incorporating the Hamiltonian structure. The seminal paper in this direction is PEGO & WEINSTEIN [38] (hereafter PW). They looked to retain the derivative of $I$ in $(1.1)$ and find its role in the linear stability problem, but work around the GSS spectral hypothesis. By combining the Evans function, $D(\lambda)$, with the Hamiltonian structure, and the energy-momentum characterization of solitary waves, PW proved that it has the following properties at $\lambda = 0$,
\[ D(0) = 0, \quad D'(0) = 0, \quad D''(0) = \frac{dI}{dc}. \quad (1.2) \]

The derivative of $I$ in $(1.1)$ appears in this formula in a natural way, but the GSS spectral condition is not used in the proof. This result is useful as it is straightforward, when the evolution equation is well-posed, to normalize the Evans function so that it satisfies $D(\lambda) \to 1$ as $\lambda \to +\infty$ along the real axis. Hence when $dI/dc < 0$ the existence of an unstable stability exponent is assured by the intermediate value theorem. The theory was applied to scalar-valued PDEs such as generalized KdV, BBM equation, and Boussinesq equation, and in all cases the formula $(1.2)$ was applicable.

BRIDGES & DERKS [8, 9, 10, 11] extended the PW theory and showed that, in the presence of symmetry, there may be an additional factor in the second derivative in $(1.2)$
\[ D''(0) = \chi \frac{dI}{dc}. \quad (1.3) \]

When $D(\lambda) \to 1$ as $\lambda \to +\infty$ along the real axis it is the negativity of the full product that gives existence of an unstable eigenvalue. The factor $\chi$ is calculated independently of the derivative of $I$ and is not just a scale factor. An explicit formula was found for the factor $\chi$ but the presence of symmetry, other than translation invariance in space, was an essential part of the proof in [8, 10, 11]. Moreover the theory relied on the “system at infinity” having only one positive and one negative real (spatial) eigenvalue when $\lambda = 0$ (see §4 and §5 for the definition of “spatial eigenvalue” and “system at infinity”). Several examples were given with $\chi$ taking both positive and negative values, showing that an additional factor is essential in general.

In this paper the assumptions of additional symmetry and one-dimensional stable manifold in the system at infinity are removed. A new formula for the second derivative is found in the form
\[ D''(0) = \chi \Pi \frac{dI}{dc}. \quad (1.4) \]

The factor $\Pi$ is associated with the transversal intersection of the stable and unstable manifolds which form the solitary wave, characterized as a homoclinic orbit. In the case of a two-dimensional
stable and unstable manifold, the new factor is
\[ \Pi = \Omega(a^-, a^+) , \] (1.5)

where \( a^- \) and \( a^+ \) are \( \xi \)-dependent tangent vectors to the stable and unstable manifolds respectively and \( \Omega \) is a symplectic form associated with a \( c \)-dependent spatial symplectic structure (defined in §4). When the dimension of the stable and unstable manifolds is greater than two the formula (1.5) expands to be the determinant of a matrix of symplectic forms (see §9 and [11, 16]). The importance of \( \Omega(a^-, a^+) \) as a symplectic invariant of homoclinic orbits was discovered by Lazutkin [24], and hence we call it the Lazutkin invariant, and its properties and connection with the parity of the Maslov index are proved by Chardard & Bridges [16].

There are a number of hypotheses that go into the result (1.4) but the most important are firstly that no symmetry (other than translation invariance) is assumed, and secondly the system at infinity is not restricted to one (spatial) eigenvalue with positive real part in the limit \( \lambda \to 0 \). In the body of the paper we will restrict attention to the case where the homoclinic orbit is the intersection between a two-dimensional stable and two-dimensional unstable manifold, with the generalization to arbitrary dimension discussed in §9.

The role of transversality in the Evans function formulation of the linear stability problem for solitary waves here is new but not that surprising. In the case of dissipative PDEs, Alexander & Jones [2] prove that the first derivative of the Evans function can be characterized in terms of a coefficient of transversality. This theory is abstracted to an orientation index in §9.4 of Kapitula & Promislow [30]. Chardard & Bridges [16] prove that in gradient systems, with Hamiltonian steady part, the first derivative of the Evans function can be expressed in terms of transversality. However, in all previous work transversality shows up in the first derivative, and here symplecticity in the time direction comes into play, and it shows up in the second derivative and includes the additional factors \( \chi \) and the momentum function \( dI/dc \).

In order to give the result (1.4) some generality we need an abstract class of Hamiltonian PDEs. By way of comparison, the class of Hamiltonian PDEs in PW [38] is
\[ u_t = J \nabla_u H(u, u_x), \quad u \in X, \] (1.6)

for some function space \( X \), where \( J: X^* \to X \) is the co-symplectic (or Poisson) operator, \( u \) is scalar-valued, and \( H: X \to \mathbb{R} \) is the Hamiltonian function. However, reduction of (1.6) to a steady problem is an ODE,
\[ cu_\xi = \nabla_u H(u, u_\xi), \quad \xi = x + ct. \]

With modest hypotheses on \( J \) this equation is generated by a Lagrangian with density \( L(u, u_\xi) \). Invoking a Legendre transform then brings in a second hidden symplectic structure and a finite-dimensional Hamiltonian system. Denote the second “spatial symplectic operator” by \( K \). This spatial symplectic structure is essential for both defining symplectic transversality and for the proof of the formula (1.3).

It is clear that the interplay between two symplectic structures is an essential part of the analysis: the time evolution and the energy-momentum characterization of the solitary wave use the temporal symplectic structure, whereas transversality of the homoclinic orbit representation of the solitary wave is defined using the spatial symplectic structure. The Evans function is defined using both symplectic structures. Hence, introducing a finite-dimensional representation of \( J \), and new coordinates, leads to a formulation of the Hamiltonian PDE in terms of multisymplectic
structure (e.g. [8, 10]; canonical structures and history are given in [7, 12]). The canonical form for a multisymplectic Hamiltonian PDE of interest here is

\[ M \dot{Z} + K Z_x = \nabla S(Z), \quad Z \in \mathbb{R}^{2n}, \tag{1.7} \]

where \( M \) and \( K \) are symplectic operators which are taken to be constant and \( S \) is a generalized Hamiltonian function with \( M \) a finite dimensional representation on the phase space \( \mathbb{R}^{2n} \) of the infinite-dimensional operator \( J \) in (1.6). Steady solutions \( Z(x,t) = \hat{Z}(\xi), \xi = x + ct, \) are orbits of the finite-dimensional Hamiltonian system

\[ (K + cM) \hat{Z}_\xi = \nabla S(\hat{Z}), \quad \hat{Z} \in \mathbb{R}^{2n}. \tag{1.8} \]

In this system a solitary wave is represented by a homoclinic orbit. The theory will be developed for the case \( n = 2 \) which is the lowest dimension of interest, and limits the proliferation of indices, with comments on the general \( n > 2 \) case in the concluding remarks. The abstract form (1.7) is quite satisfactory for the theory and represents a wide range of Hamiltonian PDEs [7, 8, 12, 10, 11]. However, we go one step further in this paper and introduce an abstract class of multisymplectic Hamiltonian PDEs. Given a smooth pseudo-Riemannian manifold there is a natural form on the total exterior algebra bundle whose variation produces a coordinate-free version of the left-hand side of (1.7). This construction generalizes the symplectic structure on the cotangent bundle of a Riemannian manifold in classical mechanics. With this strategy we get a coordinate-free formulation as well as the canonical form (1.7). In fact the partial differential operator generated is a Dirac operator. It is made nonlinear by adding a gradient on the right-hand side. We call the class of PDEs generated on the total exterior algebra bundle multisymplectic Dirac operators. This class of Hamiltonian PDEs includes as special cases the coupled mode equation which appears in fluid dynamics [18, 27, 28, 29] and optics [40, 19, 3], the massive-Thirring model [36], and a class of coupled nonlinear wave equations.

Solitary waves are relative equilibria; that is, solutions of the Hamiltonian PDE that are equilibria in a moving frame of reference. Hence, in looking for a motivation for the formula (1.4), we first consider the spectral problem for relative equilibria of Hamiltonian ODEs and establish that

\[ D''(0) = (-1)^{\text{Morse}} \frac{dI}{dc}, \tag{1.9} \]

where the exponent is the Morse index of the constrained critical point problem (the number of strictly negative eigenvalues of the constrained second variation). In the context of ODEs the proof of (1.9) uses elementary linear algebra. This result ties in with the GSS theory because it contains a weak form of the GSS spectral condition. It is weak in that only the parity of the number of negative eigenvalues enters the formula.

On the other hand, solitary waves in the energy-momentum construction, may or may not have a well-defined Morse index. So the result (1.9) is not expected to generalize to solitary waves. However, using Theorem 10.1 in [16] we can go one step further and relate the new characterization of \( \Pi \) to the Maslov index of the solitary wave

\[ \text{sign}(\Pi) = (-1)^{\text{Maslov}}, \tag{1.10} \]

where in this case the Maslov index of the solitary wave is defined using the Souriau characterization (cf. §9 of [16]). Solitary waves, with exponential decay at infinity, always have a well-defined
Maslov index, but may not have a well-defined Morse index. The Maslov index will not feature in this paper, as the emphasis here is on transversality, but some elaboration of (1.10) is given in §6.

An outline of the paper is as follows. In Section 2 the special case (1.9) of the derivative formula is proved for relative equilibria of Hamiltonian ODEs. In Section 3 an abstract class of multisymplectic Hamiltonian PDEs is introduced. Section 4 is the starting point for proving the main results on stability of solitary waves. Here the abstract class of solitary waves is introduced as well as the properties of the linearization about these waves. Section 5 constructs the Evans function and develops the interplay with symplecticity. Section 6 proves the main result on $D''(0)$ confirming (1.4). Section 7 gives an example where all the details are worked out explicitly. In §8 the difficulties with the case of complex $\mu$–eigenvalues are discussed. Finally in the concluding remarks Section 9 some generalizations are pointed out.

2 Instability of relative equilibria of ODEs

A solitary wave solution of a Hamiltonian PDE is a relative equilibrium in the following sense. Focussing on the form (1.6) for description, suppose the Hamiltonian function and symplectic structure do not depend explicitly on the spatial coordinate, $x$. Then $u(x + s, t)$ is a solution for any $s$ whenever $u(x, t)$ is, and we say that the Hamiltonian PDE is equivariant with respect to the group $G = \mathbb{R}$, the group of real numbers. When $s = ct$ with $c$ a constant, and $u$ is otherwise independent of $t$, the solution is a relative equilibrium of the form $u(x, t) := \hat{u}(x + ct)$; that is, it is an equilibrium when viewed from a frame of reference moving at constant speed along the group $\mathbb{R}$. Noether theory then gives the existence of an invariant associated with the translation symmetry that is called the momentum, here denoted by $I$. It is a functional and depends on $u$, but when $I$ is evaluated on a family of relative equilibria it becomes a function of $c$ only, and it is this function that appears in the derivative formula (1.3).

The main aims of this section are to show, in the simplest possible setting, how the product structure of $D''(0)$ arises, and to show how the geometry of relative equilibria (RE) enters the analysis. Although this section is restricted to ODEs, the abstract structure is the same as that of the solitary wave stability problem with the Evans function replaced here by an elementary characteristic function. The group is simplified to the one-parameter compact group $S^1$.

Consider the finite-dimensional Hamiltonian system on $M := \mathbb{R}^{2n}$:

$$MZ_t = \nabla H(Z), \quad Z \in M, \quad M = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix},$$

(2.1)

where $H : M \to \mathbb{R}$ is a given smooth function. The system (2.1) is assumed to be equivariant with respect to $S^1$: that is, there is an orthogonal action, $G_\theta$, of $S^1$ on $M$ satisfying

$$G_\theta^T M G_\theta = M \quad \text{and} \quad H(G_\theta Z) = H(Z) \quad \forall \theta \in S^1.$$  

(2.2)

In this section $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{R}^{2n}$.

**Proposition 2.1.** Let $g(Z) = \frac{d}{d\theta} G_\theta Z \big|_{\theta=0}$ for $Z \in M$, then there exists (symplectic Noether theory) a functional $I : M \to \mathbb{R}$ satisfying $Mg(Z) = \nabla I(Z)$.

**Proof.** In the Hamiltonian setting, Noether theory states that contraction of the symplectic form by the generator $g(Z)$ is the gradient of a function (e.g. §2.7 of [31]). In this case, the existence
of $I$ follows from a linear algebra argument. Since the action of the group is orthogonal, the tangent vector $g(Z)$ can be expressed as $g(Z) = SZ$ with $S^T = -S$. Now $M$ and $S$ are two skew-symmetric matrices that commute (proved by differentiating the first of (2.2) with respect to the group parameter) and so their product is symmetric. Define $I(Z) = \frac{1}{2} \langle MSZ, Z \rangle$, then

$$Mg(Z) = MSZ = \nabla I(Z),$$

(2.3)

completing the proof. ■

Now, suppose there exists a family of RE of the above system. An RE is a solution of (2.1) that is aligned with the group orbit and has constant speed (cf. Chapter 4 of [34]); that is, a solution of the form

$$\hat{Z}(t) = G_{\theta(t)}U(c) \quad \text{with} \quad \theta(t) = ct + \theta^o.$$

(2.4)

Substitution of this form into the governing equation (2.1) gives

$$0 = M\hat{Z}_t - \nabla H(\hat{Z}) = \dot{\theta}Mg(G_{\theta(t)}U) - \nabla H(G_{\theta(t)}U) = G_{\theta(t)}(c\nabla I(U) - \nabla H(U)),$$

and so $U(c)$ and $c$ are defined by the equations

$$\nabla H(U) = c\nabla I(U) \quad \text{and} \quad I(U) = I_0,$$

(2.5)

that is; $U(c) \in M$ can be characterised as a critical point of $H$ on level sets, $I = I_0$, of the functional $I$, with $c$ as a Lagrange multiplier. This family of RE is assumed to exist for all $c \in \mathcal{C}$ where $\mathcal{C}$ is some open subset of $\mathbb{R}$. A RE in the family is said to be non-degenerate on $\mathcal{C}$ when $\frac{dI}{dc} \neq 0$ where $I$ is evaluated on the family $U(c)$. This family of RE is a finite-dimensional analogue of a travelling wave.

### 2.1 Linear stability of the family of RE

The linear stability equation for the family of RE is formulated by linearizing (2.1) about (2.4)

$$MZ_t = D^2H(\hat{Z}(t))Z,$$

(2.6)

with $\hat{Z}(t)$ defined in (2.4)-(2.5). The group can be factored out by letting $Z(t) = G_{\theta(t)}W(t)$, then

$$0 = MZ_t - D^2H(\hat{Z}(t))Z = \dot{\theta}Mg(G_{\theta(t)}W) + MG_{\theta(t)}W_t - D^2H(G_{\theta(t)}U)Z = G_{\theta(t)}(cMg(W) + MW_t - G_{\theta(t)}D^2H(G_{\theta(t)}U)G_{\theta(t)}W) = G_{\theta(t)}(cD^2I(U)W + MW_t - D^2H(U)W),$$

using the first of (2.2), noting that $\nabla I(W) = D^2I(U)W$, and using the identity

$$D^2H(G_{\theta}U) = G_{\theta}D^2H(U)G_{\theta}^T,$$
obtained by differentiating the second of \((2.2)\) with respect to \(\theta\), twice. Hence, the linearized equation \((2.6)\) is equivalent to the constant coefficient ODE:

\[
MW_t = L(U, c) W, \quad \text{where} \quad L(U, c) = D^2 H(U(c)) - c D^2 I(U(c)),
\]

with associated spectral equation \(L(U, c) W = \lambda MW\). Let

\[
D(\lambda) = \det[L(U, c) - \lambda M],
\]

then we have the following sufficient condition for instability: \textit{if there exists a} \(\lambda \in \mathbb{C} \) \textit{with} \(\text{Re}(\lambda) > 0\) \textit{and} \(D(\lambda) = 0\), \textit{the RE \((2.4)\) is linearly (spectrally) unstable}. \(D(\lambda)\) is a finite-dimensional analogue of the Evans function.

The operator \(L\) has a zero eigenvalue with eigenvector \(g(U)\). To show this, act on \((2.5)\) with \(G_\theta\), use equivariance, differentiate with respect to \(\theta\), and set \(\theta = 0\),

\[
0 = \frac{d}{d\theta} \left( \nabla H(G_\theta U) - c \nabla (G_\theta U) \right) \bigg|_{\theta=0} \Rightarrow Lg(U) = 0. \tag{2.9}
\]

It is assumed that the zero eigenvalue of \(L\) is simple for \(c \in \mathcal{C}\). Differentiating the first of \((2.5)\) with respect to \(c\) and using \((2.3)\) gives the generalized eigenvector

\[
Lc = Mg(U).
\]

Normalize the length of these two vectors and define

\[
\zeta_1 = \frac{g(U)}{\|g(U)\|} \quad \text{and} \quad \zeta_2 = \frac{U_c}{\|g(U)\|}.
\]

Then they generate the Jordan chain of length two

\[
L\zeta_1 = 0 \quad \text{and} \quad L\zeta_2 = M\zeta_1. \tag{2.10}
\]

With the simple zero eigenvalue \((2.9)\) and \(dI/dc \neq 0\), the Jordan chain has length exactly two. To have length three would require solvability of

\[
L\zeta_3 = M\zeta_2. \tag{2.11}
\]

But solvability of this equation requires

\[
0 = \langle \zeta_1, M\zeta_2 \rangle = -\frac{1}{\|g(U)\|^2} \langle Mg(U), U_c \rangle = -\frac{1}{\|g(U)\|^2} \langle \nabla I(U), U_c \rangle = -\frac{1}{\|g(U)\|^2} \frac{dI}{dc}. \tag{2.12}
\]

Hence with \(dI/dc \neq 0\) this equation is not solvable.

### 2.2 Derivatives of the characteristic function

We are now in a position to prove the following finite-dimensional analogue of \((1.2)\) with second derivative \((1.4)\).

**Theorem 2.2.** Suppose a smooth family of RE exists for all \(c \in \mathcal{C}\), with \(dI/dc \neq 0\), and suppose \(L\) has a simple zero eigenvalue. Then the characteristic function \(D(\lambda)\) has the following derivatives at the origin

\[
D(0) = 0, \quad D'(0) = 0, \quad D''(0) = \mu(L) \frac{dI}{dc},
\]
where $\mu(L)$ is the product of the nonzero eigenvalues of $L$.

**Remark.** The sign of $\mu(L)$ is the parity of the number of negative eigenvalues so with a suitable scaling of $D(\lambda)$ an equivalent formula for $D''(0)$ is

$$D''(0) = (-1)^\text{Morse} \frac{dI}{dc},$$

where Morse is the Morse index of $L$, confirming (1.9) in the introduction.

**Proof.** Since $L$ has a simple zero eigenvalue

$$D(0) = \det[L] = 0.$$

Differentiate $D(\lambda)$ in (2.8) using the formula for the derivative of a determinant

$$D'(\lambda) = -\text{Tr}\left((L - \lambda M)^\# M\right) \Rightarrow D'(0) = -\text{Tr}(L^\# M), \quad (2.13)$$

where $L^\#$ is the adjugate of $L$. When $L$ has only one zero eigenvalue with unit length eigenvector $\zeta_1$ then $L^\#$ is the rank one matrix

$$L^\# = \mu(L)\zeta_1\zeta_1^T \quad \text{with} \quad \mu(L) = \prod_{j=2}^{2n} \mu_j. \quad (2.14)$$

$\mu(L)$ is the product of the nonzero eigenvalues, $\mu_j$, of $L$ (taking the zero eigenvalue to be $\mu_1$). This formula is stated and proved as Theorem 3 on page 48 of MAGNUS & NEUDECKER [33]. Substitute $L^\#$ into (2.13),

$$D'(0) = -\text{Tr}(L^\# M) = -\mu(L)\text{Tr}(\zeta_1\zeta_1^T M) = -\mu(L)\langle \zeta_1, M\zeta_1 \rangle = 0,$$

since $M$ is skew symmetric. For $D''(0)$, differentiate $D'(\lambda)$ in (2.13)

$$D''(0) = -\text{Tr}\left(\frac{d}{d\lambda}(L - \lambda M)^\# \bigg|_{\lambda=0} M\right).$$

The adjugate is defined by

$$(L - \lambda M)(L - \lambda M)^\# = (L - \lambda M)^\# (L - \lambda M) = D(\lambda)I, \quad (2.15)$$

where $I$ is the identity on $\mathbb{R}^{2n}$. Now differentiate (2.15) with respect to $\lambda$, set $\lambda$ to zero, and define

$$\dot{L} := \frac{d}{d\lambda}(L - \lambda M)^\# \bigg|_{\lambda=0}.$$

This gives the following equations for $\dot{L}$

$$LL = ML^\# \quad \text{and} \quad L^T = -\dot{L},$$

with skew-symmetry following from commutativity in (2.15). Combining (2.10), (2.14) and skew-symmetry of $L$ gives

$$\dot{L} = \mu(L)(\zeta_2\zeta_1^T - \zeta_1\zeta_2^T).$$
Substitute into $D''(0)$

$$
D''(0) = -\mu(L) \text{Tr}
\left((\zeta_2 \zeta_1^T - \zeta_1 \zeta_2^T)M\right)
= -\mu(L) \left(\langle \zeta_1, M \zeta_2 \rangle - \langle \zeta_2, M \zeta_1 \rangle\right)
= 2\mu(L) \langle \zeta_2, M \zeta_1 \rangle
= 2 \frac{\mu(L)}{\|g(U)\|^2} \frac{dI}{dc},
$$

with the last expression following from (2.3) and

$$
\frac{dI}{dc} = \langle \nabla I(U), U_c \rangle = \langle M g(U), U_c \rangle = \|g(U)\|^2 \langle M \zeta_1, \zeta_2 \rangle.
$$

Scaling $D(\lambda)$ by a positive constant then completes the proof.

**Remark.** Since (2.8) is “the Evans function” in this example, scaling $D(\lambda)$ by a positive constant should be interpreted as a trivial re-definition of it to

$$
\tilde{D}(\lambda) = \frac{1}{2} \|g(U)\|^2 D(\lambda).
$$

and it is this new Evans function $\tilde{D}(\lambda)$ that has $\tilde{D}''(0) = \mu(L)dI/dc$.

**Corollary 2.3.** When $(-1)^{\text{Morse}} \frac{dI}{dc} < 0$ the family of RE has an unstable eigenvalue.

**Proof.** The condition assures that $D(\lambda)$ is negative for $\lambda$ near zero. For large and real $\lambda$ the characteristic function has the asymptotic form

$$
D(\lambda) = (-1)^{2n} \text{det}(M) \lambda^{2n} + \cdots,
$$

and so $D(\lambda) > 0$ for $\lambda$ real, positive, and sufficiently large. By the intermediate value theorem $D(\lambda)$ has at least one positive real root.

Theorem 2.2 connects $dI/dc$ to the spectral problem and is a finite dimensional version of the derivative formula (1.4). The connection between $dI/dc$ and critical point type for relative equilibria appears in the literature from various perspectives (e.g. Maddocks & Sachs [32] and references therein), and is an elementary example of the critical point theory in infinite-dimensional spaces in GSS [25, 26].

The plan is to extrapolate Theorem 2.2 to the context of solitary waves. Before proceeding with that proof, the next section shows that the canonical form for Hamiltonian PDEs (1.7) is not only a useful representation of well known PDEs, it is also a universal class.

### 3 A class of multisymplectic Hamiltonian PDEs

The class of multisymplectic Hamiltonian PDEs (1.7) is a natural starting point for the theory. Indeed all the theory in [8, 10, 11] is based on this class of PDEs. In this section it is shown that this class of PDEs can be obtained naturally and coordinate free from a pseudo-Riemannian manifold, showing that the class is universal in addition to being practical.

This approach is a generalization of the cotangent bundle of a manifold as a natural and coordinate free generator of symplectic structure. Let $M$ be a smooth manifold, which for simplicity
is taken to be $\mathbb{R}^n$. Let $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ be local coordinates for $T^*M \cong \mathbb{R}^{2n}$. The cotangent bundle hosts a canonical one form $p \cdot dq$ with associated functional

$$\int_{t_1}^{t_2} p \cdot q_t \, dt. \tag{3.1}$$

The first variation of this functional

$$0 = \delta \int_{t_1}^{t_2} p \cdot q_t \, dt = \int_{t_1}^{t_2} (\delta p \cdot q_t + q_t \cdot \delta p_t) \, dt,$$

with fixed endpoints on variations, $\delta q(t_1) = \delta q(t_2) = 0$, generates the operator

$$J \frac{d}{dt} \text{ with } J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}.$$

Two observations about this operator are of interest: firstly, it generates a Hamiltonian system by introducing the gradient of $H(q, p)$, a given smooth function,

$$J \frac{d}{dt} \left( \begin{array}{c} q \\ p \end{array} \right) = \nabla H,$$

and secondly it is a one-dimensional “Dirac operator”

$$J \frac{d}{dt} \circ J \frac{d}{dt} = -I_{2n} \otimes \frac{d^2}{dt^2}.$$

The strategy here is to generalize this construction to generate abstract multisymplectic Hamiltonian PDEs. The main difference in the PDE case is that the manifold $M$ is the base manifold representing space-time, and the fiber is built on the total exterior algebra bundle rather than just the cotangent bundle.

The starting point is a smooth pseudo-Riemannian manifold $M$ (which here is taken to be the flat space $M = \mathbb{R}^{q,p}$), with constant signature metric which can be represented by

$$[ [u, v]_1 := \langle R u, v \rangle, \tag{3.2}$$

with $\langle \cdot, \cdot \rangle$ a Euclidean inner product, and

$$R = \text{diag}(1, \ldots, 1, -1, \ldots, -1). \tag{3.3}$$

The metric $[\cdot, \cdot]_1$ induces a metric on each of the spaces $\wedge^k(T^*_xM)$. The induced metrics are denoted by

$$[u^{(k)}, v^{(k)}]_k \text{ for } u^{(k)}, v^{(k)} \in A^k(M), \tag{3.4}$$

where $A^k(M)$ is the space of mappings from $M$ into $\wedge^k(T^*_xM)$. Concatenating these spaces gives the total exterior algebra (TEA) bundle denoted by $A(M) := \bigcup_{k=0}^{n} A^k(M)$. Let $r = \text{dim} \left( \wedge^k(T^*_xM) \right)$, then we have the Euclidean space representation of the lifted metric

$$[u^{(k)}, v^{(k)}]_k = \langle R^{(r)} u^{(k)^r}, v^{(k)^r} \rangle, \tag{3.5}$$

10
where \( u^{(k)^r} \) is the lift to \( \mathbb{R}^r \) of \( u^{(k)} \).

The generalization of \( p \cdot dq \) on the cotangent bundle is the following form on the total exterior algebra bundle

\[
\Theta(Z) = \sum_{k=1}^{m} u^{(k)} \wedge \star d u^{(k-1)},
\]

where \( \star \) is Hodge star and \( d \) is an exterior derivative. See [7] for the origin of this form in the case of a positive definite metric, which generates an elliptic partial differential operator (PDO), and see [14] for the case of general indefinite metric, which generates a hyperbolic PDO.

Define the Lagrangian density

\[
L(Z) = \Theta(Z) - S(Z) \text{vol},
\]

where \( S : \mathcal{A}(M) \rightarrow \mathbb{R} \) is a given generalized Hamilton function. Taking the first variation of the functional

\[
\delta \int \Theta(Z) - S(Z) \text{vol} = 0,
\]

generates a nonlinear Dirac operator

\[
J_\partial Z = R^{(r)} \nabla S(Z).
\]

When written out in coordinates, and pre-multiplying by \( R^{(r)} \), the PDE becomes

\[
\sum_{j=1}^{n} R^{(r)} J_j \partial_{x_j} Z = \nabla S(Z).
\]

This PDE is now in standard form for a multisymplectic Hamiltonian PDE [12]. Indeed it is a new class of multisymplectic Hamiltonian PDEs. It is not difficult to show that each \( R^{(r)} J_j \) is symplectic (skew-symmetric and non-degenerate), acting on a space of dimension \( 2^n \), and so \( n \)-independent symplectic structures are generated.

### 3.1 Multisymplectic Dirac operator based on \( M = \mathbb{R}^{1,1} \)

The case of \( M = \mathbb{R}^{1,1} \) with metric tensor \( R = \text{diag}(1,-1) \) is the case of interest in this paper. Take coordinates \((t,x)\) and volume form \( \text{vol} = dt \wedge dx \). Then differential forms in the TEA bundle are of the form \( Z = (\phi, u, v) \) with \( \phi \) a scalar-valued function,

\[
u = u_1 dt + u_2 dx \quad \text{and} \quad v := v dt \wedge dx,
\]

where, to simplify notation, \( v \) is both a form and a coordinate. The PDO in this case acting on \( Z \in \mathcal{A}(M) \) is

\[
J_\partial Z = \begin{pmatrix}
0 & \delta & 0 \\
\delta & 0 & 0 \\
0 & \delta & 0
\end{pmatrix}
\begin{pmatrix}
\phi \\
u \\
v
\end{pmatrix},
\]

and it can be expressed in coordinates as

\[
J_\partial = J_1 \partial_t + J_2 \partial_x
\]
with
\[
J_1 = \begin{pmatrix}
0 & -1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0
\end{pmatrix}, \quad J_2 = \begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & -1 \\
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0
\end{pmatrix}.
\tag{3.11}
\]

The pair \(\{J_1, J_2\}\) generates the Clifford algebra \(\mathcal{C}_{1,1}\),
\[
J_i J_j + J_j J_i = -2\mathcal{R}_{ij} I_4.
\]

The Dirac property and the connection with the d’Alembertian is
\[
J_\partial \circ J_\partial = (J_1 \partial_t + J_2 \partial_x)^2 = J_1^2 \partial_t + (J_1 J_2 + J_2 J_1) \partial_t x + J_2^2 \partial_{xx} = -(\partial_t - \partial_{xx}) \otimes I_4.
\]

Introducing a scalar-valued function \(S : \mathcal{A}(M) \to \mathbb{R}\), a nonlinear Dirac equation is generated in the canonical form
\[
J_\partial Z = \mathcal{R}^{(4)} \nabla S(Z), \quad Z \in \mathcal{A}(M).
\tag{3.12}
\]

The induced metric in this case is
\[
\mathcal{R}^{(4)} = \text{diag}(1, 1, -1, -1) = [+1] \oplus \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \oplus [-1].
\tag{3.13}
\]

This form follows by constructing the induced metric on each of the vector spaces \(\wedge^j\) and then concatenating. The space \(\wedge((\mathbb{R}^{1,1})\) is isomorphic to \(\mathbb{R}^{2,2}\) with metric \(\langle \mathcal{R}^{(4)}, \cdot, \cdot \rangle\).

The operator \(J_1\) is skew-symmetric and \(J_2\) is symmetric and they are both invertible. The induced skew symmetric operators are
\[
M := \mathcal{R}^{(4)} J_1 = \begin{pmatrix}
0 & -1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & -1 & 0
\end{pmatrix}, \quad K := \mathcal{R}^{(4)} J_2 = \begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & -1 \\
-1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix}.
\tag{3.14}
\]

The system
\[
MZ_t + KZ_x = \nabla S(Z), \quad Z \in \mathcal{A}(M),
\tag{3.15}
\]

is then in standard form for a multisymplectic Hamiltonian PDE, and the two operators \(M\) and \(K\) define independent symplectic vector spaces.

When \(M\) is invertible as above, the PDE can be written in evolution form
\[
Z_t + CZ_x = \nabla S(Z), \quad Z(x, 0) = Z_0(x),
\]

with \(C = M^{-1} K\). When \([M, K] = 0\), as is the case for \(M\) and \(K\) in (3.14), the matrix \(C\) is symmetric and the operator \(Z_t + CZ_x\) is hyperbolic. There are a range of results in the literature on existence and well-posedness of equations in this form in general, and Dirac equations in particular (e.g. PELINOVSKY [39] and references therein).
3.2 The coupled mode equation

The coupled-mode equation (CME) which appears in fluid mechanics [18, 27, 28, 29] and optics [40, 19, 3] can be characterised as a multisymplectic Dirac operator on $A(\mathbb{R}^{1,1})$ in the form (3.12). In the literature, the CME is represented in complex-amplitude form

\begin{align*}
  i(A_t + A_x) + \alpha B + \tau |A|^2 A + \nu |B|^2 A + \mu B^2 \overline{A} &= 0, \\
  i(B_t - B_x) + \alpha A + \tau |B|^2 B + \nu |B|^2 B + \mu A^2 \overline{B} &= 0.
\end{align*}

(3.16)

In this equation the coefficients $\alpha$, $\tau$, $\nu$ and $\mu$ are real-valued and $A(x,t)$ and $B(x,t)$ are complex valued functions. Introduce coordinates $(\phi, u, v)$ in $A^0 \times A^1 \times A^2$ and to link more closely with the CME coordinates, take

$w = (w_1, w_2) := (\phi, u_1)$ and $v = (v_1, v_2) := (u_2, v)$.

The system (3.16) is transformed using

\begin{align*}
  A &:= A_1 + iA_2 = w_1 - v_2 + i(w_2 - v_1) \\
  B &:= B_1 + iB_2 = w_1 + v_2 + i(w_2 + v_1).
\end{align*}

In these coordinates the CME becomes

\begin{align*}
  \begin{pmatrix}
    0 & -1 & 0 & 0 \\
    1 & 0 & 0 & 0 \\
    0 & 0 & 0 & 1 \\
    0 & 0 & 1 & 0
  \end{pmatrix}
  \begin{pmatrix}
    w_1 \\
    w_2 \\
    v_1 \\
    v_2
  \end{pmatrix}
  +
  \begin{pmatrix}
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & -1 \\
    1 & 0 & 0 & 0 \\
    0 & -1 & 0 & 0
  \end{pmatrix}
  \begin{pmatrix}
    w_1 \\
    w_2 \\
    v_1 \\
    v_2
  \end{pmatrix}
  =
  \mathcal{R}(4)
  \begin{pmatrix}
    \partial S/\partial w_1 \\
    \partial S/\partial w_2 \\
    \partial S/\partial v_1 \\
    \partial S/\partial v_2
  \end{pmatrix}.
\end{align*}

(3.17)

or, with $Z = (w, v)$ now identified with $\mathbb{R}^4$

\begin{align*}
  \mathbf{J}_1 Z_t + \mathbf{J}_2 Z_x = \mathcal{R}(4) \nabla S(Z),
\end{align*}

(3.18)

using (3.11) with $\mathcal{R}(4) = \text{diag}(1, 1, -1, -1)$.

A special case of (3.16) arises in optics with $\tau = \gamma$, $\nu = 2\gamma$, and $\mu = 0$. It is the one-dimensional model that rules nonlinear wave propagation around a forbidden frequency band gap (cf. Sugny et al. [40]). An even more special case is the massive Thirring model (MTM) where $\tau = \mu = 0$,

\begin{align*}
  i(A_t + A_x) + \alpha B + \nu |B|^2 A &= 0, \\
  i(B_t - B_x) + \alpha A + \nu |A|^2 B &= 0.
\end{align*}

(3.19)

The transformed system for MTM is (3.17) with

\begin{align*}
  S(Z) = -\frac{1}{2} \alpha (w \cdot \mathbf{w} - v \cdot \mathbf{v}) - \frac{1}{4} \nu (w \cdot \mathbf{w} + v \cdot \mathbf{v})^2 + \nu (w_1 v_2 + w_2 v_1)^2.
\end{align*}

(3.20)

3.3 Coupled second-order nonlinear wave equations

The pair of coupled second order nonlinear wave equations

\begin{align*}
  \phi_{tt} - \phi_{xx} + V\phi = 0 \quad \text{and} \quad v_{tt} - v_{xx} - Vv = 0,
\end{align*}

(3.21)
where $V(\phi, v)$ is a given smooth function, can also be transformed to the canonical form (3.15). Introduce new coordinates $(\phi, u_1, u_2, v)$ via

$$u_1 = \phi_t - v_x \quad \text{and} \quad u_2 = \phi_x - v_t.$$ 

Then with

$$S(Z) = \frac{1}{2}(u_1^2 - u_2^2) + V(\phi, v),$$

the coupled equations (3.21) are represented by (3.15). An analysis of the Evans function for an example in this class of nonlinear wave equations in given in §7.

### 3.4 Reversibility

The canonical operators $M$ and $K$ in (3.14) are reversible in the following sense. There exists an involution

$$R = \text{diag}(1, -1, -1, 1),$$

with

$$RM = -MR \quad \text{and} \quad RK = -KR.$$ 

The implication of this symmetry is that if $S$ is reversible, that is $S(RZ) = S(Z)$, then $RZ(-x, -t)$ is a solution whenever $Z(x, t)$ is a solution.

### 4 Solitary waves and linearization

The class of PDEs that we take as a starting point for the development of the theory of linear stability of solitary waves is

$$MZ_t + KZ_x = \nabla S(Z), \quad Z \in \mathbb{R}^4,$$

with $S : \mathbb{R}^4 \to \mathbb{R}$ a smooth scalar-valued function, $\nabla$ the gradient on $\mathbb{R}^4$, and $M, K$ are $4 \times 4$ skew-symmetric matrices. The extension to higher dimension phase space is straightforward in principle, modulo a proliferation of indices, and is discussed in §9.

The skew-symmetric matrices are required to satisfy

$$\det(J(c)) \neq 0, \quad \text{for all } c \in \mathcal{C}, \quad J(c) := K + cM.$$ 

The structure (4.1) and property (4.2) are hypothesis (H1). When $J(c)$ has a non-trivial kernel the theory goes through with the phase space restricted to the complement of the kernel of $J(c)$. This is rare but does happen; an example is in [13].

In (4.2), $\mathcal{C}$ is just an open subset of $\mathbb{R}$. In examples, the existence of solitary waves will inform the definition of $\mathcal{C}$. With assumption (H1), $(\mathbb{R}^4, \Omega)$ is a symplectic vector space with symplectic form

$$\Omega(u, v) = \langle J(c)u, v \rangle, \quad \forall u, v \in \mathbb{R}^4.$$ 

The operator $J(c)$ is not canonical and $J(c)^2 \neq -I$, so does not define a complex structure either. However, neither of these properties are required for the theory here.

In addition to the multisymplectic Dirac operators in §3.1, the class of PDEs (4.1) also includes the case where $M$ is of rank two and $K$ is of rank four. The KdV equation and NLS equation are multisymplectic Hamiltonian PDEs of this latter type [8, 10].
4.1 Solitary wave solutions

The abstract form (4.1) is equivariant with respect to the translation group with action

$$T_s Z(x, t) = Z(x + s, t), \quad \forall s \in \mathbb{R},$$

that is, $T_s Z(x, t)$ is a solution of (4.1) whenever $Z(x, t)$ is solution. This latter property follows since $M$, $K$ and $S(Z)$ do not depend explicitly on $x$. A relative equilibrium associated with this group is a solution of the form

$$Z(x, t) = T_s \hat{Z}(x) := \hat{Z}(\xi) \text{ with } \xi = x + ct,$$

(4.4)

where $c \in \mathbb{R}$. This relative equilibrium solution is called a solitary wave when the following asymptotic conditions are operational

$$\lim_{\xi \to \pm \infty} \|\hat{Z}(\xi)\| = 0,$$

(4.5)

with the convergence exponential. As in §2, the relative equilibrium $\hat{Z}$ can be characterized as a constrained critical point problem. Let

$$H(\hat{Z}) = \int_{-\infty}^{+\infty} \left[ S(\hat{Z}) - \frac{1}{2} \langle K \hat{Z}_\xi, \hat{Z} \rangle \right] \ d\xi \quad \text{and} \quad I(\hat{Z}) = \int_{-\infty}^{+\infty} \frac{1}{2} \langle M \hat{Z}_\xi, \hat{Z} \rangle \ d\xi.$$

(4.6)

The functional $I(\hat{Z})$ is called the momentum of a solitary wave as it is the conserved functional associated, via Noether’s Theorem, to the $x-$translation symmetry of (4.1). Indeed, the symplectic Noether argument is the same as (2.3): the symplectic operator $M$ acting on the generator of the translation group, $\hat{Z}_\xi$, generates the gradient of a functional

$$M \hat{Z}_\xi = \nabla I(\hat{Z}),$$

(4.7)

where here $\nabla I$ is the gradient in $\mathbb{R}^4 \otimes L^2(\mathbb{R})$. Analogous to (2.3), the left-hand side of (4.7) is the product of two commuting skew-symmetric operators $M$ and $\partial_\xi$. The operator $M$ appears in both $I(\hat{Z})$ and the governing equation (4.1) and this will be useful for connecting $dI/dc$ to the Evans function.

Solitary wave solutions then correspond to critical points of $H(\hat{Z})$ restricted to level sets of the function $I(\hat{Z})$, with Euler-Lagrange equation $\nabla H(\hat{Z}) - c \nabla I(\hat{Z}) = 0$, which works out to

$$0 = \nabla H(\hat{Z}) - c \nabla I(\hat{Z}) = \nabla S(\hat{Z}) - K \hat{Z}_\xi - c M \hat{Z}_\xi = \nabla S(\hat{Z}) - J(c) \hat{Z}_\xi,$$

that is, the solitary wave is a homoclinic orbit of the Hamiltonian ODE

$$J(c) \hat{Z}_\xi = \nabla S(\hat{Z}).$$

(4.8)

In the critical point argument, the speed $c$ is a Lagrange multiplier. Solitary waves come in one parameter families parameterized by $c$; that is, $\hat{Z}(\xi, c)$, and the family is non-degenerate when

$$\frac{d}{dc} I \circ \hat{Z} \neq 0 \quad \text{for all } c \in \mathcal{C}.$$

(4.9)

This derivative will be represented by $\frac{dt}{dc}$. The homoclinic orbit solution of (4.8) is assumed to be transversely constructed for all $c \in \mathcal{C}$. The concept of “transversely constructed” is discussed below in (4.27).

It is assumed that there exists a solitary wave solution of the form (4.4)-(4.5) satisfying (4.8), transversely constructed, with (1.9), and a smooth function of $\xi$ and $c$, for all $c \in \mathcal{C}$. This is hypothesis (H2).
4.1.1 Reversible solitary waves

The equation for solitary waves (4.8) is said to be reversible if there exists an involution \( R \) with
\[
RJ(c) = -J(c)R \quad \text{and} \quad S(R\hat{Z}) = S(\hat{Z}).
\]
The first identity is satisfied with (3.23) when \( J(c) = K + cM \) with \( K \) and \( M \) satisfying (3.24).

**Definition 4.1.** A solitary wave solution \( \hat{Z}(\xi, c) \) is a “reversible solitary wave” when the equation (4.8) is reversible with involution \( R \) and the solitary wave satisfies
\[
R\hat{Z}(-\xi, c) = \hat{Z}(\xi, c).
\]

4.2 Linearization and the operator \( L \)

The second variation of \( H - cI \) is a linear operator
\[
L(\xi, c) := D^2H(\hat{Z}) - cD^2I(\hat{Z}).
\]
It is this operator to which the GSS spectral condition is applied. Here the operator \( L \) will play an important role in the Evans function theory but the spectrum of \( L \), other than its zero eigenvalue, will not enter the theory, being replaced by transversality and the coefficient \( [1.5] \). Using (4.6) another representation of \( L \) is
\[
LW = D^2S(\hat{Z})W - KW_\xi - cMW_\xi = B(\xi, c)W - J(c)W_\xi,
\]
with
\[
B(\xi, c) = D^2S(\hat{Z}(\xi, c)).
\]
Differentiating (4.8) with respect to \( \xi \) shows that the tangent vector to the solitary wave is in the kernel of \( L \)
\[
L\hat{Z}_\xi = 0.
\]
It is assumed that \( \text{Ker}(L) \cap L^2(\mathbb{R}) = \text{span}\{\hat{Z}_\xi\} \).

The above properties of the linearization constitute hypothesis (H3).

In the analysis of the Evans function an equation for \( \hat{Z}_c \) will be needed. Differentiate (4.8) with respect to \( c \)
\[
(K + cM)(\hat{Z}_c)_\xi + M\hat{Z}_\xi = B(\xi, c)\hat{Z}_c,
\]
or
\[
L\hat{Z}_c = M\hat{Z}_\xi.
\]
Note the similarity with the second equation in the ODE case (2.10).

The pair (4.14) and (4.15) form a Jordan chain of length exactly two. This property is confirmed by noting that length three would require the existence of \( W \) satisfying \( LW = \hat{M}\hat{Z}_c \). The proof is the same as that in (2.11) and (2.12). By Hypotheses (H2) and (H3) and the symmetry of \( L \), the solvability condition for \( LW = \hat{M}\hat{Z}_c \) is
\[
0 = \int_{-\infty}^{+\infty} \langle \hat{Z}_\xi, M\hat{Z}_c \rangle \, d\xi = -\int_{-\infty}^{+\infty} \langle \nabla I(\hat{Z}), \hat{Z}_c \rangle \, d\xi = -\frac{dI}{dc},
\]
which is non-zero by hypothesis (H2).
4.3 Eigenvalues and Eigenvectors in the system at infinity

We begin the setup of the Evans function by taking a dynamical systems viewpoint of the linearization (4.12). The linearized ODE is

\[ J(c)Z_\xi = B(\xi, c)Z. \] (4.17)

Due to the boundary condition (4.5) the operator \( B \) is asymptotic to a constant matrix

\[ \lim_{\xi \to \pm \infty} B(\xi, c) = B_\infty(c). \]

The “system at infinity” for the steady problem is \( J(c)Z_\xi = B_\infty(c)Z \) which can be solved explicitly,

\[ Z(\xi) = \sum_{j=1}^{4} z_j \zeta_j e^{\mu_j \xi}, \]

where \( z_j \) are arbitrary complex constants, and \( \mu_j(0, c) \) are eigenvalues determined by

\[ \Delta(\mu, 0; c) := \det [B_\infty(c) - \mu J(c)] = 0. \]

The zero in one of the arguments anticipates the introduction of the stability exponent \( \lambda \) in the next section. The vectors \( \zeta_j \) are the eigenvectors satisfying

\[ [B_\infty(c) - \mu_j(c, 0)J(c)] \zeta_j(c, 0) = 0, \quad j = 1, \ldots, 4. \] (4.18)

The adjoint eigenvectors are of the form \( J(c)\eta_j(c, 0) \) with \( \eta_j(c, 0) \) satisfying

\[ [B_\infty(c) + \mu_j(c, 0)J(c)] \eta_j(c, 0) = 0, \quad j = 1, \ldots, 4. \] (4.19)

The adjoint eigenvectors are normalized as

\[ \langle J(c)\eta_i(c, 0), \zeta_j(c, 0) \rangle := \Omega(\eta_i, \zeta_j) = \delta_{i,j}, \quad i, j = 1, \ldots, 4. \] (4.20)

It is assumed that the spectrum of \( B_\infty(c) \) has a two-two splitting. Introducing a numbering the splitting is represented as \( \mu_1(0, c) \) and \( \mu_2(0, c) \) with negative real part, and \( \mu_3(0, c) \) and \( \mu_4(0, c) \) with positive real part. For now, the eigenvalues are assumed to be real and are shown schematically in Figure 1. The case where the eigenvalues form a complex quartet is discussed below in §8.

![Figure 1: Typical position of spatial exponents in the complex \( \mu \)-plane.](image)

When \( \lambda = 0 \) the sum of the eigenvalues is zero since

\[ \mu_1(c, 0) + \mu_2(c, 0) + \mu_3(c, 0) + \mu_4(c, 0) = \text{Tr}(J(c)^{-1}B_\infty(c)) = 0, \] (4.21)

using skew-symmetry of \( J(c) \) and symmetry of \( B_\infty(c) \).
4.4 Orientation

Let \( \{ e_1, \ldots, e_4 \} \) be any fixed (independent of \( c \)) basis, not necessarily the standard basis, for the phase space, \( \mathbb{R}^4 \) and fix an orientation by defining

\[
\text{vol} := e_1 \wedge e_2 \wedge e_3 \wedge e_4.
\]  

(4.22)

The eigenvectors are related to the orientation by,

\[
\zeta_1(c, 0) \wedge \zeta_2(c, 0) \wedge \zeta_3(c, 0) \wedge \zeta_4(c, 0) = K(c, 0) \text{vol}, \quad \text{with } K \neq 0, \quad \forall \ c \in \mathcal{C}.
\]  

(4.23)

The simple and real \( \mu \)-eigenvalue assumption, the normalization (4.20), and the orientation property (4.23) constitute hypothesis (H4).

4.5 The \( \xi \)-dependent stable and unstable subspaces

Consistent with the two-two splitting are solutions of the \( \xi \)-dependent equation (4.17)

\[
E^s(\xi, 0) = \text{span} \{ \hat{Z}_\xi, a^+ \} \quad \text{and} \quad E^u(\xi, 0) = \text{span} \{ \hat{Z}_\xi, a^- \},
\]  

(4.24)

where \( \hat{Z}_\xi \) decays exponentially as \( \xi \to \pm \infty \) and \( a^\pm \) are the other solutions which satisfy

\[
\lim_{\xi \to +\infty} e^{-\mu_1 \xi} a^+(\xi, c) = \zeta_1(c, 0) \quad \text{and} \quad \lim_{\xi \to -\infty} e^{-\mu_4 \xi} a^-(\xi, c) = \zeta_4(c, 0),
\]  

(4.25)

with the convergence exponential. In general \( a^\pm \) are not bounded as \( \xi \to \mp \infty \).

Remark. Here we have associated \( \hat{Z}_\xi \) with the exponents \( \mu_2 \) and \( \mu_3 \), and have associated \( a^\pm \) with the exponents \( \mu_1 \) and \( \mu_4 \). It could be the other way around. However we will show in §5 that the Evans function is independent of the permutation \( 3 \leftrightarrow 4 \).

Existence of \( a^\pm \) follows from the stable/unstable manifold theorem. This strategy for constructing \( a^\pm \) is used by GALVAO & GELFREICH \[23\] and is effective in both the case of real eigenvalues and complex quartets. The existence of \( a^\pm \) can also be proved directly, by analyzing the linear system (4.17) using asymptotic theory for linear ODEs \[20, 30\].

Proposition 4.1. \( E^s(\xi, 0) \) and \( E^u(\xi, 0) \) are Lagrangian subspaces with respect to the symplectic structure \( \Omega \) in (4.3).

Proof. The proof is given for \( E^s \). It is required to show that

\[
\Omega(\hat{Z}_\xi, a^+) = 0 \quad \text{for all } \xi \in \mathbb{R}.
\]  

(4.26)

Differentiate this expression, use (4.17), and the skew symmetry of \( J(c) \),

\[
\frac{d}{d\xi} \Omega(\hat{Z}_\xi, a^+) = \langle J(c) \hat{Z}_\xi, a^+ \rangle + \langle J(c) \hat{Z}_\xi, (a^+) \rangle = 0,
\]  

using symmetry of \( B(\xi, c) \). This proves that \( \Omega(\hat{Z}_\xi, a^+) \) is a constant for all \( \xi \in \mathbb{R} \). Now use the fact that \( \hat{Z}_\xi \) and \( a^+ \) both go to zero as \( \xi \to \infty \) to conclude (4.26). A similar proof confirms that \( E^u \) is a Lagrangian subspace. \( \blacksquare \)
4.6 Transversality, the Lazutkin invariant, and orientation

The Lazutkin invariant of a homoclinic orbit is defined in (1.5). It is a property of the intersection between the stable and unstable manifolds which form the homoclinic orbit, and it is a symplectic invariant \[24, 23, 16\]. The fact that \( \Pi \) is independent of \( \xi \) follows from

\[
\frac{d}{d\xi} \Omega(a^-, a^+) = \Omega((a^-)_\xi, a^+) + \Omega(a^-, (a^+)_\xi)
= \langle B(\xi, c)a^-, a^+ \rangle - \langle a^-, B(\xi, c)a^+ \rangle
= \langle B(\xi, c)a^-, a^+ \rangle - \langle B(\xi, c)a^-, a^+ \rangle = 0.
\]

A homoclinic orbit is said to be transversely constructed when

\[
\Pi := \Omega(a^-, a^+) \neq 0 \quad \text{for all } c \in \mathcal{C}.
\] \hspace{1cm} (4.27)

To fix the sign of \( \Pi \) we have to synchronize the orientations of the stable and unstable subspace, relative to the ambient space. The unstable subspace satisfies

\[
\text{span}\{\hat{Z}_\xi, a^-\} \to \text{span}\{\zeta_3, \zeta_4\} \quad \text{as} \quad \xi \to -\infty.
\] \hspace{1cm} (4.28)

Hence there exists a constant \( C^- \) with the property that

\[
e^{-(\mu_3+\mu_4)\xi} \hat{Z}_\xi(-\xi) \wedge a^-(\xi) \to C^-\zeta_3 \wedge \zeta_4 \quad \text{as} \quad \xi \to -\infty.
\] \hspace{1cm} (4.29)

Similarly, the stable subspace satisfies

\[
\text{span}\{\hat{Z}_\xi, a^+\} \to \text{span}\{\zeta_1, \zeta_2\} \quad \text{as} \quad \xi \to +\infty.
\] \hspace{1cm} (4.30)

Hence there exists a constant \( C^+ \) with the property that

\[
e^{-(\mu_1+\mu_2)\xi} \hat{Z}_\xi(\xi) \wedge a^+(\xi) \to C^+\zeta_1 \wedge \zeta_2 \quad \text{as} \quad \xi \to +\infty.
\] \hspace{1cm} (4.31)

Take \( \xi \) to \( -\xi \) in this formula, and use the eigenvalue sum formula (4.21),

\[
e^{-(\mu_3+\mu_4)\xi} \hat{Z}_\xi(-\xi) \wedge a^+(\xi) \to C^+\zeta_1 \wedge \zeta_2 \quad \text{as} \quad \xi \to -\infty.
\] \hspace{1cm} (4.32)

Take the wedge product between (4.29) and (4.32) and take the limit \( \xi \to -\infty \),

\[
e^{-2(\mu_3+\mu_4)\xi} \hat{Z}_\xi(-\xi) \wedge a^+(\xi) \wedge \hat{Z}_\xi(\xi) \wedge a^- = C^+C^-K(c, 0) \text{vol},
\] \hspace{1cm} (4.33)

using (4.23). We take the sign of \( C^+C^-K(c, 0) \) to be fixed throughout the analysis. This fixes the sign of \( \Pi \) as the sign of \( a^- \) can not be changed without changing the sign of \( a^+ \) and vice versa. These properties of \( \Pi \) constitute Hypothesis (H5).
4.7 Asymptotics of $\hat{Z}_\xi$ and the factor $\chi$

The asymptotic behaviour of $\hat{Z}_\xi$ as $\xi \to -\infty$ is associated with either the $\mu_3$ or the $\mu_4$ eigenvalue. We will see below in §5 that the Evans function is independent of this choice. Hence, we will assume without loss of generality that $\hat{Z}_\xi$ is associated with $\mu_3$ as $\xi \to -\infty$.

There exists real numbers $\chi^\pm$ with

$$\lim_{\xi \to -\infty} e^{-\mu_3 \xi} \hat{Z}_\xi = \chi^- \zeta_3 \quad \text{and} \quad \lim_{\xi \to +\infty} e^{\mu_3 \xi} \hat{Z}_\xi = \chi^+ \eta_3. \quad (4.34)$$

These expressions can be inverted to give formula for $\chi^\pm$

$$\chi^- = \lim_{\xi \to -\infty} e^{-\mu_3 \xi} \Omega(\eta_3, \hat{Z}_\xi) \quad \text{and} \quad \chi^+ = \lim_{\xi \to +\infty} e^{\mu_3 \xi} \Omega(\hat{Z}_\xi, \zeta_3). \quad (4.35)$$

The factor $\chi$ in the derivative of the Evans function is defined by

$$\chi = (\chi^+ \chi^-)^{-1}. \quad (4.36)$$

**Proposition 4.2.** Suppose the basic state is reversible in the sense of Definition 4.1. Then

$$\text{sign}(\chi) = \text{sign}[\Omega(\zeta_3, R\zeta_3)]. \quad (4.37)$$

**Proof.** Differentiating (4.10) with respect to $\xi$ gives

$$-R\hat{Z}_\xi(-\xi, c) = \hat{Z}_\xi(\xi, c). \quad (4.38)$$

Now use the formula (4.34) for the $\hat{Z}_\xi$ asymptotics as $\xi \to -\infty$, take $\xi \to -\xi$ in this equation, and act on it with $R$,

$$\lim_{\xi \to +\infty} e^{\mu_3 \xi} R\hat{Z}_\xi(-\xi, c) = \chi^- R\zeta_3. \quad \text{Substituting (4.38)} \quad \text{then gives}$$

$$\lim_{\xi \to +\infty} e^{\mu_3 \xi} \hat{Z}_\xi(\xi, c) = -\chi^- R\zeta_3. \quad (4.39)$$

Now use the second asymptotic formula in (4.34) $\xi \to +\infty$

$$\lim_{\xi \to +\infty} e^{\mu_3 \xi} \hat{Z}_\xi(\xi, c) = \chi^+ \eta_3. \quad (4.40)$$

Combining (4.39) and (4.40),

$$-\chi^- R\zeta_3 = \chi^+ \eta_3. \quad (4.41)$$

On the other hand $\zeta_3$ and $\eta_3$ satisfy

$$[B - \mu_3 J(c)]\zeta_3 = 0 \quad \text{and} \quad [B + \mu_3 J(c)]\eta_3 = 0.$$
The fact that \( BR = RB \) follows from differentiating \( S(RZ) = S(Z) \) twice and setting \( Z = 0 \). Therefore there exists \( \vartheta \in \mathbb{R} \) such that

\[
R\zeta_3 = \vartheta\eta_3. 
\] (4.42)

Act on both sides of (4.42) with \( J(c) \) and take the pairing with \( \zeta_3 \) to get

\[
\langle J(c)R\zeta_3, \zeta_3 \rangle = \vartheta\langle J(c)\eta_3, \zeta_3 \rangle, 
\]

and so \( \vartheta = -\Omega(\zeta_3, R\zeta_3) \). Applying this formula and (4.42) to (4.41),

\[
-\chi^-\vartheta\eta_3 = \chi^+\eta_3 \Rightarrow -\vartheta\chi^- = \chi^+ \Rightarrow \Omega(\zeta_3, R\zeta_3)(\chi^-)^2 = \chi^-1.
\]

Comparing signs on each side of this expression proves the proposition.

\[\blacksquare\]

### 4.8 Summary of the key geometric properties of a wave

To summarize, there are three key properties of the family of solitary waves that will feed into the stability analysis: (a) the derivative of the momentum \( dI/dc \) in (4.9), (b) the transversality coefficient \( \Pi \) in (1.5), and (c) the asymptotic property (4.34) which generates \( \chi \) in (4.36).

### 5 Linear stability and the Evans function

The linearization of the PDE (4.1) about the solitary wave solution (4.4) is

\[
MZ_t + J(c)Z_\xi = B(\xi, c)Z, \quad Z \in \mathbb{R}^4. 
\] (5.1)

Introduce the spectral ansatz \( Z(x, t) = e^{\lambda t}u(\xi, \lambda) \). Then the eigenvalue problem for \( \lambda \in \mathbb{C} \) is

\[
u_\xi = A(\xi, \lambda)u, \quad u \in \mathbb{C}^4, \quad \lambda \in \Lambda, \] (5.2)

for some open set \( \Lambda \subset \mathbb{C} \), with

\[
A(\xi, \lambda) := J(c)^{-1}(B(\xi, c) - \lambda M). 
\] (5.3)

**Definition 5.1.** There exists an unstable eigenvalue, in the linearization about the solitary wave, if there exists a solution of (5.2), for some \( \lambda \in \mathbb{C} \) with \( \operatorname{Re}(\lambda) > 0 \), and \( u(\cdot, \lambda) \in \mathbb{R}^4 \otimes L^2(\mathbb{R}) \).

The asymptotic condition (4.5) assures that

\[
\int_{-\infty}^{+\infty} \|A(\xi, \lambda) - A^\infty(\lambda)\|d\xi < +\infty, \quad \lambda \in \Lambda. 
\] (5.4)

In this integral the “system at infinity” is defined by

\[
A^\infty(\lambda) := \lim_{\xi \to \pm\infty} A(\xi, \lambda) \quad \forall \lambda \in \Lambda. 
\] (5.5)

with the dependence on \( c \) suppressed for brevity. This limit is assumed to exist for any fixed \( c \) and uniformly for \( \lambda \in \Lambda \). The spectrum of \( A^\infty(\lambda) \) consists of four eigenvalues with the two-two splitting

\[
\operatorname{Re}(\mu_1) \leq \operatorname{Re}(\mu_2) < 0 < \operatorname{Re}(\mu_3) \leq \operatorname{Re}(\mu_4) \quad \forall \lambda \in \Lambda. 
\] (5.6)
The set Λ is defined below.

The eigenvalue problem \([A^\infty(\lambda) - \mu I]\zeta = 0\) is unwrapped into
\[
[B^\infty(c) - \lambda M - \mu_j(c, \lambda)J(c)]\zeta_j(c, \lambda) = 0, \quad j = 1, \ldots, 4, \quad \forall \lambda \in \Lambda.
\] (5.7)

The extension of (4.23) to include the \(\lambda\)-dependent eigenvectors is
\[
V(c, \lambda) := \zeta_1(c, \lambda) \wedge \zeta_2(c, \lambda) \wedge \zeta_3(c, \lambda) \wedge \zeta_4(c, \lambda) = K(c, \lambda) vol,
\] (5.8)

with \(K(c, \lambda) \neq 0\), for all \(c \in C\) and \(\lambda \in \Lambda\).

The adjoint eigenvalue problem \([A^\infty(\lambda)H - \mu I]\psi = 0\), where the superscript \(H\) denotes complex-conjugate transpose, is unwrapped to
\[
[B^\infty(c) + \lambda M + \mu_j(c, \lambda)J(c)]\eta_j(c, \lambda) = 0, \quad j = 1, \ldots, 4, \quad \forall \lambda \in \Lambda,
\] (5.9)

with \(\eta_j = J(c)^{-1}\overline{\psi}_j\). The adjoint eigenvectors are normalized as
\[
\langle J(c)\eta_i(c, \lambda), \zeta_j(c, \lambda) \rangle := \Omega(\eta_i, \zeta_j) = \delta_{i,j}, \quad i, j = 1, \ldots, 4.
\] (5.10)

With this normalization, we can define a dual to \(V\) in (5.8) as
\[
V^*(c, \lambda) := J(c)\eta_1(c, \lambda) \wedge J(c)\eta_2(c, \lambda) \wedge J(c)\eta_3(c, \lambda) \wedge J(c)\eta_4(c, \lambda).
\] (5.11)

The continuous spectrum is defined by
\[
\sigma^{cont} = \left\{ \lambda \in \mathbb{C} : \det[B^\infty(c) - \lambda M - i\kappa J(c)] = 0, \quad \kappa \in \mathbb{R} \right\}.
\] (5.12)

Normally, it is assumed that the continuous spectrum is purely imaginary, \(\sigma^{cont} \subset i\mathbb{R}\), thereby limiting instability considerations to unstable eigenvalues. Here, no special assumption is imposed on the continuous spectrum since our main interest is in the derivatives of the Evans function at \(\lambda = 0\), and the existence of unstable point spectra. But we will need \(\lambda = 0\) in the set \(\Lambda\).

The set \(\Lambda\) is an open set in the complex plane, including the origin, such that for all \(\lambda \in \Lambda\) the four eigenvalues of \(A^\infty(\lambda)\) are simple and satisfy the constraints (5.6). The above properties of the eigenvalues and eigenvectors at infinity and the definition of the set \(\Lambda\) constitutes hypothesis (H6). The assumption of simple eigenvalues can be relaxed by working on exterior algebra spaces, or using maximally analytic eigenvectors [11].

5.1 Constructing the Evans function

There are many equivalent ways of defining the Evans function (e.g. Chapters 8–10 in [30]). The direct approach is to take the wedge product of the individual vector-valued solutions of (5.2). We will first define the Evans function that way, and then introduce an equivalent definition which pairs solutions of (5.2) with solutions of the adjoint equation, and brings in the \(J(c)\)-symplectic structure.

Using standard asymptotic theory for ODEs [20, 30] there are four \((\xi, \lambda)\)-dependent vectors satisfying
\[
(u_j)_\xi = A(\xi, \lambda)u_j, \quad j = 1, \ldots, 4,
\] (5.13)
with the asymptotic properties
\[ \lim_{\xi \to +\infty} e^{-\mu_1(\lambda)\xi} u_1(\xi, \lambda) = \zeta_1(\lambda), \quad \lim_{\xi \to +\infty} e^{-\mu_2(\lambda)\xi} u_2(\xi, \lambda) = \zeta_2(\lambda), \]
\[ \lim_{\xi \to -\infty} e^{-\mu_3(\lambda)\xi} u_3(\xi, \lambda) = \zeta_3(\lambda), \quad \lim_{\xi \to -\infty} e^{-\mu_4(\lambda)\xi} u_4(\xi, \lambda) = \zeta_4(\lambda), \]
(5.14)
suppressing the dependence on \(c\) as it is now secondary. The stable subspace is represented by \(u_1 \land u_2\), with
\[ \lim_{\xi \to +\infty} u_1(\xi, \lambda) \land u_2(\xi, \lambda) = 0, \]
with the convergence exponential, and the unstable subspace is represented by \(u_3 \land u_4\), with
\[ \lim_{\xi \to -\infty} u_3(\xi, \lambda) \land u_4(\xi, \lambda) = 0, \]
with the convergence exponential. If these two spaces intersect then there is a solution which decays exponentially as \(\xi \to \pm \infty\), thereby generating an eigenfunction with eigenvalue \(\lambda\). The Evans function captures these interesections and hence eigenvalues. It is defined by
\[ D(\lambda) \operatorname{vol} = e^{-\tau(c)\lambda} u_1(\xi, \lambda) \land u_2(\xi, \lambda) \land u_3(\xi, \lambda) \land u_4(\xi, \lambda), \]
(5.15)
with \(\tau(c)\lambda\) related to the trace of \(A(\xi, \lambda)\) by
\[ \operatorname{Tr}(A(\xi, \lambda)) = \operatorname{Tr}(J(c)^{-1}B(\xi, \lambda) - \lambda J(c)^{-1}M) = -\lambda \operatorname{Tr}(J(c)^{-1}M) := \lambda \tau(c), \]
(5.16)
since \(J(c)\) is skew-symmetric and \(B(\xi, c)\) is symmetric. The function \(D(\lambda)\) is independent of \(\xi\) and an analytic function of \(\lambda\) for all \(\lambda \in \Lambda\) and \(\lambda \in D^{-1}(0)\) is an element of the point spectrum \([1,30]\).

Here an equivalent definition of the Evans function, in terms of individual vectors of \((5.2)\) and its symplectic adjoint, are used.

**Proposition 5.1.** The symplectic adjoint of \(u_\xi = A(\xi, \lambda)u\) is
\[ w_\xi = A(\xi, -\lambda)w, \quad w \in \mathbb{C}^4, \]
(5.17)
and \(w(\xi, \lambda)\) is an analytic functions of \(\lambda\) for \(\lambda \in \Lambda\).

**Proof.** Start with a construction of the adjoint of \(A(\xi, \lambda)\). The symplectic adjoint is that linear operator \(A^{*\text{symp}}\) satisfying
\[ \Omega(A^{*\text{symp}}w, u) = \Omega(w, Au). \]
Calculating, using \(A(\xi, \lambda) = J(c)^{-1}(B(\xi, c) - \lambda M)\), gives
\[ \Omega(w, Au) = \langle Jw, Au \rangle = \langle A^T Jw, u \rangle = -\langle (B(\xi, c) + \lambda M)J^{-1}Jw, u \rangle = -\langle JJ^{-1}(B(\xi, c) + \lambda M)w, u \rangle = -\langle JA(\xi, -\lambda)w, u \rangle = \Omega(A^{*\text{symp}}w, u), \]
giving \(A^{*\text{symp}} = -A(\xi, -\lambda)\). Now bring in the derivative
\[ \Omega(w_\xi, u) + \Omega(w, u_\xi) = \frac{d}{d\xi} \Omega(w, u), \]
23
and so
\[
\Omega(w, u_\xi - A(\xi, \lambda)u) = \Omega(w, u_\xi) - \Omega(w, A(\xi, \lambda)u)
\]
\[
= -\Omega(w_\xi, u) - \Omega(A^{\mathsf{symplectic}}, w, u) + \frac{d}{d\xi} \Omega(w, u)
\]
\[
= -\Omega(w_\xi, u) + \Omega(A(\xi, -\lambda)w, u) + \frac{d}{d\xi} \Omega(w, u)
\]
or
\[
\Omega(w, u_\xi - A(\xi, \lambda)u) + \Omega(w_\xi - A(\xi, -\lambda)w, u) = \frac{d}{d\xi} \Omega(w, u)
\]
At this point, one can either bring in integration, or note that if \(u\) satisfies (5.2) and \(w\) satisfies (5.17) then \(\Omega(w, u)\) is independent of \(\xi\).

Another way to prove the form of the adjoint (5.17) is to start with the formal adjoint of (5.2),
\[
W_\xi = -A(\xi, \lambda)^H W, \quad W \in \mathbb{C}^4, \quad \lambda \in \Lambda,
\]
where the superscript \(H\) signifies complex conjugate transpose, and \(W\) is not an analytic function of \(\lambda\). Pre-multiply by \(J(c)^{-1}\) and use the special form of \(A(\xi, \lambda)\) in (5.3),
\[
(J(c)^{-1}W)_\xi = -J(c)^{-1}A(\xi, \lambda)^H W = J(c)^{-1}[B(\xi, c) + \bar{\Lambda}M](J(c)^{-1}W),
\]
or
\[
(J(c)^{-1}W)_\xi = A(\xi, -\bar{\lambda})(J(c)^{-1}W).
\]
Define
\[
w(\xi, \lambda) = J(c)^{-1}W(\xi, \lambda),
\]
then the vector-valued functions \(w(\xi, \lambda)\) are analytic and satisfy (5.17).

There are four solutions of (5.17) with the asymptotic properties
\[
\lim_{\xi \to -\infty} e^{+\mu_1(\lambda)\xi} w_1(\xi, \lambda) = \eta_1(\lambda), \quad \lim_{\xi \to -\infty} e^{+\mu_2(\lambda)\xi} w_2(\xi, \lambda) = \eta_2(\lambda),
\]
\[
\lim_{\xi \to +\infty} e^{+\mu_3(\lambda)\xi} w_3(\xi, \lambda) = \eta_3(\lambda), \quad \lim_{\xi \to +\infty} e^{+\mu_4(\lambda)\xi} w_4(\xi, \lambda) = \eta_4(\lambda),
\]
where \(\eta_j, j = 1, \ldots, 4\) are adjoint eigenvectors (5.9).

**Theorem 5.2.** With the orientation (4.22) and the normalizations (5.10), the Evans function (5.15) can be transformed to the representation
\[
D(\lambda)\text{vol} = \det \begin{bmatrix} \Omega(w_3, u_3) & \Omega(w_3, u_1) \\ \Omega(w_4, u_3) & \Omega(w_4, u_1) \end{bmatrix} \mathcal{V}(\xi, \lambda).
\]
The proof of the theorem is based on two remarkable formulae, which bring the symplectic structure into the Evans function (5.15),
\[
e^{-\tau(c)\lambda \xi} u_1(\xi, \lambda) \wedge u_2(\xi, \lambda) \wedge \mathcal{V}^* = J(c)w_3(\xi, \lambda) \wedge J(c)w_4(\xi, \lambda),
\]
and
\[
J(c)\eta_3 \wedge J(c)\eta_4 = \zeta_1 \wedge \zeta_2 \wedge \mathcal{V}^*.
\]
where \(\wedge\) is the interior product, and \(\mathcal{V}^*\) is the dual four-form defined in (5.11). The proof of these two formulae and Theorem 5.2 requires an excursion into exterior algebra and they are given in Appendix A.
5.2 The equivalence class of Evans functions

The two Evans functions \( D^A(\lambda) \) and \( D^B(\lambda) \) are equivalent. Evans functions form an equivalence class. Two representations \( D^A(\lambda) \) and \( D^B(\lambda) \), of an Evans function, are equivalent if there exists a non-vanishing analytic function \( C(\lambda) \) such that \( D^A(\lambda) = C(\lambda)D^B(\lambda) \) for all \( \lambda \in \Lambda \). When orientation of \( D^A(\lambda) \) and \( D^B(\lambda) \) along the real axis is of interest, then \( C(\lambda) \big|_{\lambda \in \mathbb{R}} \) is required to be real and positive.

We will select a particular representation from this class. Using \( \mathcal{V}(c,\lambda) = K(c,\lambda)\text{vol} \) from (5.8), the explicit formula for the scalar-valued function \( D(\lambda) \) in (5.21) is

\[
D(\lambda) = K(c,\lambda)\det \begin{bmatrix}
\Omega(w_3, u_3) & \Omega(w_3, u_4) \\
\Omega(w_4, u_3) & \Omega(w_4, u_4)
\end{bmatrix}.
\]

However, as noted in (5.8), \( K(c,\lambda) \) is non-zero for all \( c \in \mathcal{C} \) and all \( \lambda \in \Lambda \). Hence it can be factored out giving the equivalent function

\[
D(\lambda) := \det \begin{bmatrix}
\Omega(w_3, u_3) & \Omega(w_3, u_4) \\
\Omega(w_4, u_3) & \Omega(w_4, u_4)
\end{bmatrix}.
\] (5.24)

Here and henceforth, it is this function that will be used as “the Evans function”. The proof of the following property of the Evans function is evident from inspection.

**Proposition 5.3.** The Evans function (5.24) is invariant under the permutation 3 ↔ 4.

This property will be useful in the proof of the derivative formula as it shows that it does not matter whether \( \hat{Z}_\xi \) is asymptotic to the direction \( \zeta_3 \) or \( \zeta_4 \) as \( \xi \to -\infty \).

The factor \( \chi \) can be eliminated by defining an equivalent Evans function with \( \chi \) built in

\[
D_R(\lambda) = \det \begin{bmatrix}
\Omega(\chi^+w_3, \chi^-u_3) & \Omega(\chi^+w_3, u_4) \\
\Omega(\chi^-w_3, u_3) & \Omega(\chi^-w_4, u_4)
\end{bmatrix}.
\] (5.25)

This Evans function is equivalent to (5.24). The advantage is that the derivative formula (1.4) simplifies to

\[
D''_R(0) = \Pi \frac{dI}{dc}.
\]

However, we prefer to use the definition (5.24) of the Evans function as a starting point and bring the factor \( \chi \) out explicitly in the derivative formula.

If a class of PDEs of interest has solitary waves with \( \Pi \) and \( \chi \) of one sign, then an equivalent Evans function can be defined without \( \Pi \) or \( \chi \), e.g.

\[
D_{\text{new}}(\lambda) = \frac{1}{\chi\Pi} D(\lambda).
\]

In this case \( D_{\text{new}}''(0) = dI/dc \). For example, this case is relevant for the Evans function in PW (e.g. equation (1.2)).

5.3 The \( \lambda \to 0 \) limit of the stable and unstable spaces

As above, we take, without loss of generality, \( u_3(\xi, \lambda) \to \text{span}\{\hat{Z}_\xi\} \). Hence we have the following limits

\[
\lim_{\lambda \to 0} u_3(\xi, \lambda) = \frac{1}{\chi^-} \hat{Z}_\xi \quad \text{and} \quad \lim_{\lambda \to 0} u_4(\xi, \lambda) = a^- (\xi).
\] (5.26)
The first limit is due to the fact that
\[ e^{-\mu_3(\lambda)} u_3 \to \zeta_3 \quad \text{but} \quad e^{-\mu_3(\lambda)} \hat{Z}_\xi \to \chi_\xi \zeta_3, \]
using (5.14) and (4.34). The \( \lambda \)-limit of the second term in (5.26) follows from (5.14) and (4.25).

Similarly, the adjoint eigenfunctions have \( \lambda \to 0 \) limits
\[ \lim_{\lambda \to 0} w_3(\xi, \lambda) = \frac{1}{\chi^+} \hat{Z}_\xi \quad \text{and} \quad \lim_{\lambda \to 0} w_4(\xi, \lambda) = a^+(\xi). \] (5.27)

The form of the eigenfunctions \( u_j \) and adjoint eigenfunctions \( w_j \), the limits (5.26) and (5.26), and the form of the Evans function (5.24) constitute hypothesis (H7).

6 Derivatives of the Evans function

In this section the main result of the paper is proved, the connection between \( D''(0) \), transversality, \( \chi \), and \( dI/dc \) that was asserted in the introduction in formula (1.4).

**Theorem 6.1.** The Evans function (5.24), associated with the linearization of the class of PDEs (4.1), about the solitary wave solutions (4.4), under the hypotheses (H1)-(H7), has the following derivatives at \( \lambda = 0 \),
\[ D(0) = 0, \quad D'(0) = 0, \quad \text{and} \quad D''(0) = \chi \Pi \frac{dI}{dc}, \] (6.1)
where \( \Pi \) is the transversality coefficient (1.5), \( \chi \) is the \( \hat{Z}_\xi \) coefficient (4.36), and \( I(c) \) is the momentum (4.6) evaluated on the \( c \)-dependent family of solitary waves.

**Proof.** The fact that \( D(0) = 0 \) follows from the fact that \( L \) has a zero eigenvalue (4.14). However, the proof in the context of the Evans function is a bit more interesting, as it brings in the Lagrangian subspace property of \( E^s \) and \( E^u \). The proof proceeds with the evaluation of \( D(\lambda) \) in (5.24) at \( \lambda = 0 \),
\[ D(0) = \det \begin{bmatrix} 0 & 0 \\ 0 & -\Pi \end{bmatrix}. \] (6.2)

The zeros in the first column and row are confirmed by noting that \( u_3 \) and \( w_3 \) satisfy (5.26) and (5.27). Now use skew-symmetry of \( \Omega \) and the Lagrangian subspace property of the stable and unstable subspaces (Proposition 4.1) to conclude
\[ \Omega(w_3, u_3)|_{\lambda=0} = \chi \Omega(\hat{Z}_\xi, \hat{Z}_\xi) = 0 \]
\[ \Omega(w_3, u_4)|_{\lambda=0} = (\chi^+)^{-1} \Omega(\hat{Z}_\xi, a^-) = 0 \]
\[ \Omega(w_4, u_3)|_{\lambda=0} = (\chi^-)^{-1} \Omega(a^+, \hat{Z}_\xi) = 0 \]
\[ \Omega(w_4, u_4)|_{\lambda=0} = \Omega(a^+, a^-) = -\Pi. \]

Substitution into (5.24) then confirms the zero structure in (6.2).

To prove the properties of the first and second derivatives of \( D(\lambda) \), define the entries of the matrix in \( D(\lambda) \) as
\[ D(\lambda) = (d_1(\lambda)d_2(\lambda) - d_3(\lambda)d_4(\lambda)) \] (6.3)
where
\[ d_1(\lambda) = \Omega(w_3, u_3) \quad d_2(\lambda) = \Omega(w_4, u_4), \]
\[ d_3(\lambda) = \Omega(w_3, u_3), \quad d_4(\lambda) = \Omega(w_4, u_3). \]

It follows from (6.2) that
\[ d_1(0) = d_3(0) = d_4(0) = 0, \quad \text{and} \quad d_2(0) = -\Pi. \quad (6.4) \]

Computing the first derivative
\[ D'(\lambda) = (d'_1(\lambda)d_2(\lambda) + d_1(\lambda)d'_2(\lambda) - d'_3(\lambda)d_4(\lambda) - d_3(\lambda)d'_4(\lambda)). \]

Evaluating at \( \lambda = 0 \) and using (6.4)
\[ D'(0) = -d'_1(0)\Pi. \quad (6.5) \]

Now
\[ d'_1(\lambda) = \Omega(\partial_\lambda w_3, u_3) + \Omega(w_3, \partial_\lambda u_3). \quad (6.6) \]

For \( \partial_\lambda u_3 \) and \( \partial_\lambda w_3 \), start with their defining equation,
\[ J(u_3)_3 = [B - \lambda M]u_3 \quad \text{and} \quad J(w_3)_3 = [B + \lambda M]w_3, \]

and differentiate with respect to \( \lambda \),
\[ J(u_3)_{3\lambda} = [B - \lambda M]u_3 - Mu_3 \quad \text{and} \quad J(w_3)_{3\lambda} = [B + \lambda M]w_3 + Mw_3. \quad (6.7) \]

Set \( \lambda = 0 \),
\[ L(u_3)_3 = Mu_3 \bigg|_{\lambda=0} = (\chi^-)^{-1}M\tilde{Z}_\xi \]
\[ L(w_3)_3 = -Mw_3 \bigg|_{\lambda=0} = -(\chi^+)^{-1}M\tilde{Z}_\xi, \]

using (5.26) and (5.27). Now use equation (4.15), giving
\[ (u_3)_3 \bigg|_{\lambda=0} = (\chi^-)^{-1}\tilde{Z}_c + C_a\tilde{Z}_\xi \quad \text{and} \quad (w_3)_3 \bigg|_{\lambda=0} = -(\chi^+)^{-1}\tilde{Z}_c + C_b\tilde{Z}_\xi, \quad (6.8) \]

with \( C_a \) and \( C_b \) arbitrary constants.

Substitute the expressions (6.8) into (6.6) evaluated at \( \lambda = 0 \),
\[ d'_1(0) = \left[ \Omega(\partial_\lambda w_3, u_3) + \Omega(w_3, \partial_\lambda u_3) \right] \bigg|_{\lambda=0} \]
\[ = \left[ \Omega(\chi^+)^{-1}\tilde{Z}_c + C_b\tilde{Z}_\xi, (\chi^-)^{-1}\tilde{Z}_\xi) + \Omega((\chi^+)^{-1}\tilde{Z}_c, (\chi^-)^{-1}\tilde{Z}_c + C_a\tilde{Z}_\xi) \right] \]
\[ = 2\lambda\Omega(\tilde{Z}_\xi, \tilde{Z}_c). \]

This latter term is zero. To see this, first show that it is independent of \( \xi \),
\[ \frac{d}{d\xi} \Omega(\tilde{Z}_c, \tilde{Z}_\xi) = \langle J(\tilde{Z}_c)_\xi, \tilde{Z}_\xi \rangle - \langle \tilde{Z}_c, J(\tilde{Z}_\xi)_\xi \rangle 
= \langle B\tilde{Z}_c, \tilde{Z}_\xi \rangle + \langle M\tilde{Z}_\xi, \tilde{Z}_\xi \rangle - \langle \tilde{Z}_c, B\tilde{Z}_\xi \rangle 
= 0, \]

27
and so $\Omega(\hat{Z}_\xi, \hat{Z}_c)$ is a constant, but this constant clearly vanishes at $\xi = \pm \infty$ and so the form is zero for all $\xi$. This proves that $d'_1(0) = 0$ and so $D'(0) = 0$ in (6.5).

The second derivative is

$$D''(\lambda) = \left(d''_1(\lambda)d_2(\lambda) + d'_1(\lambda)d'_2(\lambda) + d_1(\lambda)d''_2(\lambda) + d''_1(\lambda)d'_2(\lambda) \right) - d''_3(\lambda)d_4(\lambda) - d'_3(\lambda)d'_4(\lambda) - d_3(\lambda)d''_4(\lambda) - d''_3(\lambda)d'_4(\lambda).$$  \hspace{1cm} (6.9)

Evaluating (6.9) at $\lambda = 0$ eliminates the second derivatives of $d_j$ for $j = 2, 3, 4$, leaving

$$D''(0) = (d''_1(0)d_2(0) + 2d'_1(0)d'_2(0) - 2d'_1(0)d'_2(0)).$$  \hspace{1cm} (6.10)

The second term is zero due to $d'_1(0) = 0$ as was shown above. A similar argument can be used to show that $d''_2(0) = 0$ as follows,

$$d''_2(0) = (\chi^{-1}\Omega(\hat{Z}_c, a^-) + (\chi^+)^{-1}\Omega(\hat{Z}_\xi, (u_i)_{\lambda=0}|_{\lambda=0}).$$  \hspace{1cm} (6.11)

The sum of the two terms is constant (since $d_3(\lambda)$ and $d'_3(\lambda)$ are independent of $\xi$). The first term goes to zero as $\xi \to -\infty$ as both $\hat{Z}_c$ and $a^-$ go to zero. For the second term $\hat{Z}_\xi$ also goes to zero as $\xi \to -\infty$, so all that is needed is that $(u_i)_{\lambda=0}|_{\lambda=0}$ be bounded as $\xi \to -\infty$. But $u_i|_{\lambda=0}$ goes to zero exponentially as $\xi \to -\infty$ and $\partial_\lambda u_i$ will only add a polynomial in $\xi$ to the exponential decay, giving zero for the second term in (6.11) as well.

Hence the second derivative (6.10) reduces to

$$D''(0) = d_2(0)d'_2(0) = -\Pi d''_2(0).$$  \hspace{1cm} (6.12)

To compute $d''_1(0)$ start with $d'_1(\lambda)$ in (6.6). Using (6.7), we can write,

$$\partial_\xi \langle J(w_3)_{\lambda}, u_3 \rangle = \langle \left[ B + \lambda M \right](w_3)_{\lambda}, u_3 \rangle + \langle Mw_3, u_3 \rangle$$

$$\quad - \langle (w_3)_{\lambda}, [B - \lambda M]u_3 \rangle,$$

$$\partial_\xi \langle Jw_3, (u_3)_{\lambda} \rangle = \langle \left[ B + \lambda M \right]w_3, (u_3)_{\lambda} \rangle + \langle w_3, Mu_3 \rangle$$

$$\quad - \langle w_3, [B - \lambda M](u_3)_{\lambda} \rangle,$$

which combine to leave us with:

$$\partial_\xi \langle J(w_3)_{\lambda}, u_3 \rangle = \langle Mw_3, u_3 \rangle = -\partial_\xi \langle Jw_3, (u_3)_{\lambda} \rangle.$$  \hspace{1cm} (6.14)

If we now take some $R > 0$ then we can integrate the first part of this over the range $\xi \in [0, R]$ and the second part over $\xi \in [-R, 0]$ to get:

$$\left[ \langle J(w_3)_{\lambda}, u_3 \rangle \right]_{\xi=0}^{\xi=R} = \int_0^R \langle Mw_3, u_3 \rangle \, d\xi,$$

$$\left[ - \langle Jw_3, (u_3)_{\lambda} \rangle \right]_{\xi=-R}^{\xi=0} = \int_{-R}^0 \langle Mw_3, u_3 \rangle \, d\xi.$$  \hspace{1cm} (6.15)

They can be combined to give:

$$d'_1(\lambda) \bigg|_{\xi=0} = - \int_{-R}^R \langle Mw_3, u_3 \rangle \, d\xi + \langle J(w_3)_{\lambda}, u_3 \rangle \bigg|_{\xi=R} + \langle Jw_3, (u_3)_{\lambda} \rangle \bigg|_{\xi=-R}. \hspace{1cm} (6.17)$$

28
(Note that although this value of \( d_1' (\lambda) \) is specifically evaluated at \( \xi = 0 \), since \( d_1 \) is independent of \( \xi \) it will take this value for all \( \xi \).) Taking the limit \( R \to \infty \) allows us to write this as

\[
d_1' (\lambda) = - \int_{-\infty}^{+\infty} \langle M w_3, u_3 \rangle \, d\xi + \ell (\lambda), \tag{6.18}
\]

where the function \( \ell (\lambda) \) is defined as

\[
\ell (\lambda) = \lim_{\xi \to +\infty} \langle J (w_3)_\lambda, u_3 \rangle + \lim_{\xi \to -\infty} \langle J w_3, (u_3)_\lambda \rangle.
\]

Differentiate this function with respect to \( \lambda \) to get an expression for \( d_1'' (\lambda) \):

\[
d_1'' (\lambda) = - \int_{-\infty}^{+\infty} \langle M (w_3)_\lambda, u_3 \rangle + \langle M w_3, (u_3)_\lambda \rangle \, d\xi + \ell' (\lambda). \tag{6.19}
\]

Now

\[
\ell' (\lambda) = \lim_{\xi \to +\infty} \left[ \langle J (w_3)_\lambda, u_3 \rangle + \langle J (w_3)_\lambda, (u_3)_\lambda \rangle \right] + \lim_{\xi \to -\infty} \left[ \langle J (w_3)_\lambda, (u_3)_\lambda \rangle + \langle J w_3, (u_3)_\lambda \rangle \right]. \tag{6.20}
\]

However, since

\[
\lim_{\xi \to +\infty} e^{w_3 (\lambda) \xi} w_3 = \eta_3 (\lambda)
\]

we can deduce that

\[
w_3 e^{w_3 (\lambda) \xi} = \eta_3 + o (1) \quad \text{for } \xi \to +\infty.
\]

This in turn implies that

\[
(w_3)_\lambda e^{w_3 (\lambda) \xi} = p (\xi, \lambda) \eta_3 + o (1) \quad \text{for } \xi \to +\infty,
\]

where \( p (\xi, \lambda) \) is a quadratic polynomial in \( \xi \). Since the exponential term will dominate the quadratic polynomial this tells us that

\[
\lim_{\xi \to +\infty} (w_3)_\lambda = 0. \tag{6.21}
\]

The same argument can be used to show that

\[
\lim_{\xi \to -\infty} (u_3)_\lambda = 0. \tag{6.22}
\]

Now set \( \lambda = 0 \) in (6.20) and use these asymptotic properties. The first term becomes

\[
\lim_{\xi \to +\infty} \left[ \langle J (w_3)_\lambda |_{\lambda = 0}, (\lambda^-)^{-1} \hat{Z}_\xi \rangle + \langle J (- (\lambda^+)^{-1} \hat{Z}_c + C_b \hat{Z}_\xi), (\lambda^-)^{-1} \hat{Z}_c + C_a \hat{Z}_\xi \rangle \right] = 0,
\]

using (6.21) and the vanishing of \( \Omega (\hat{Z}_\xi, \hat{Z}_c) \). Set \( \lambda = 0 \) in the second term in (6.20),

\[
\lim_{\xi \to -\infty} \left[ \langle J (- (\lambda^+)^{-1} \hat{Z}_c + C_b \hat{Z}_\xi), (\lambda^-)^{-1} \hat{Z}_c + C_a \hat{Z}_\xi \rangle + \langle (\lambda^+)^{-1} \langle J \hat{Z}_\xi, (u_3)_\lambda |_{\lambda = 0} \rangle \rangle \right] = 0,
\]

using (6.22) and \( \Omega (\hat{Z}_\xi, \hat{Z}_c) = 0 \). Therefore \( \ell' (0) = 0 \) and

\[
d_1'' (0) = - \int_{-\infty}^{+\infty} \langle M (- (\lambda^+)^{-1} \hat{Z}_c + C_b \hat{Z}_\xi), (\lambda^-)^{-1} \hat{Z}_c \rangle + (\lambda^+)^{-1} \langle M \hat{Z}_\xi, (\lambda^-)^{-1} \hat{Z}_c + C_a \hat{Z}_\xi \rangle \, d\xi
\]

\[
= -2 \chi \int_{-\infty}^{+\infty} \langle \hat{M} \hat{Z}_\xi, \hat{Z}_c \rangle \, d\xi
\]

\[
= -2 \chi \frac{dI}{dc}, \tag{6.23}
\]

29
using the expression for $\frac{dI}{dc}$ in (4.16). Substituting this result into (6.12) then gives

$$D''(0) = 2 \chi \Pi \frac{dI}{dc}.$$  

Scaling $D$ to eliminate the 2 then completes the proof of the Theorem.  

Remark. Since (5.24) is “the Evans function”, scaling $D$ to eliminate the 2 should be interpreted as a trivial re-definition $D \mapsto \frac{1}{2} D$.

Using Theorem 6.1 above and Theorem 10.1 in [16] an alternative formula for the second derivative in terms of the Maslov index is obtained,

**Corollary 6.2.** Under the above hypotheses, an alternative formula for $D''(0)$ is,

$$D''(0) = \chi \Pi \frac{dI}{dc}.$$  

(6.24)

The Maslov index of an orbit, say a homoclinic or periodic orbit, is a generalization of the Morse index, and equals the Morse index in special cases (e.g. COX, ET AL. [21], LATUSHKIN & SUKHTAIEV [31]). The use of the Maslov index in the stability of waves is now a burgeoning subject (e.g. BECK, ET AL. [4], CHARDARD ET AL. [17, 16], and references therein). To go into the details of the Maslov index would take us too far afield. The new observation here is just the connection between the parity of the Maslov index and the Evans function.

**Corollary 6.3.** Let $d_\infty = \text{sign}(D(\lambda_\infty))$, for some real $\lambda_\infty \gg 1$. Then there exists an unstable real eigenvalue of the spectral problem (5.2) when $d_\infty D''(0) < 0$, or

$$\chi \Pi \frac{dI}{dc} d_\infty < 0.$$  

(6.25)

We will not study $d_\infty$ here as there are well-established strategies for the determining the large $\lambda$ behaviour of the Evans function in the Hamiltonian context (e.g. §1(g) of [38], §8 of [10], and Appendix B of [11]), and in the non-Hamiltonian context (e.g. §5B of [1] and Chapter 9 of [30]).

**7 Example: a coupled wave equation**

To illustrate the theory it is applied to a nonlinear wave equation in the class introduced in §3.3,

\[ \phi_{tt} - \phi_{xx} + \partial_\phi V(\phi, v) = 0, \]
\[ v_{tt} - v_{xx} - \partial_v V(\phi, v) = 0, \]

with

\[ V(\phi, v) = 2\phi^2 - 2\phi^3 - 2v^2 + v^3 + \frac{p}{2}(2\phi - v)^2. \]

Hypothesis (H1) is satisfied by rewriting the equations in the form

$$MZ_t + KZ_x = \nabla S(Z), \quad Z \in \mathbb{R}^4$$  

(7.2)

with

$$Z = \begin{pmatrix} \phi \\ u_1 \\ u_2 \\ v \end{pmatrix}, \quad M = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$  

(7.3)
where
\[ u_1 = \phi_t - v_x, \quad u_2 = \phi_x - v_t, \]
and
\[ S(Z) = \frac{1}{2}(u_1^2 - u_2^2) + 2\phi^3 - 2\phi^3 - 2v^2 + v^3 + \frac{p}{2}(2\phi - v)^2. \] (7.4)
The symplectic operator \( J(c) \) is
\[
J = K + cM = \begin{pmatrix}
0 & -c & 1 & 0 \\
c & 0 & 0 & -1 \\
-1 & 0 & 0 & c \\
0 & 1 & -c & 0
\end{pmatrix},
\] (7.5)
with \( \det(J(c)) = (1 - c^2)^2 \). The other part of hypothesis (H1) is satisfied by taking \( c \in \mathbb{R} \setminus \pm 1 \) but the set \( C \) will be restricted further below for the existence of the family of solitary waves.

An exact solitary wave solution is given by
\[
\phi = \hat{\phi}(\xi, c) \quad \text{and} \quad v = 2\hat{\phi}(\xi, c), \quad \xi = x + ct
\] (7.6)
where
\[
\hat{\phi}(\xi, c) = \text{sech}^2(\alpha \xi), \quad \alpha = \frac{1}{\sqrt{1 - c^2}}.
\] (7.7)
Hence the set \( C \) is taken to be
\[ C = \{ c : -1 < c < +1 \}. \] (7.8)
In terms of the \( Z \)-coordinates the solitary wave solution is
\[
\hat{Z}(\xi, c) = \begin{pmatrix}
\hat{\phi} \\
-(2 - c)\hat{\phi}_\xi \\
(1 - 2c)\hat{\phi}_\xi \\
2\hat{\phi}
\end{pmatrix},
\] (7.9)
and satisfies
\[
J(c)\hat{Z}_\xi = \nabla S(Z). \] (7.10)
The hypothesis (H2) is satisfied: the family of solitary waves satisfies (7.10) and is a smooth function of \( \xi \) and \( c \) for \( c \in C \). Transversality, and the non-degeneracy of \( dI/dc \) will be verified below.

Moreover this solitary wave is reversible. With the reversor \( R \), defined in (3.23), the operator \( J(c) \) anti-commutes with \( R \) and \( S(Z) \) in (7.4) is \( R \)-invariant. Hence the governing equation (7.10) is reversible. Acting on the solitary wave with \( R \) gives
\[
R\hat{Z}(\xi, c) = \begin{pmatrix}
\hat{\phi}(\xi, c) \\
+(2 - c)\hat{\phi}(\xi, c)_\xi \\
-(1 - 2c)\hat{\phi}(\xi, c)_\xi \\
2\hat{\phi}(\xi, c)
\end{pmatrix} = \begin{pmatrix}
\hat{\phi}(\xi, c) \\
-(2 - c)\hat{\phi}(\xi, c)_\xi \\
(1 - 2c)\hat{\phi}(\xi, c)_\xi \\
2\hat{\phi}(\xi, c)
\end{pmatrix} = \hat{Z}(\xi, c),
\] (7.11)
confirming Definition 4.1.
7.1 Linearization about the solitary wave

The linearization about the solitary wave solution is

\[ JZ_\xi = [B(\xi, c) - \lambda M] Z, \quad Z \in \mathbb{C}^4 \]  \hspace{1cm} (7.12)

with

\[
[B(\xi, c) - \lambda M] = \begin{pmatrix}
4p + 4 - 12\hat{\phi}(\xi, c) & \lambda & 0 & -2p \\
-\lambda & 1 & 0 & 0 \\
0 & 0 & -1 & -\lambda \\
-2p & 0 & \lambda & p - 4 + 12\hat{\phi}(\xi, c)
\end{pmatrix}.
\]  \hspace{1cm} (7.13)

The system at infinity is

\[ JZ_\xi = [B^\infty(c) - \lambda M] Z, \quad Z \in \mathbb{C}^4, \]  \hspace{1cm} (7.14)

with \([B^\infty(c) - \lambda M]\) the same as (7.13) but with \(\hat{\phi}\) set to zero. The eigenvalues of the system at infinity are defined by \(\Delta(\mu, \lambda) = 0\) with

\[
\Delta(\mu, \lambda) = \det [B^\infty(c) - \lambda M - \mu J(c)]
\]
\[
= (\mu^2 - \varrho^2)^2 - (8 + 3p)(\mu^2 - \varrho^2) + 16 + 12p
\]

where \(\varrho = \lambda + c\mu\). The continuous spectrum is

\(\sigma^{\text{cont}} = \{\lambda \in \mathbb{C} : \Delta(i\kappa, \lambda) = 0, \quad \kappa \in \mathbb{R}\}\)
\[
= \left\{ \lambda = -i\kappa \pm i\sqrt{4 + \frac{3}{2}p + \kappa^2 \pm \frac{3}{2}p}, \quad \kappa \in \mathbb{R} \right\}.
\]

There are four branches with a gap about the origin,

\(\sigma^{\text{cont}} = \left\{ \lambda = i\tilde{\kappa} : \{\tilde{\kappa} \in \mathbb{R} \setminus (-2\alpha^{-1}, 2\alpha^{-1}) \} \cup \{\tilde{\kappa} \in \mathbb{R} \setminus (-\alpha^{-1}\sqrt{4 + 3p}, \alpha^{-1}\sqrt{4 + 3p})\} \right\}.
\]

With \(p > 0\) the gap consists of \(-2\alpha^{-1} < \tilde{\kappa} < 2\alpha^{-1}\). A schematic of the branches of continuous spectra, showing the gap about the origin, is illustrated in Figure 2. The set \(\Lambda\) can be taken to be the right-half plane along with a neighbourhood of the origin.

![Figure 2: Schematic of the four branches of continuous spectra on the imaginary axis.](image-url)
Now set $\lambda = 0$ and set $\Delta(\mu, 0) = 0$, giving the four $\mu$–eigenvalues

$$
\mu_1 = -\sqrt{\frac{4 + 3p}{1 - c^2}}, \quad \mu_2 = -\sqrt{\frac{4}{1 - c^2}}, \quad \mu_3 = \sqrt{\frac{4}{1 - c^2}}, \quad \mu_4 = \sqrt{\frac{4 + 3p}{1 - c^2}}.
$$

With $p > 0$ and $c \in \mathcal{C}$ they are simple and real, and a schematic of the eigenvalue positions is shown in Figure 3. The tangent vector $\hat{Z}_\xi$ is bi-asymptotic to the $\mu$–eigenvalues of slowest decay

$$
\begin{pmatrix}
\mu_1 \\
\mu_2 \\
\mu_3 \\
\mu_4
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
\text{Re}(\mu) \\
\text{Im}(\mu)
\end{pmatrix}
$$

Figure 3: Typical position of spatial exponents in the complex $\mu$–plane when $p > 0$.

($\mu_2$ and $\mu_3$). On the other hand, when $p < 0$ (but $p > -\frac{4}{3}$) the tangent vector $\hat{Z}_\xi$ is bi-asymptotic to the $\mu$–eigenvalues of maximal decay ($\mu_1$ and $\mu_4$).

The eigenvectors when $\lambda = 0$ are

$$
\begin{align*}
\zeta_1 &= z_1 \begin{pmatrix} 2 \\ (2 - c) \mu_1 \\ 2 \end{pmatrix}, \quad \zeta_2 = z_2 \begin{pmatrix} 1 \\ (c - 2) \mu_2 \\ 2 \end{pmatrix}, \quad \zeta_3 = z_3 \begin{pmatrix} 1 \\ (c - 2) \mu_3 \\ 2 \end{pmatrix}, \quad \zeta_4 = z_4 \begin{pmatrix} 2 \\ (2 - c) \mu_4 \\ 1 \end{pmatrix},
\end{align*}
$$

with the $z_j$ arbitrary real numbers, and the adjoint eigenvectors are

$$
\begin{align*}
\eta_1 &= h_1 \begin{pmatrix} 2 \\ (1 - 2c) \mu_1 \\ 1 \end{pmatrix}, \quad \eta_2 = h_2 \begin{pmatrix} 1 \\ (2 - c) \mu_2 \\ 2 \end{pmatrix}, \quad \eta_3 = h_3 \begin{pmatrix} 1 \\ (2 - c) \mu_3 \\ 2 \end{pmatrix}, \quad \eta_4 = h_4 \begin{pmatrix} 2 \\ (1 - 2c) \mu_4 \\ 1 \end{pmatrix},
\end{align*}
$$

with $h_j$ arbitrary real numbers. However, the normalization $\Omega(\eta_i, \zeta_j) = \delta_{i,j}$ gives the constraints

$$
\begin{align*}
h_1 z_1 &= \frac{\alpha}{6\sqrt{4 + 3p}}, \quad h_2 z_2 = -\frac{\alpha}{12}, \quad h_3 z_3 = \frac{\alpha}{12}, \quad h_4 z_4 = -\frac{\alpha}{6\sqrt{4 + 3p}}.
\end{align*}
\tag{7.15}
$$

The full bevy of eigenvector expressions is given for information, as this is more detail than is needed. The function $K(c, 0)$ is calculated by taking the wedge product of the eigenvectors

$$
\zeta_1 \wedge \zeta_2 \wedge \zeta_3 \wedge \zeta_4 = K(c, 0) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3 \wedge \mathbf{e}_4, \tag{7.16}
$$

with $\mathbf{e}_1, \ldots, \mathbf{e}_4$ the standard basis. Calculating gives

$$
\begin{align*}
K(c, 0) &= \left. -9(1 - c^2)((\mu_1 - \mu_4)(\mu_2 - \mu_3)) z_1 z_2 z_3 z_4 \right|_{\lambda=0} \\
&= \left. -36 z_1 z_2 z_3 z_4 \sqrt{\alpha^2 \lambda^2 + 4 \sqrt{\alpha^2 \lambda^2 + 4 + 3p}} \right|_{\lambda=0} \\
&= -72 z_1 z_2 z_3 z_4 \sqrt{4 + 3p}.
\end{align*}
$$

33
This completes the verification of Hypothesis (H4).

Since the solitary wave is reversible \((7.11)\), we have that sign(\(\chi\)) is determined from
\[
\Omega(\zeta_3, R\zeta_3) = -6\mu_2(z_3)^2(1 - c^2) < 0.
\]
(7.17)

By Proposition 4.2, \(\chi < 0\). Although not necessary, an explicit formula for \(\chi\) can be obtained as well. Differentiating \((7.9)\) with respect to \(\xi\) and calculating gives
\[
\lim_{\xi \to -\infty} e^{-\mu_3 \xi} \tilde{Z}_\xi = \frac{8\alpha}{z_3} \tilde{\phi}_\xi \quad \text{and} \quad \lim_{\xi \to +\infty} e^{\mu_3 \xi} \tilde{Z}_\xi = -\frac{8\alpha}{h_3} \eta_\xi.
\]
Hence
\[
\chi^{-} = \left(\frac{8\alpha}{z_3}\right) \left(-\frac{8\alpha}{h_3}\right) = -64\frac{\alpha^2}{z_3 h_3} = -768\alpha \Rightarrow \chi = -\frac{1}{768\alpha} \quad (7.18)
\]
using \((7.15)\).

### 7.2 Derivative of the momentum \(dI/dc\)

It is now immediate from Theorem 6.1 that \(D(0) = D'(0) = 0\). To compute \(D''(0)\), the two factors \(\Pi\) and \(\frac{dI}{dc}\) need to be computed. The easier of the two is \(I(c)\) and its derivative:

\[
I = \frac{1}{2} \int_{-\infty}^{+\infty} \langle M \tilde{Z}_\xi, \tilde{Z} \rangle \, d\xi
\]
\[
= \frac{1}{2} \int_{-\infty}^{+\infty} (2 - c) \hat{\phi} \hat{\phi}_\xi - (2 - c) \hat{\phi}^2 + 2(1 - 2c) \hat{\phi}^2 - 2(1 - 2c) \hat{\phi} \hat{\phi}_\xi \, d\xi
\]
\[
= \frac{3c}{2} \int_{-\infty}^{+\infty} \hat{\phi} \hat{\phi}_\xi - \hat{\phi}^2 \, d\xi
\]
\[
= -3c \int_{-\infty}^{+\infty} \hat{\phi}^2 \, d\xi
\]
\[
= -\frac{16}{5} \frac{c}{\sqrt{1 - c^2}},
\]
(7.19)

after integrating by parts and using the exact solution \((7.7)\). Differentiating with respect to \(c\),
\[
\frac{df}{dc} = -\frac{16}{5} (1 - c^2)^{-\frac{3}{2}} = -\frac{16}{5} \alpha^3
\]
(7.20)

which is strictly negative for \(c \in \mathcal{C}\). This completes the verification of most of hypothesis (H2). Transversality will be verified below.

### 7.3 Computing the transversality coefficient \(\Pi\)

To compute the transversality coefficient \(\Pi\), the tangent vectors \(a^\pm\) are needed. Because of the nature of this example, these tangent vectors can be computed explicitly. Express the solutions of \((7.12)\) by \(Z = (\hat{\phi}, \hat{u}_1, \hat{u}_2, \hat{v})\). Then \(\tilde{u}_1\) and \(\tilde{u}_2\) can be obtained from \((7.12)\) as
\[
\tilde{u}_1 = \lambda \hat{\phi} + c \hat{\phi}_\xi - \hat{v}_\xi \quad \text{and} \quad \tilde{u}_2 = -\lambda \hat{v} - c \hat{v}_\xi + \hat{\phi}_\xi.
\]
(7.21)
This allows us to rewrite the first and fourth components of (7.12) as the coupled pair of equations:

\[
\begin{align*}
(1 - c^2)\ddot{\phi}_{xx} - 2c\lambda\dot{\phi}_x + (\tilde{\varphi} - 4p - \lambda^2)\dot{\phi} + 2p\ddot{\phi} &= 0, \\
(1 - c^2)\ddot{\psi}_{xx} - 2c\lambda\dot{\psi}_x + (\tilde{\varphi} + p - \lambda^2)\dot{\psi} - 2p\ddot{\psi} &= 0,
\end{align*}
\]

(7.22a, 7.22b)

where \( \tilde{\varphi} := 12\dot{\phi} - 4 \). If we now introduce the transformation

\[
\dot{\phi} = e^{\alpha^2c\lambda\xi}(\psi_1 + 2\psi_2), \quad \dot{\psi} = e^{\alpha^2c\lambda\xi}(2\psi_1 + \psi_2),
\]

then equations (7.22) will become

\[
\begin{align*}
\alpha^{-2}(\psi_1 + 2\psi_2)_{\xi\xi} + (\tilde{\varphi} - 4p - \alpha^2\lambda^2)(\psi_1 + 2\psi_2) + 2p(2\psi_1 + \psi_2) &= 0, \\
\alpha^{-2}(2\psi_1 + \psi_2)_{\xi\xi} + (\tilde{\varphi} + p - \alpha^2\lambda^2)(2\psi_1 + \psi_2) - 2p(\psi_1 + 2\psi_2) &= 0.
\end{align*}
\]

(7.23a, 7.23b)

These equations decouple to give

\[
\begin{align*}
\alpha^{-2}(\psi_1)_{\xi\xi} + (\tilde{\varphi} - \alpha^2\lambda^2)\psi_1 &= 0, \\
\alpha^{-2}(\psi_2)_{\xi\xi} + (\tilde{\varphi} - 3p - \alpha^2\lambda^2)\psi_2 &= 0.
\end{align*}
\]

(7.24a, 7.24b)

which can be rearranged as

\[
\begin{align*}
\alpha^{-2}(\psi_1)_{\xi\xi} + 12\text{sech}^2(\alpha\xi)\psi_1 &= (4 + \alpha^2\lambda^2)\psi_1, \\
\alpha^{-2}(\psi_2)_{\xi\xi} + 12\text{sech}^2(\alpha\xi)\psi_2 &= (4 + 3p + \alpha^2\lambda^2)\psi_2.
\end{align*}
\]

(7.25)

The analysis of these equations for \( \lambda \neq 0 \) and the construction of the Evans function is given in Appendix B. Here we are interested in the case \( \lambda = 0 \) and the construction of \( a^\pm \).

Now set \( \lambda = 0 \) in the equations (7.25) for \( \psi_1 \) and \( \psi_2 \). These exact solutions can be used to confirm the kernel condition (4.14), verifying Hypothesis (H3). \( a^\pm \) will be produced by the following two \( \psi_2 \) solutions (now denoted by \( \psi^\pm \)) of (7.25)

\[
\psi^\pm = e^{\mp\alpha\gamma\xi} \left[ \pm \frac{p\gamma}{5} + \left( 1 + \frac{6p}{5} \right) \tanh(\alpha\xi) \pm \gamma \tanh^2(\alpha\xi) + 3\tanh^3(\alpha\xi) \right]
\]

where \( \gamma = \sqrt{4 + 3p} \). The exponents are related to the \( \mu \)–eigenvalues by

\[
\mu_1(c, 0) = -\alpha\gamma \quad \text{and} \quad \mu_4(c, 0) = +\alpha\gamma.
\]

By reversing the transformations to express \( \tilde{\phi}, \tilde{u}_1, \tilde{u}_2, \tilde{v} \) in terms of \( \psi^\pm \) we find that

\[
a^+ = \begin{pmatrix} 2\psi^+ \\ (2c - 1)\psi^+ \\ (2 - c)\psi^+ \end{pmatrix}, \quad a^- = \begin{pmatrix} 2\psi^- \\ (2c - 1)\psi^- \\ (2 - c)\psi^- \end{pmatrix}.
\]

Substitution into the formula for \( \Omega \) then gives

\[
\Omega(a^-, a^+) = 4(1 - c^2)\psi^-\psi^+ + (2c - 1)^2\psi^-\psi^+ - (2 - c)^2\psi^-\psi^+ + (c^2 - 1)\psi^-\psi^+ = 3(1 - c^2)\psi^-\psi^+ + (4c^2 - 4c + 1 - 4 + 4c - c^2)\psi^-\psi^+ = 3\alpha^{-2}(\psi^-\psi^+ - \psi^-\psi^+).
\]

(7.26)
It is easy to check that this expression is independent of $\xi$ so we can evaluate it at any value of $\xi$ we choose. If we take $\xi = 0$ then since

$$\psi^{\pm} + \alpha e^{\pm \gamma \xi} \text{sech}^2(\alpha \xi) \left[ 1 + \frac{6p}{5} \pm 2\gamma \tanh(\alpha \xi) + 3 \tanh^2(\alpha \xi) \right]$$

we get

$$\Pi := \Omega(a^-, a^+) = 3 \alpha^{-1} \left[ \left( 1 + \frac{6p}{5} - \frac{p \gamma^2}{5} \right) \left( 1 + \frac{p \gamma}{5} \right) - \left( -\frac{p \gamma}{5} \right) \left( 1 + \frac{6p}{5} - \frac{p \gamma^2}{5} \right) \right]$$

$$= \frac{6p \gamma}{5 \alpha} \left( 1 + \frac{6p - p(4 - 3p)}{5} \right)$$

$$= \frac{6p \gamma}{25 \alpha} (5 + 2p - 3p^2)$$

$$= \frac{6p}{25 \alpha} \sqrt{4 + 3p(5 - 3p)(1 + p)}. \quad (7.27)$$

It remains to check the orientation condition (4.33). The asymptotics of each of the solutions in the unstable direction, that is, as $\xi \to -\infty$ is

$$\lim_{\xi \to -\infty} e^{-\mu_3 \xi} \tilde{Z}_\xi(\xi, c) = \frac{8\alpha}{z_3} \zeta_3(c, 0), \quad \lim_{\xi \to -\infty} e^{-\mu_4 \xi} a^-(\xi) = -\frac{\mathcal{P}}{z_4} \zeta_4(c, 0),$$

where $\mathcal{P} = \frac{1}{5}[(p + 5)\sqrt{4 + 3p} + 6p + 10]$, with $3p + 4 > 0$, and as $\xi \to +\infty$,

$$\lim_{\xi \to +\infty} e^{-\mu_2 \xi} \tilde{Z}_\xi(\xi, c) = -\frac{8\alpha}{z_2} \zeta_2(c, 0), \quad \lim_{\xi \to +\infty} e^{-\mu_1 \xi} a^+(\xi) = -\frac{\mathcal{P}}{z_1} \zeta_1(c, 0).$$

Combining gives

$$\lim_{\xi \to -\infty} e^{-(\mu_3 + \mu_4) \xi} \tilde{Z}_\xi(\xi, c) \wedge a^-((\xi) = -\frac{8\alpha \mathcal{P}}{z_3 z_4} \zeta_3(c, 0) \wedge \zeta_4(c, 0),$$

and

$$\lim_{\xi \to +\infty} e^{-(\mu_1 + \mu_2) \xi} \tilde{Z}_\xi(\xi, c) \wedge a^+(\xi) = \frac{8\alpha \mathcal{P}}{z_1 z_2} \zeta_1(c, 0) \wedge \zeta_2(c, 0),$$

and so, using $\mu_1 + \mu_2 + \mu_2 + \mu_4 = 0$,

$$\lim_{\xi \to -\infty} e^{-2(\mu_3 + \mu_4) \xi} \tilde{Z}_\xi(-\xi) \wedge a^+(-\xi) \wedge \tilde{Z}_\xi(\xi) \wedge a^-((\xi) = C^- C^+ K(c, 0) \text{vol}, \quad (7.28)$$

with

$$C^- C^+ K(c, 0) = -\frac{64 \alpha^2 \mathcal{P}^2}{z_1 z_2 z_3 z_4} \left( -288 z_1 z_2 z_3 z_4 \sqrt{4 + 3p} \right) = 18432 \alpha^2 \mathcal{P}^2 \sqrt{4 + 3p} > 0.$$
7.4 Summary and $D''(0)$

The three key properties of the solitary wave that feed into $D''(0)$ are

\[
\chi = -\frac{1}{768\alpha} < 0
\]

\[
\frac{dI}{dc} = -\frac{16}{5}(1 - c^2)^{-3/2} < 0
\]

\[
\Pi = \frac{6p}{25\alpha}\sqrt{4 + 3p(5 - 3p)(1 + p)},
\]

with the sign of $\Pi$ dependent on sign($5 - 3p$) when $p > 0$.

Combining these formulae and applying Theorem 6.1 gives $D(0) = 0$, $D'(0) = 0$, and

\[
D''(0) = \frac{\alpha p}{500}\sqrt{4 + 3p(5 - 3p)(1 + p)}.
\] (7.29)

With the assumption of $p$ positive, the second derivative is positive for $0 < p < 5/3$ and negative for $p > 5/3$. Hence, there is at least one interval of $p$ with unstable eigenvalues. To determine whether $p < \frac{5}{3}$ or $p > \frac{5}{3}$ is the unstable region we need $d_\infty$. The sign of $d_\infty$ can be obtained abstractly (see comments below Corollary 6.3), but here we have enough information to obtain it explicitly (see §7.5 below), and we find that

\[
d_\infty = \text{sign}(D(\lambda_\infty)) = +1, \quad \lambda_\infty \gg 1.
\]

Hence, applying Corollary 6.3 gives that the solitary wave is unstable when $5 - 3p < 0$ or $p > \frac{5}{3}$. The theory is inconclusive in the case $p < \frac{5}{3}$. However, we will compute the exact Evans function below and find that for $p < \frac{5}{3}$ there are two unstable eigenvalues. Indeed, we will find that the graph of the Evans function, along the real line, has the form shown in Figure 4. This result shows

both the strength and weakness of the theory. The strength is that with limited information about the solitary wave, and the correct sign, the existence of an unstable eigenvalue is proved. The weakness is that it can not capture an even number of unstable eigenvalues, and misses unstable eigenvalues that are off the real line.

7.5 An exact Evans function

The Evans function for this example can be constructed explicitly, and the details are given in Appendix B. The result is

\[
D_E(\lambda) = \frac{3S_{\alpha}\lambda^2}{16(225)^2}(3 + x^2)(5 - x^2)(3 + 3p + x^2)(3p + x^2)(5 - 3p - x^2),
\] (7.30)
with \( x^2 = \alpha^2 \lambda^2 \) and
\[
S = \sqrt{4 + x^2} \sqrt{4 + 3p + x^2},
\]
and the subscript \( E \) is for exact. Taking the second derivative
\[
D''_E(0) = \frac{\alpha p}{500} \sqrt{4 + 3p(5 - 3p)}(p + 1),
\]
in agreement with \( D''(0) \) in (7.29).

We see from the exact expression that there are many more eigenvalues that go unnoticed, particularly on the imaginary axis in the gap of the continuous spectrum. The set of eigenvalues for \( 5 - 3p > 0 \) are
\[
x = \pm i \sqrt{3}, \quad x = 0^2, \quad x = \pm \sqrt{5}, \quad x = \pm i \sqrt{3 + 3p}, \quad x = \pm i \sqrt{3p}, \quad x = \pm \sqrt{5 - 3p}.
\]
When \( 0 < p < 1 \) the three positive purely imaginary eigenvalues satisfy
\[
0 < i \frac{\sqrt{3p}}{\alpha} < i \frac{\sqrt{3}}{\alpha} < i \frac{\sqrt{3 + 3p}}{\alpha}.
\]
However, when \( p > 1/3 \), the largest imaginary eigenvalue is in the continuous spectrum. A schematic showing all the eigenvalues, as well as the branches of continuous spectrum, in the case \( 0 < p < 1/3 \) is shown in Figure 5.

![Figure 5](image.png)

Figure 5: Schematic of the branches of continuous spectra on the imaginary axis, along with the double zero eigenvalue at the origin, and the nonzero eigenvalues, for \( 0 < p < 1/3 \).

8 Complex \( \mu \)-eigenvalues in the system at infinity

The case where the \( \mu \)-eigenvalues, in the system at infinity when \( \lambda = 0 \), form a complex quartet, has a number of interesting features that are outside the scope of this paper. In this section, some of the issues are highlighted and then encapsulated in an open question.
A schematic of the $\mu-$eigenvalues in this case when $\lambda = 0$ and $\lambda \neq 0$ is shown in Figure 6. When $\lambda = 0$ the four eigenvalues form a complex quartet, symmetric about the origin numbered so that

\[
\mu_1 = -\beta + i\alpha, \quad \mu_2 = \overline{\mu_1}, \quad \mu_3 = \beta + i\alpha, \quad \mu_4 = \overline{\mu_3},
\]  
(8.1)

with $\alpha > 0$ and $\beta > 0$. The basic state $\tilde{Z}(\xi, c)$ will have the asymptotic form

\[
\lim_{\xi \to \infty} e^{2\beta \xi} \tilde{Z}(\xi, c) = (c_1 \cos(\alpha \xi + \nu) + c_2 \sin(\alpha \xi + \nu)) \zeta_1 + (-c_1 \sin(\alpha \xi + \nu) + c_2 \cos(\alpha \xi + \nu)) \zeta_2,
\]  
(8.2)

for some constants $c_1, c_2$ and phase shift $\nu$, with $\zeta_1$ and $\zeta_2$ real valued and satisfying

\[
[B^\infty(c) - (-\beta + i\alpha)J(c)](\zeta_1 + i\zeta_2) = 0.
\]

The ratio of the constants $c_1/c_2$, $c_2 \neq 0$, is fixed on a given solitary wave. Solitary waves of this type are called “oscillatory solitary waves”. The exponential decay assures that these solitary waves are still integrable on the real line, and so the momentum (4.6) will be well defined.

The question is how to compute the three key ingredients in the derivative formula: $dI/dc$, $\Pi$, and $\chi$, and more importantly, how to adjust the proof of the derivative formula in Theorem 6.1.

The calculation of $\chi$ is the least clear as the reality of the asymptotics is used in the definition of $\chi$ in §4.7. On the other hand, with the momentum well defined, $dI/dc$ is in principal computable. In fact, a calculation of the invariants $H$ and $I$ for a class of oscillatory solitary waves is given in Buryak & Akhmediev [15].

The tangent vector $\tilde{Z}_\xi$ will have a similar asymptotic form to (8.2) with different constants, and $a^+$ is any tangent vector transverse to $\tilde{Z}_\xi$ with $a^+ \to 0$ as $\xi \to +\infty$. There is a similar definition of $a^-$ as $\xi \to -\infty$. Hence construction and normalization of $\Pi$ is similar to the real $\mu-$eigenvalue case in §4.3 and §4.6. Indeed, an example of the construction of $a^\pm$ for the case of a homoclinic orbit with complex quartet eigenvalues in the system at infinity, and an explicit construction of the Lazutkin invariant for this case, is given in Gaivao & Gelfreich [23]. There they use the Lazutkin invariant (called “the homoclinic invariant” in [23]) to measure the splitting of separaties in the unfolding of the Hamiltonian Hopf bifurcation.

When $\lambda \neq 0$ the asymptotics (5.14) are defined exactly as before

\[
\lim_{\xi \to +\infty} e^{-\mu_1(\lambda)\xi} u_1(\xi, \lambda) = \zeta_1(\lambda), \quad \lim_{\xi \to +\infty} e^{-\mu_2(\lambda)\xi} u_2(\xi, \lambda) = \zeta_2(\lambda),
\]

\[
\lim_{\xi \to -\infty} e^{-\mu_3(\lambda)\xi} u_3(\xi, \lambda) = \zeta_3(\lambda), \quad \lim_{\xi \to -\infty} e^{-\mu_4(\lambda)\xi} u_4(\xi, \lambda) = \zeta_4(\lambda),
\]  
(8.3)
with the complex vectors $\zeta_j$ defined as in (5.7). But the relation between $u_j(\xi,0)$ and $\hat{Z}_\xi$ and $a^\pm$ will be more complex, something like

\begin{align*}
c_1 u_3(\xi,0) + \overline{c_1} u_4(\xi,0) &= \hat{Z}_\xi \\
c_2 u_3(\xi,0) + \overline{c_2} u_4(\xi,0) &= a^- \\
c_3 w_3(\xi,0) + \overline{c_3} w_4(\xi,0) &= \hat{Z}_\xi \\
c_4 w_3(\xi,0) + \overline{c_4} w_4(\xi,0) &= a^+.
\end{align*}

The number of constants may be reducible, but the form of the limit still complicates the proof of the derivative formula in §6.

The existence of oscillatory solitary waves is a difficult subject itself. Moreover when such a homoclinic orbit does exist, and is transversely constructed, then, by a theorem of Devaney [22], there exists a countable set of multi-pulse homoclinic orbits nearby in parameter space. This complicates the stability question.

On the other hand, the Maslov index of multi-pulse solitary waves, when the system at infinity has complex $\mu -$eigenvalues, is computed in [17], and so $\Pi$ can be obtained in principle as the parity of the Maslov index (subject to orientation and normalization). It is clear from those results that the Maslov index can take any natural number on oscillatory solitary waves with more and more humps, suggesting a complicated stability picture.

The open question, regarding the stability of oscillatory solitary waves, from the viewpoint of this paper, is: does the derivative formula carry over to this case

$$D'(0) = \chi \Pi \frac{dI}{dc},$$

and if so what is the definition of $\chi$?

9 Concluding remarks

There are two assumptions that are worthy of further comment: the simple $\mu -$eigenvalue assumption, and the restriction of four-dimensional phase space.

When the $\mu -$eigenvalues are simple at $\lambda = 0$ they are simple for $\lambda$ near zero and so, with the emphasis of the paper on the derivatives of $D(\lambda)$ at $\lambda = 0$, the assumption of simple eigenvalues is adequate. For global (in $\lambda$) analysis, this assumption can be relaxed by working on exterior algebra spaces, where $\zeta_1 \wedge \zeta_2$ and $\zeta_3 \wedge \zeta_4$ can be constructed to be analytic for all $\lambda \in \Lambda$ (cf. Appendix A for the formulation in terms of exterior algebra), or maximally analytic individual eigenvectors can be constructed [11].

The assumption of a four dimensional phase space in the steady problem (4.8) is sufficient to capture the essence of the theory, and avoids unnecessary complexity. When the phase space has dimension $2n$ with $n > 2$, the theory has a straightforward generalization when all the $\mu -$eigenvalues are real and simple when $\lambda = 0$. In this case the off diagonal terms in the matrix (5.24) are of higher order in $\lambda$ and so the derivative formula is based on the diagonal entries, and moreover it does not matter to which eigenvalue/eigenvector $\hat{Z}_\xi$ is asymptotic to as $\xi \to \pm \infty$. The only change is that $\Pi$ will have to be adapted to the new dimension of the stable and unstable manifolds. The
2\!−\!2 splitting would be replaced by an \( n \!−\! n \) splitting and (4.24) would be replaced by
\[
E^s(\xi, 0) = \text{span}\{\hat{Z}_\xi, a_1^+, \ldots, a_{n-1}^+\} \quad \text{and} \quad E^u(\xi, 0) = \text{span}\{\hat{Z}_\xi, a_1^-, \ldots, a_{n-1}^-\},
\]
with \( \Pi \) replaced by the Lazutkin-Treschev invariant
\[
\Pi = \det \begin{bmatrix}
\Omega(a_1^-, a_1^+) & \cdots & \Omega(a_1^-, a_{n-1}^+)
\vdots & \ddots & \vdots
\Omega(a_{n-1}^-, a_1^+) & \cdots & \Omega(a_{n-1}^-, a_{n-1}^+)
\end{bmatrix},
\]
(cf. Treschev [11] and Chardard & Bridges [16]). Hence, subject to the generalizations required in the Evans function construction, we expect that Theorem 6.1 generalizes with \( \Pi \) replaced by (9.2). The higher-dimensional case is better treated as it arises in applications. The case where the phase space has dimension 8, with a four-four splitting of eigenvalues is studied by Burchell [14]. This case is relevant for the study of the transverse instability of solitary waves in the case of multisymplectic Dirac operators in two space dimensions and time.

**Acknowledgements**

The first author acknowledges support from an EPSRC Doctoral Training Partnership, grant numbers EP/M508160/1 and EP/R513350/1. The second author is grateful to Frédéric Chardard (Université Jean Monnet) for helpful discussions at an early stage of this project.

---

**Appendix**

**A Matrix representation of the Evans function**

For the purposes of this appendix \( V \) is a four-dimensional complex vector space with dual \( V^* \), as the explicit use of a dual space makes some of the calculations in exterior algebra clearer. A basis for \( V \) is \( \{e_1, e_2, e_3, e_4\} \), which is not necessarily the standard basis, and a basis for \( V^* \) is \( \{e_1^*, e_2^*, e_3^*, e_4^*\} \), normalized in the usual way \( \delta_{ij} = \langle e_i^*, e_j \rangle \).

Exterior algebra spaces \( \bigwedge^k(V) \) and \( \bigwedge^k(V^*) \), \( k = 1, 2, 3, 4 \), are defined in the usual way with induced pairings, \( [\cdot, \cdot]_k \).

The volume form on \( V \) is (4.22) and on \( V^* \) it is
\[
\text{vol}^* = e_1^* \wedge e_2^* \wedge e_3^* \wedge e_4^*.
\]
The normalized basis then gives that \( [\text{vol}^*, \text{vol}]_4 = 1 \).

We will also find use for the equivalent bases
\[
V = \text{span}\{\zeta_1, \ldots, \zeta_4\} \quad \text{and} \quad V^* = \text{span}\{J(c)\eta_1, \ldots, J(c)\eta_4\},
\]
(A-1)
where $\zeta_j$ and $\eta_j$ are the eigenvectors defined in (5.7) and (5.9) respectively, with their associated four-forms $\mathcal{V}(c, \lambda)$ in (5.8) and $\mathcal{V}^*(c, \lambda)$ in (5.11).

Although $V$ is a complex vector space, the pairing between $V^*$ and $V$ does not involve conjugation (to be consistent with the definition of the dual vectors $\eta_j$ as in (5.9), where conjugation is built in). An example of the pairing on $V^2$ is

$$[a \wedge b, c \wedge d]_2 = \det \begin{bmatrix} [a, c]_1 & [a, d]_1 \\ [b, c]_1 & [b, d]_1 \end{bmatrix}, \text{ for any } a, b \in V^*, \ c, d \in V,$$

for a decomposable element with obvious extension to arbitrary elements in $V^2$. When reverting back to the notation in the body of the paper, $[\cdot, \cdot]_1$ is replaced by $\langle \cdot, \cdot \rangle$.

The interior product $\iota$ is the adjoint operator associated with the of the wedge product, for example

$$[c \iota (a \wedge b), d]_1 = [a \wedge b, c \wedge d]_2, \quad a, b \in V^*, \ c, d \in V.$$

It is a contraction operator and is sometimes written $c \iota (a \wedge b) = i_c(a \wedge b)$.

The formula (5.21) for the Evans function can be proved using the Hodge star operator as in [9 10]. Here a new proof is given using the interior product. It is a little more general in that it does not require an inner product in its definition.

The starting point is the Evans function in the form (5.15), relative to the orientation (4.22). The interior product is used to transform it as follows, suppressing the arguments for brevity,

$$D(\lambda)\text{vol} = e^{-\tau(c)\lambda}u_1 \wedge u_2 \wedge u_3 \wedge u_4$$

$$= [\text{vol}^*, e^{-\tau(c)\lambda}u_1 \wedge u_2 \wedge u_3 \wedge u_4]_4 \text{vol}$$

$$= [\mathcal{V}^*(c, \lambda), e^{-\tau(c)\lambda}u_1 \wedge u_2 \wedge u_3 \wedge u_4]_4 \mathcal{V}(c, \lambda)$$

$$= [e^{-\tau(c)\lambda}e^{\iota_2}(u_1 \wedge u_2) \iota \mathcal{V}^*(c, \lambda), u_3 \wedge u_4]_2 \mathcal{V}(c, \lambda),$$

using $1 = [\text{vol}^*, \text{vol}]_4 = [\mathcal{V}^*, \mathcal{V}]_4$. Define the left argument in the pairing as

$$W(\xi, \lambda) = e^{-\tau(c)\lambda}u_1(\xi, \lambda) \wedge u_2(\xi, \lambda) \iota \mathcal{V}^* \in \wedge^2(V^*).$$

$W$ is the left-hand side of the formula (5.22). Differentiate it with respect to $\xi$

$$W_\xi = -\tau(c)\lambda e^{-\tau(c)\lambda}u_1(\xi, \lambda) \wedge u_2(\xi, \lambda) \iota \mathcal{V}^* + e^{-\tau(c)\lambda}A^{(2)}(u_1(\xi, \lambda) \wedge u_2(\xi, \lambda)) \iota \mathcal{V}^*,$$

using the definition $A^{(2)}(a \wedge b) := A(a \wedge b) + a \wedge A(b)$ for any $a, b \in V$, and the governing equation for $u_j$ in (5.13). Now look at the second term. It can be split using the identity

$$A^{(2)}(a \wedge b) \wedge c \wedge d + a \wedge b \wedge A^{(2)}(c \wedge d) = \text{Trace}(A)a \wedge b \wedge c \wedge d.$$

Let $a, b \in V$ be arbitrary and pair $a \wedge b \in \wedge^2(V)$ with the operator in the second term on the right-hand side of (A-4),

$$[A^{(2)}(u_1 \wedge u_2) \iota \mathcal{V}^*, a \wedge b]_2 = [\mathcal{V}^*, A^{(2)}(u_1 \wedge u_2) \wedge a \wedge b]_4$$

$$= \text{Trace}(A)[\mathcal{V}^*, u_1 \wedge u_2 \wedge a \wedge b]_4 - [\mathcal{V}^*, u_1 \wedge u_2 \wedge A^{(2)}(a \wedge b)]_4$$

$$= \tau(c)\lambda[(u_1 \wedge u_2) \iota \mathcal{V}^*, a \wedge b]_2 - [(u_1 \wedge u_2) \iota \mathcal{V}^*, A^{(2)}(a \wedge b)]_2$$

$$= \tau(c)\lambda[(u_1 \wedge u_2) \iota \mathcal{V}^*, a \wedge b]_2$$

$$- [(A^{(2)})^T[(u_1 \wedge u_2) \iota \mathcal{V}^*], a \wedge b]_2.$$
Since this holds for any \( a, b \in V \) we get that
\[
A^{(2)}(u_1 \wedge u_2) \wedge V^* = \tau(c)\lambda(u_1 \wedge u_2) \wedge V^* - (A^{(2)})^T \left[(u_1 \wedge u_2) \wedge V^*\right].
\]

Substitute into (A-4)
\[
W_\xi = -\tau(c)\lambda e^{-\tau(c)\lambda \xi}(u_1(\xi, \lambda) \wedge u_2(\xi, \lambda)) \wedge V^* + e^{-\tau(c)\lambda \xi}A^{(2)}(u_1(\xi, \lambda) \wedge u_2(\xi, \lambda)) \wedge V^*
= -\tau(c)\lambda W + \left(\tau(c)\lambda W - (A^{(2)})^T W\right),
\]
or
\[
W_\xi = -(A^{(2)}(\xi, \lambda))^T W. \tag{A-5}
\]
This equation shows that \( W \) satisfies the adjoint equation on \( \wedge^2(V^*) \).

Now write the right hand side of (5.22) as
\[
\Upsilon(\xi, \lambda) = J(c)w_3(\xi, \lambda) \wedge J(c)w_4(\xi, \lambda). \tag{A-6}
\]
Differentiate \( \Upsilon \) and use the governing equation for \( w \) in (5.17),
\[
\Upsilon_\xi = J(c)(w_3)_\xi \wedge J(c)w_4 + J(c)w_3 \wedge J(c)(w_4)_\xi
= J(c)A(\xi, -\lambda)w_3 \wedge J(c)w_4 + J(c)w_3 \wedge J(c)A(\xi, -\lambda)w_4
= (B(\xi, c) + \lambda M)w_3 \wedge J(c)w_4 + J(c)w_3 \wedge (B(\xi, c) + \lambda M)w_4
= -A(\xi, \lambda)^T J(c)w_3 \wedge J(c)w_4 - J(c)w_3 \wedge A(\xi, \lambda)^T J(c)w_4,
\]
which simplifies to
\[
\Upsilon_\xi = -(A^{(2)}(\xi, \lambda))^T \Upsilon. \tag{A-7}
\]
It is clear from (A-5) and (A-7) that \( W \) and \( \Upsilon \) satisfy the same adjoint equation on \( \wedge^2(V^*) \). But to show equality as in (5.22), it is necessary to study their asymptotics. Using the asymptotic properties of \( w_3 \) and \( w_4 \) in (5.20),
\[
\lim_{\xi \to +\infty} e^{(\mu_3 + \mu_4)\xi} \Upsilon(\xi, \lambda) = \lim_{\xi \to +\infty} \left(e^{\mu_3 \xi} J(c)w_3 \right) \wedge \left(e^{\mu_4 \xi} J(c)w_4 \right) = J(c)\eta_3 \wedge J(c)\eta_4.
\]
On the other hand, using \( \tau(c)\lambda = \mu_1 + \mu_2 + \mu_3 + \mu_4 \),
\[
e^{(\mu_3 + \mu_4)\xi} W(\xi, \lambda) = e^{-(\mu_1 + \mu_2)\xi} u_1(\xi, \lambda) \wedge u_2(\xi, \lambda) \wedge V^*.
\]
Taking the limit \( \xi \to +\infty \) and using the asymptotics of \( u_1 \) and \( u_2 \),
\[
\lim_{\xi \to +\infty} e^{(\mu_3 + \mu_4)\xi} W(\xi, \lambda) = \lim_{\xi \to +\infty} e^{-(\mu_1 + \mu_2)\xi} u_1(\xi, \lambda) \wedge u_2(\xi, \lambda) \wedge V^* = \zeta_1 \wedge \zeta_2 \wedge V^*.
\]
This is where the second formula in (5.23) comes into play. It remains to prove that
\[
J(c)\eta_3 \wedge J(c)\eta_4 = \zeta_1 \wedge \zeta_2 \wedge V^*. \tag{A-8}
\]
This identity is not immediately obvious, but a proof can be given using elementary linear algebra. Expand the right-hand side of (A-8) in terms of a basis for \( \wedge^2(V^*) \) using (A-1),
\[
\zeta_1 \wedge \zeta_2 \wedge V^* = c_1 J(c)\eta_1 \wedge J(c)\eta_2 + c_2 J(c)\eta_1 \wedge J(c)\eta_3 + c_3 J(c)\eta_1 \wedge J(c)\eta_4
+ c_4 J(c)\eta_2 \wedge J(c)\eta_3 + c_5 J(c)\eta_2 \wedge J(c)\eta_4 + c_6 J(c)\eta_3 \wedge J(c)\eta_4. \tag{A-9}
\]
A basis for $\bigwedge^2(V)$ is the six-dimensional set $\{\zeta_1 \wedge \zeta_2, \ldots, \zeta_3 \wedge \zeta_4\}$. Now pair both sides of (A-9) with each basis element from $\bigwedge^2(V)$ in turn. One finds that $c_1, \ldots, c_5$ are all zero, and $c_6 = 1$. This proves (5.23), and completes the proof of the formula (5.22).

Now apply the identity $W = \Upsilon$ to the Evans function in (A-2),

$$D(\lambda) = e^{-\tau(c)\lambda\xi} u_1 \wedge u_2 \wedge u_3 \wedge u_4$$

$$= \left[ \Omega(w_3, u_3) \Omega(w_4, u_4) \right] \bigwedge \left[ \Omega(w_2, u_3) \Omega(w_4, u_4) \right]$$

$$= \det \left[ \Omega(w_3, u_3) \Omega(w_3, u_4) \right] \bigwedge \left[ \Omega(w_2, u_3) \Omega(w_4, u_4) \right]$$

thereby completing the proof of Theorem 5.2.

\[ \blacksquare \]

**B  Exact Evans function for the example**

In this appendix the details of the construction of the exact Evans function for the example in §7 are given. When $\lambda \neq 0$ the eigenvalues are

$$\begin{align*}
\mu_1 &= \alpha^2 \lambda - \alpha \sqrt{\alpha^2 \lambda^2 + 4 + 3p} \\
\mu_2 &= \alpha^2 \lambda - \alpha \sqrt{\alpha^2 \lambda^2 + 4} \\
\mu_3 &= \alpha^2 \lambda + \alpha \sqrt{\alpha^2 \lambda^2 + 4} \\
\mu_4 &= \alpha^2 \lambda + \alpha \sqrt{\alpha^2 \lambda^2 + 4 + 3p}.
\end{align*}$$

The $\mu$-eigenvalues $\mu(c, \lambda)$ are simple for all $\lambda \in \Lambda$, satisfying hypothesis (H6). The eigenvectors are

$$\zeta_1 = z_1 \left( \begin{array}{c} 2 \\ 2\mu_1 - \mu_1 \\ 1 \end{array} \right), \quad \zeta_2 = z_2 \left( \begin{array}{c} 2 \\ 2\mu_1 - 4\mu_2 \\ 4 \end{array} \right), \quad \zeta_3 = z_3 \left( \begin{array}{c} 2 \\ 2\mu_3 - 4\mu_3 \\ 4 \end{array} \right), \quad \zeta_4 = z_4 \left( \begin{array}{c} 2 \\ 2\mu_4 - \mu_4 \\ 1 \end{array} \right),$$

where $\varrho_j = \lambda + c\mu_j$. The real numbers $z_1, \ldots, z_4$ are arbitrary.

Now construct the exact solutions, starting with (7.23). The two exact solutions with $\lambda \neq 0$ for $\psi_1$ are

$$\psi_1^\pm = e^{\mp \alpha \sqrt{4 + \alpha^2 \lambda^2}} A^\pm(\xi),$$

with $T = \tanh(\alpha \xi)$ and

$$A^\pm(\xi) = \pm a_0 + a_1 T \pm a_2 T^2 + T^3,$$

where,\[a_0 = -\frac{\alpha^2 \lambda^2}{15\sqrt{4 + \alpha^2 \lambda^2}}, \quad a_1 = 1 + \frac{2}{5} \alpha^2 \lambda^2, \quad a_2 = -\sqrt{4 + \alpha^2 \lambda^2}.\]
These two solutions are associated with the eigenvalues

$$\psi_1^- : \mu_2 = c\alpha^2\lambda - \alpha\sqrt{4 + \alpha^2\lambda^2} \quad \text{and} \quad \psi_1^+ : \mu_3 = c\alpha^2\lambda + \alpha\sqrt{4 + \alpha^2\lambda^2}.$$ 

Similarly, the two exact solutions with $\lambda \neq 0$ for $\psi_2$ are

$$\psi_2^\pm = e^{\mp\alpha\sqrt{4 + 3p + \alpha^2\lambda^2}}b^\pm(\xi),$$

with $T = \tanh(\alpha\xi)$ and

$$B^\pm(\xi) = \pm b_0 + b_1T \pm b_2T^2 + T^3,$$

where

$$b_0 = -\left(\frac{3p + \alpha^2\lambda^2}{15}\right)\sqrt{4 + 3p + \alpha^2\lambda^2}, \quad b_1 = \frac{1}{5}(5 + 6p + 2\alpha^2\lambda^2), \quad b_2 = -\sqrt{4 + 3p + \alpha^2\lambda^2}.$$ 

These two solutions are associated with the eigenvalues

$$\psi_2^- : \mu_1 = c\alpha^2\lambda - \alpha\sqrt{4 + 3p + \alpha^2\lambda^2} \quad \text{and} \quad \psi_2^+ : \mu_4 = c\alpha^2\lambda + \alpha\sqrt{4 + 3p + \alpha^2\lambda^2}.$$ 

To simplify the construction of the Evans function, use the formula (5.15) evaluated at $\xi = 0$,

$$D_E(\lambda) \text{vol} = u_1(0, \lambda) \wedge u_2(0, \lambda) \wedge u_3(0, \lambda) \wedge u_4(0, \lambda).$$

(B-3)

Hypothesis (H7) is implicitly satisfied by using the original, and equivalent, definition of the Evans function.

Using (7.21) and the above expressions for $\psi_j^\pm$ the exact solutions for $u_j(0, \lambda)$ are

$$u_1(0, \lambda) = \begin{pmatrix} -2b_0 \\ \lambda b_0 + (2 - c)(\alpha b_1 - \mu_1 b_0) \\ -b_0 \end{pmatrix}, \quad u_2(0, \lambda) = \begin{pmatrix} -a_0 \\ -\lambda a_0 + (c - 2)(\alpha a_1 - \mu_2 a_0) \\ 2\lambda a_0 + (1 - 2c)(\alpha a_1 - \mu_2 a_0) \end{pmatrix},$$

and

$$u_3(0, \lambda) = \begin{pmatrix} \lambda a_0 + (c - 2)(\alpha a_1 + \mu_3 a_0) \\ -2\lambda a_0 + (1 - 2c)(\alpha a_1 + \mu_3 a_0) \\ 2a_0 \end{pmatrix}, \quad u_4(0, \lambda) = \begin{pmatrix} 2b_0 \\ -b_0 + (2 - c)(\alpha b_1 + \mu_4 b_0) \\ 2b_0 \end{pmatrix}.$$ 

Substituting these expressions into (B-3) leads to

$$D_E(\lambda) \text{vol} = \Gamma(\lambda) \begin{pmatrix} 1 \\ \lambda \\ -2\lambda \end{pmatrix} \wedge \begin{pmatrix} 0 \\ c - 2 \\ 1 - 2c \end{pmatrix} \wedge \begin{pmatrix} 2 \\ 2\lambda \\ -\lambda \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 2c - 1 \\ 2 - c \end{pmatrix},$$

with

$$\Gamma(\lambda) = a_0b_0(2\alpha a_1 + \mu_3 a_0 - \mu_2 a_0)(2\alpha b_1 + \mu_4 b_0 - \mu_1 b_0).$$

The latter wedge product evaluates to $9(c^2 - 1)\text{vol} = -9\alpha^{-2}\text{vol}$ where here $\text{vol}$ is the standard volume form. Therefore

$$D_E(\lambda) = -\frac{9}{\alpha^2}a_0b_0(2\alpha a_1 + \mu_3 a_0 - \mu_2 a_0)(2\alpha b_1 + \mu_4 b_0 - \mu_1 b_0).$$

45
After some simplification we find
\[ D_E(\lambda) = -\frac{36S}{(225)^2}(3 + x^2)(3 + 3\rho + x^2)(3\rho + x^2)(5 - 3\rho - x^2), \]
with \( x^2 = \alpha^2\lambda^2 \) and
\[ S = \sqrt{4 + x^2(4 + 3\rho + x^2)}. \]
To synchronize with the Evans function used in §7 we replace \( D_E(\lambda) \) by its equivalent \( D_E(\lambda) \mapsto 4\chi D_E(\lambda) \),
\[ D_E(\lambda) = -\frac{36S}{(225)^2}(3 + x^2)(3 + 3\rho + x^2)(3\rho + x^2)(5 - 3\rho - x^2), \]
which simplifies to
\[ D_E(\lambda) = \frac{3S}{16(225)^2}\alpha\lambda^2(3 + x^2)(3 + 3\rho + x^2)(3\rho + x^2)(5 - 3\rho - x^2), \quad (B-4) \]
This is the exact Evans function that is used in (7.30).

References

[1] J.W. Alexander, R. Gardner & C.K.R.T. Jones. A topological invariant arising in the stability analysis of traveling waves, J. reine angew. Math. 410 167–212 (1990).

[2] J.W. Alexander & C.K.R.T. Jones. Existence and stability of asymptotically oscillatory double pulses, J. reine angew. Math. 446 49–79 (1994).

[3] I.V. Barashenkov, D.E. Pelinovksy, & E.V. Zemlyanaya. Vibrations and oscillating instabilities of gap solitons, Phys. Rev. Lett. 80 5117–5120 (1998).

[4] M. Beck, G. Cox, C. Jones, Y. Latushkin, K. McQuighan, & A. Sukhtayev. Instability of pulses in gradient reaction-diffusion systems: a symplectic approach, Phil. Trans. Roy. Soc. Lond. A 376 20170187 (2018).

[5] T.B. Benjamin. The stability of solitary waves, Proc. Roy. Soc. Lond. A 328 153–183 (1972).

[6] J. Bona. On the stability theory of solitary waves, Proc. Roy. Soc. Lond. A 344 363–374 (1975).

[7] T.J. Bridges. Canonical multi-symplectic structure on the total exterior algebra bundle, Proc. Roy. Soc. Lond. A 462 1531–1551 (2006).

[8] T.J. Bridges & G. Derks. Unstable eigenvalues and the linearization about solitary waves and fronts with symmetry, Proc. Roy. Soc. London A 455 2427–2469 (1999).

[9] T.J. Bridges & G. Derks. Hodge duality and the Evans function, Phys. Lett. A 251 363–372 (1999).

[10] T.J. Bridges & G. Derks. The symplectic Evans matrix, and the instability of solitary waves and fronts, Arch. Rat. Mech. Anal. 156 1–87 (2001).
[11] T.J. Bridges & G. Derks. Constructing the symplectic Evans matrix using maximally analytic individual vectors, Proc. Royal Soc. Edin. A 133 505–526 (2003).

[12] T.J. Bridges, P.E. Hydon, & J.K. Lawson. Multisymplectic structures and the variational bicomplex, Math. Proc. Camb. Phil. Soc. 148 159–178 (2010).

[13] T.J. Bridges & A. Mielke. A proof of the Benjamin-Feir instability, Arch. Rat. Mech. Anal. 133 145–198 (1995).

[14] T.J. Burchell. Nonlinear wave equations, Clifford Analysis, and the stability of solitary waves, PhD Thesis, University of Surrey (2022).

[15] A.V. Buryak & N.N. Akhmediev. Stability criterion for stationary bound states of solitons with radiationless oscillating tails, Phys. Rev. E 51 3572–3578 (1995).

[16] F. Chardard & T.J. Bridges. Transversality of homoclinic orbits, the Maslov index and the symplectic Evans function, Nonlinearity 28 77–102 (2015).

[17] F. Chardard, F. Dias & T.J. Bridges. On the Maslov index of multi-pulse homoclinic orbits, Proc. Roy. Soc. Lond. A 465 2897–2910 (2009).

[18] P. Christodoulides, R. Grimshaw & C. Demetriades. Three-fluid system short-wave instability and gap solitons, IAENG Int. J. Appl. Math. 41 IJAM-41-3-10 (2011).

[19] M. Chugunova & D.E. Pelinovsky. Block-diagonalization of the symmetric first-order coupled-mode systems, SIAM J. Appl. Dyn. Sys. 5 66–83 (2006).

[20] W.A. Coppel. Stability and Asymptotic Behaviour of Differential Equations, D.C. Heath and Co: Boston (1965).

[21] G. Cox, C.K.R.T. Jones, Y. Latushkin, & A. Sukhaiyev. Morse and Maslov indices for multidimensional Schrödinger operators with matrix-valued potentials, Trans. Amer. Math. Soc. 368 8145–8207 (2016).

[22] R.L. Devaney. Homoclinic orbits in Hamiltonian systems, J. Diff. Eqns. 21 431–438 (1976).

[23] J.P. Gaivao & V. Gelfreich. Splitting of separatrices for the Hamiltonian-Hopf bifurcation with the Swift-Hohenberg equation as an example, Nonlinearity 24 677–698 (2011).

[24] V.G. Gelfreich & V.F. Lazutkin. Splitting of separatrices: perturbation theory and exponential smallness, Russ. Math. Surveys 56 499–558 (2001).

[25] M. Grillakis, J. Shatah, & W. Strauss. Stability theory of solitary waves in the presence of symmetry, I, J. Func. Anal. 74 160–197 (1987).

[26] M. Grillakis, J. Shatah, & W. Strauss. Stability theory of solitary waves in the presence of symmetry, II, J. Func. Anal. 94 308–348 (1990).

[27] R. Grimshaw. Models for long-wave instability due to a resonance between two waves, in Trends in Appl. of Math. to Mech., (Eds. G. Iooss, O. Gues & A. Nouri), Chapman & Hall/CRC Monographs 106 183–192 (2000).
[28] R. Grimshaw & P. Christodoulides. *Short-wave instability in a three-layer stratified shear flow*, Quart. J. Mech. Appl. Math. **54**, 375–388 (2001).

[29] R. Grimshaw & Y. Skyrnnikov. *Long-wave instability in a three-layer stratified shear flow*, Stud. Appl. Math. **108**, 77–88 (2002).

[30] T. Kapitula & K. Promislow. *Spectral and Dynamical Stability of Nonlinear Waves*, Appl. Math. Sci. **185**, Springer-Verlag: New York (2013).

[31] Y. Latushkin & S. Sukhtaiev. *The Maslov index and the spectra of second order elliptic operators*, Adv. in Maths. **329**, 422–486 (2018).

[32] J.H. Maddocks & R.L. Sachs. *Constrained variational principles and stability in Hamiltonian systems*, IMA Volumes in Mathematics and its Applications **63**, 231–264 (1995).

[33] J.R. Magnus & H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Third Edition, John Wiley & Sons: Chichester (1999).

[34] J.E. Marsden. *Lectures on Mechanics*, LMS Lecture Series **174**, Cambridge University Press (1992).

[35] S. Morita. *Geometry of Differential Forms*, AMS: Providence (2001).

[36] S.J. Orfandis & R. Wang. *Soliton solutions of the massive Thirring model*, Phys. Lett. B **57**, 281–283 (1975).

[37] J.A. Pava. *Nonlinear Dispersive Equations: Existence and Stability of Solitary and Periodic Travelling Wave Solutions*, Mathematical Surveys & Monographs **156**, Amer. Math. Soc: Providence (2009).

[38] R.L. Pego & M.I. Weinstein. *Eigenvalues, and instabilities of solitary waves*, Phil. Trans. Roy. Soc. London A **340**, 47–94 (1992).

[39] D.E. Pelinovsky. *Survey on global existence in the nonlinear Dirac equations in one spatial dimension*, RIMS Kōkyūroku Bessatsu, RIMS, Kyoto, **B26**, 37-50 (2011).

[40] D. Sugny, A. Picozzi, S. Lagrange, & H.R. Jauslin. *Role of singular tori in the dynamics of spatiotemporal nonlinear wave systems*, Phys. Rev. Lett. **103**, 034102 (2009).

[41] D.V. Treschev. *Separatrix splitting from the point of view of symplectic geometry*, Matematicheskie Zametki **61**, 890–906 (1997).

[42] M.I. Weinstein. *Liapunov stability of ground states of nonlinear dispersive evolution equations*, Comm. Pure Appl. Math. **39**, 51–68 (1986).