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 stability exponent. The main result is the introduction of a new factor, denoted $\Pi$, in the Pego-Weinstein derivative formula

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

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. 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 Hamiltonian PDEs. 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. The new result is already present when the homoclinic orbit representation of the solitary wave lives in a four dimensional phase space, and so the theory is presented for this case, with the generalization to arbitrary dimension sketched.

1 Introduction

The stability of solitary waves of Hamiltonian partial differential equations (PDEs) can be approached many ways. One powerful approach 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. This approach goes back to Benjamin [4] and Bona [5] in the context of the Korteweg-de Vries equation, and was developed into a general and powerful approach for a class of Hamiltonian PDEs with one or more constraints by Grilakis, Shatah, & Strauss [17, 18] (hereafter GSS). There has been a vast amount of work in this direction (e.g. see Chapter 5 in Kapitula & Promislow [22] and references therein). One 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

\[1\text{Email: T.Burchell@surrey.ac.uk}\]
\[2\text{Email: T.Bridges@surrey.ac.uk}\]
\( I \), and a single Lagrange multiplier, the speed of the solitary wave denoted by \( c \), the condition is
\[
\frac{dI}{dc} > 0 \quad \Rightarrow \quad \text{solitary wave is a minimizer.} \tag{1.1}
\]
This condition is useful since \( I(c) \) is a property of the basic state and so, in principle, 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 linearized stability problem for solitary waves while incorporating the Hamiltonian structure. The seminal paper in this direction is PEGO & WEINSTEIN \[28\] (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. It was already known at that time that a novel and highly successful way to approach the linear stability problem was to use the Evans function (ALEXANDER, GARDNER & JONES \[1\]). The Evans function, denoted by \( D(\lambda) \) where \( \lambda \) is the stability exponent, is a complex analytic function whose zeros are eigenvalues of the linearized (spectral) stability problem. By combining the Evans function with the Hamiltonian structure, and the energy-momentum characterization of solitary waves, PW were able to prove that the Evans function has the following properties
\[
D(0) = 0, \quad D'(0) = 0, \quad D''(0) = \frac{dI}{dc}. \tag{1.2}
\]
The derivative of \( I \) in (1.1) appears in this formula in a natural way, but the GSS spectral condition is not used in any way 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 \( D(\lambda) \rightarrow 1 \) as \( \lambda \rightarrow +\infty \) 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 \[7, 9, 10\] extended the PW theory and showed that there is an additional factor in the second derivative in (1.2)
\[
D''(0) = \Pi_{\text{BD}} \frac{dI}{dc}. \tag{1.3}
\]
When \( D(\lambda) \rightarrow 1 \) as \( \lambda \rightarrow +\infty \) along the real axis then it is the negativity of the full product that gives existence of an unstable eigenvalue. The factor \( \Pi_{\text{BD}} \) is calculated independently of the derivative of \( I \) and is not just a scale factor. An explicit formula was found for the factor \( \Pi_{\text{BD}} \) but the presence of symmetry, other than translation invariance in space, was an essential part of the proof in \[7, 9, 10\]. Moreover the theory relied on the “system at infinity” having only one positive and one negative real (spatial) eigenvalue when \( \lambda = 0 \) (see \[4.2\] and \[5\] for the definition of “spatial eigenvalue” and “system at infinity”). Several examples were given with \( \Pi_{\text{BD}} \) taking both positive and negative values, showing that the 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 expression for the second factor is found in the form
\[
D''(0) = 2\Pi \frac{dI}{dc}. \tag{1.4}
\]
The factor Π is associated with the intersection of the stable and unstable manifolds which form
the solitary wave, characterized as a homoclinic orbit. The 2 is added for convenience, giving

\[ D(\lambda) = \Pi \lambda^2 + \mathcal{O}(\lambda^3) \] as \( \lambda \to 0 \).

Explicitly, the new factor is

\[ \Pi = \Omega(a^+, a^-), \tag{1.5} \]

where \( a^+ \) and \( a^- \) are \( x \)-dependent tangent vectors to the oriented stable and unstable manifolds
respectively and \( \Omega \) is a symplectic form associated with a \( c \)-dependent spatial symplectic structure
(defined in §4). \( \Pi \) is a symplectic invariant and an intrinsic property of the homoclinic orbit that
represents the solitary wave. The importance of \( \Omega(a^+, a^-) \) as a symplectic invariant of homoclinic
orbits was discovered by Lazutkin [16], 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 [12]. (When the dimension of the stable and unstable manifolds is greater than two
this formula expands to be the determinant of a matrix of symplectic forms [31, 12].) It is
proved in [12] that the homoclinic orbit is transversely constructed if and only if this symplectic
intersection index is nonzero. 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 \( \lambda \to 0 \).

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, and Chardard & Bridges [12] prove that in gradient systems
the first derivative of the Evans function can be expressed in terms of transversality. However, in
both cases there is no second factor like \( \frac{dI}{dc} \) in (1.4).

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 [28] is

\[ u_t = \mathcal{J} \nabla H(u), \quad u \in \mathcal{X}, \tag{1.6} \]

for some function space \( \mathcal{X} \), where \( \mathcal{J} : \mathcal{X}^* \to \mathcal{X} \) is the co-symplectic (or Poisson) operator, \( u \) is
scalar-valued, and \( H : \mathcal{X} \to \mathbb{R} \) is the Hamiltonian function. However, reduction of (1.6) to a steady
problem is an ODE and a finite dimensional Hamiltonian system. That is; there is a second hidden
symplectic structure in (1.6). In principle it is obtained via Legendre transform of the stationary
system relative to a moving frame

\[ \mathcal{FL}(\nabla H - c\mathcal{J}^{-1}u_\xi), \quad \xi = x + ct, \]

when the inverse of \( \mathcal{J} \) exists, and \( \mathcal{FL} \) denotes Legendre transform. The outcome of this Legendre
transform is a second symplectic operator and a finite-dimensional Hamiltonian system. Denote
the second “spatial symplectic operator” by \( \mathcal{K} \). 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 \( \mathcal{J} \),
and new coordinates, leads to a formulation of the Hamiltonian PDE in terms of multisymplectic structure \[7, 9\]. The canonical form for a multisymplectic Hamiltonian PDE \[11\] is

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

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 \(\mathcal{J}\) in \(1.6\). Steady solutions \(Z(x, t) = \hat{Z}(\xi), \xi = x + ct\), are orbits of the finite-dimensional Hamiltonian system

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

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 \[6, 7, 11, 9, 10\]. However, we go one step further in this paper and introduce an abstract class of multisymplectic Hamiltonian PDEs. Given an arbitrary 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 \textit{multisymplectic Dirac operators}. This class of Hamiltonian PDEs includes as special cases the coupled mode equation which appears in fluid dynamics \[13, 19, 20, 21\] and optics \[30, 14, 3\], the massive-Thirring model \[27\], and a class of coupled nonlinear wave equations.

Solitary waves are \textit{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

\[
D''(0) = 2(-1)^{\text{Morse}} \frac{dI}{dc},
\]

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 is required.

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 \[12\] we can go one step further and relate the new characterization of \(\Pi\) to the Maslov index of the solitary wave

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

where in this case the Maslov index of the solitary wave is defined using the Souriau characterization (cf. \S 9 of \[12\]). Solitary waves, with exponential decay at infinity, always have a well-defined Maslov index, but may not have a well-defined Morse index.
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. Finally in the concluding remarks Section 8 some generalizations are discussed.

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 called a relative equilibrium and is of the form $u(x,t) := \hat{u}(x+ct)$; that is, a travelling wave solution is a relative equilibrium. Symplectic 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).

In this section one-parameter relative equilibria (RE) of finite-dimensional Hamiltonian systems are studied. The abstract structure is the same as that of the solitary wave stability problem with the Evans function replaced by an elementary characteristic function, and it shows how the product structure of $D''(0)$ arises naturally. The group is simplified to the compact group $S^1$.

Consider a standard finite-dimensional Hamiltonian system on $\mathbb{R}^{2n}$:

$$ M Z_t = \nabla H(Z), \quad Z \in M := \mathbb{R}^{2n}, \quad M = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}, \quad (2.1) $$

where $H : M \to \mathbb{R}$ is a given smooth function. The system (2.1) is assumed to be symmetric. In particular, it has an orthogonal action, $G_\theta$, of $S^1$ on $M$ satisfying

$$ G_\theta^T M G_\theta = M \quad \text{and} \quad H(G_\theta Z) = H(Z) \quad \forall \theta \in S^1. \quad (2.2) $$

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

$$ M g(Z) = \nabla I(Z). \quad (2.3) $$

The existence of $I$ follows since $M$ and $g$ commute and their product is symmetric.

Now, suppose there exists a family of RE of the above system of the form

$$ \tilde{Z}(t) = G_{\theta(t)} U \quad \text{with} \quad \theta(t) = ct + \theta^o. \quad (2.4) $$

Substitution of this form into the governing equation gives the following characterisation of RE: $U \in M$ can be characterised as a critical point of $H$ on level sets of the functional $I$, with $c$ as a Lagrange multiplier,

$$ \nabla H(U) = c \nabla I(U) \quad \text{and} \quad I(U) = I_0, \quad (2.5) $$

5
where $I_0$ is some specified real number. A RE in the family is said to be non-degenerate when $rac{dI}{dc} \neq 0$ when $I$ is evaluated on the family $U(c)$.

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

$$M \dot{Z}_t = D^2 H(\hat{Z}(t)) Z,$$

(2.6)

with $\hat{Z}(t)$ defined in (2.4)-(2.5). However, it follows from the invariance of $H$ that

$$D^2 H(G_{\theta(t)} U) = G_{\theta(t)} D^2 H(U) G_{\theta(t)}^T.$$

Therefore, the substitution $Z(t) = G_{\theta(t)} W(t)$ reduces (2.6) to the constant coefficient ODE:

$$MW_t = L(U, c) W,$$

(2.7)

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

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

(2.8)

then we have the following sufficient condition for instability: *if there exists a $\lambda \in \mathbb{C}$ with $\text{Re}(\lambda) > 0$ and $D(\lambda) = 0$, the RE (2.7) is linearly (spectrally) unstable.* $D(\lambda)$ is a finite-dimensional analogue of the Evans function.

The operator $L$ has a zero eigenvalue with the tangent vector to the RE as eigenvector; that is, $Lg(U) = 0$, and it is assumed that the zero eigenvalue of $L$ is simple. The derivative of $U$ with respect to $c$ satisfies $LU_c = Mg(U)$. Normalize the length of these vectors and define

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

Then they satisfy

$$L\zeta_1 = 0 \quad \text{and} \quad L\zeta_2 = M\zeta_1.$$

(2.9)

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

**Theorem 2.1.** Suppose a smooth family of RE exists parameterized by $c$, and suppose $L$ has a simple zero eigenvalue. Then the characteristic function $D(\lambda)$ has the following derivatives at the origin

$$D(0) = 0, \quad D'(0) = 0, \quad D''(0) = 2\mu(L) \frac{dI}{dc},$$

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

$$D''(0) = 2(-1)^{\text{Morse}} \frac{dI}{dc},$$

confirming (1.9) in the introduction.

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

$$D(0) = \det[L] = 0.$$
Differentiate $D(\lambda)$ using the formula for the derivative of a determinant

$$D'(\lambda) = -\text{Tr}\left( (L - \lambda M)^\# M \right) \Rightarrow D'(0) = -\text{Tr}(L^\# M), \quad (2.10)$$

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

$$L^\# = \mu(L)\zeta_1\zeta_1^T \quad \text{with} \quad \mu(L) = \prod_{j=2}^{2n} \mu_j. \quad (2.11)$$

$\mu(L)$ 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 [25]. Substitute $L^\#$ into (2.10),

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

since $M$ is skew symmetric. For $D''(0)$, differentiate (2.10)

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

The adjugate is defined by

$$(L - \lambda M)(L - \lambda M)^\# = (L - \lambda M)^\# (L - \lambda M) = D(\lambda)I, \quad (2.12)$$

where $I$ is the identity on $\mathbb{R}^{2n}$. Now differentiate (2.12) with respect to $\lambda$, set $\lambda$ to zero, and define

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

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

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

with skew-symmetry following from commutativity in (2.12). Combining (2.9), (2.11) and skew-symmetry of $L$ gives

$$\dot{L} = \mu(L)(\zeta_2\zeta_1^T - \zeta_1\zeta_2^T).$$

Substitute into $D''(0)$

$$D''(0) = -\mu(L)\text{Tr}\left( (\zeta_2\zeta_1^T - \zeta_1\zeta_2^T) M \right)$$

$$= -\mu(L) \left( \langle \zeta_1, M\zeta_2 \rangle - \langle \zeta_2, M\zeta_1 \rangle \right)$$

$$= 2\mu(L) \langle \zeta_2, M\zeta_1 \rangle$$

$$= 2 \frac{\mu(L)}{\|g(\mathcal{U})\|^2} \frac{dI}{dc},$$

with the last expression following from (2.3) and

$$\frac{dI}{dc} = (\nabla I(\mathcal{U}), \mathcal{U}_c) = (Mg(\mathcal{U}), \mathcal{U}_c) = \|g(\mathcal{U})\|^2 \langle M\zeta_1, \zeta_2 \rangle.$$
Scaling $D(\lambda)$ by a positive constant then completes the proof. ■

**Corollary.** When $(-1)^{Morse} \frac{dI}{dc} < 0$ 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 2.1 connects $dI/dc$ to the spectral problem and is a finite dimensional version of PW [28] with full generality of the second factor. The connection between $dI/dc$ and critical point type for RE appears in the literature from various perspectives (e.g. Maddocks & Sachs [24] and references therein), and proofs in infinite dimensions are given by Maddocks [23] and Vogel [32], and when the Morse index is unity it is an elementary example of the theory in GSS [17, 18].

It is clear from this result that the appearance of a second factor in the formula for $D''(0)$ is natural in the stability analysis of RE. In PW the second factor was always unity. In [7, 9] the factor was determined using a symmetry argument. We now proceed to develop a theory for the second factor via a transversality argument, for a general class of solitary wave solutions, where the only symmetry is the translation invariance in $x$.

Before proceeding with that proof, the next section develops an abstract class of Hamiltonian PDEs starting with a pseudo-Riemannian manifold. The reader interested only in the general class of PDEs that is taken as a starting point, and not where they might come from, can skip to §4.

### 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 [7, 9, 10] 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 an arbitrary pseudo-Riemannian manifold.

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 of $M$. The cotangent bundle hosts a canonical one form $\mathbf{p} \cdot dq$ with associated functional

$$\int_{t_1}^{t_2} \mathbf{p} \cdot \mathbf{q}_t \, dt . \quad (3.1)$$

The first variation of this functional with fixed endpoints generates the operator

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

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

$$\mathbf{J} \frac{d}{dt} \begin{bmatrix} q \\ p \end{bmatrix} = \nabla H ,$$
and secondly it is a one-dimensional “Dirac operator”

\[
\mathbf{J} \frac{d}{dt} \circ \mathbf{J} \frac{d}{dt} = -\mathbf{I}_{2n} \otimes \frac{d^2}{dt^2}.
\]

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 \( M \), with constant signature metric. In the applications we have in mind a flat manifold is sufficient and by congruence transformation the metric can be assumed to be in standard diagonal form. Hence the starting point is the pseudo-Riemannian vector space \( M = \mathbb{R}^{q,p} \), with \( q + p = m \), with metric

\[
g(u,v) := \langle R u, v \rangle,
\]

where \( \langle \cdot, \cdot \rangle \) is a standard inner product and

\[
\mathcal{R} = \text{diag}(1, \ldots, 1, -1, \ldots, -1).
\]

This metric induces a metric on \( T^*_x M \) and on each of the spaces \( \bigwedge^k (T^*_x M) \). The induced metrics are denoted by

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

with \( g(u,v) \equiv [u,v]_1 \), where \( \mathcal{A}^k(M) \) is the space of differential \( k \)-forms on \( M \) (mappings from \( M \) into \( \bigwedge^k (T^*_x M) \)). Concatenating these spaces gives the total exterior algebra (TEA) bundle denoted by \( \mathcal{A}(M) := \bigcup_{k=0}^n \mathcal{A}^k(M) \).

There is a natural differential 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, \( d \) is an exterior derivative, and \( Z = (u^{(0)}, \ldots, u^{(n)}) \in \mathcal{A}(M) \). The differential form \( [3.5] \) is a generalisation of the canonical form \( \mathbf{p} \cdot d \mathbf{q} \) on the cotangent bundle in \( [3.1] \). The form \( [3.3] \) was introduced in [6] for the case of a positive definite metric, which generates an elliptic partial differential operator (PDO). Here the case of indefinite metric is considered and it generates a number of interesting new features, in addition to generating a hyperbolic PDO. It is this hyperbolic PDO that is the backbone of the nonlinear wave equations of interest here.

Let \( D \) be an open subset of \( M \) with coordinates \( x = (x_1, \ldots, x_n) \) and volume form \( \text{vol} = dx_1 \wedge \cdots \wedge dx_n \). The form \( [3.5] \) in coordinates is

\[
\Theta(Z) = \sum_{k=1}^n [u^{(k)}, du^{(k-1)}]_k \text{vol}.
\]
Proposition 3.1. Let \( W = (w^{(0)}, \ldots, w^{(n)}) \in \mathcal{A}(M) \) be an arbitrary smooth variation on \( \mathcal{D} \) with \( W \) vanishing on \( \partial \mathcal{D} \). Then
\[
\frac{d}{ds} \int_{\mathcal{D}} \Theta(Z + sW) \bigg|_{s=0} = \int_{\mathcal{D}} \left[ \Theta(Z + W) - \Theta(Z) - \Theta(W) \right]
= \int_{\mathcal{D}} \sum_{k=1}^{n} \left( [u^{(k)}, dw^{(k-1)}]_k + [w^{(k)}, du^{(k-1)}]_k \right) \text{vol}
= \int_{\mathcal{D}} \langle \mathcal{J}_\partial Z, W \rangle \text{vol} := \int_{\mathcal{D}} \langle \mathcal{R}^{(r)} \mathcal{J}_\partial Z, W \rangle \text{vol}. \tag{3.7}
\]

Proof. The first two lines are proved by direct calculation using (3.5) and (3.6). The third line is proved by using the induced pseudo-inner product on the total exterior algebra bundle
\[
\langle \langle Z, W \rangle \rangle = \sum_{k=0}^{n} [u^{(k)}, v^{(k)}]_k = \langle \mathcal{R}^{(r)} Z, W \rangle, \quad Z, W \in \mathcal{A}(M), \tag{3.8}
\]
where \( \langle \cdot, \cdot \rangle \) denotes the standard Euclidean inner product on \( \mathbb{R}^r \), with \( r = 2^n \), and \( \mathcal{R}^{(r)} \) is the representation of \( \mathcal{R} \) on \( \mathcal{A}(M) \). The coordinate-free representation of the PDO \( \mathcal{J}_\partial \) is,
\[
\mathcal{J}_\partial Z = \begin{pmatrix} 0 & \delta & 0 & 0 & \cdots & 0 \\ d & 0 & \delta & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & d & 0 & \delta \\ 0 & \cdots & 0 & 0 & d & 0 \end{pmatrix} \begin{pmatrix} u^{(0)} \\ u^{(1)} \\ \vdots \\ u^{(n-1)} \\ u^{(n)} \end{pmatrix}. \tag{3.9}
\]
The right-hand side is generated using the identity \[26\]
\[
\int_{\mathcal{D}} \left( [\beta, d\alpha]_{k+1} - [\alpha, \delta\beta]_k \right) \text{vol} = \int_{\partial \mathcal{D}} \alpha \wedge \star \beta,
\]
for any \( k \)-form \( \alpha \) and \((k+1)\)-form \( \beta \), where \( d \) is the exterior differential and \( \delta \) is the codifferential. When \( \alpha \) vanishes on \( \partial \mathcal{D} \) it simplifies to
\[
\int_{\mathcal{D}} [\beta, d\alpha]_{k+1} \text{vol} = \int_{\mathcal{D}} [\alpha, \delta\beta]_k \text{vol}. \tag{3.10}
\]
Applying this formula then completes the proof. \( \blacksquare \)

The operator \( \mathcal{J}_\partial \) is self-adjoint with respect to the inner product \( \langle \langle \cdot, \cdot \rangle \rangle \) but not with respect to the standard Euclidean inner product on \( \mathbb{R}^r \). However in the analysis it may be convenient to use the operator \( \mathcal{R}^{(r)} \mathcal{J}_\partial \) rather than \( \mathcal{J}_\partial \).

The connection with the classical idea of a Dirac operator becomes apparent when \( \mathcal{J}_\partial \) is expressed in local coordinates.

Proposition 3.2. The operator \( \mathcal{J}_\partial \) in coordinates is
\[
\mathcal{J}_\partial = \sum_{j=1}^{n} \mathcal{J}_j \frac{\partial}{\partial x_j},
\]
the matrices \( \{J_1, \ldots, J_n\} \) are elements of the Clifford algebra \( C^{\ell_{q,p}} \), and

\[
J_\partial \circ J_\partial = -\sum_{j=1}^{n} R_{jj} \frac{\partial^2}{\partial x_j^2} \otimes I_r,
\]

where \( R_{jj} \) are the diagonal entries in \((3.3)\).

**Proof.** The proof follows by writing out coordinate expressions for \( d \) and \( \delta \) in \((3.9)\) and relating the resulting matrices to the Clifford algebra identity

\[
J_i J_j + J_j J_i = -2 R_{ij} I_r.
\]

\[\blacksquare\]

**Remark.** It can be shown that each \( R^{(r)} J_j \) defines an independent symplectic vector space, but that level of detail will not be required here. The two symplectic structures in the case \( M = \mathbb{R}^{1,1} \) are given explicitly below.

By introducing a scalar-valued function \( S : \mathcal{A}(M) \to \mathbb{R} \) into the functional \((3.7)\),

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

and taking the first variation of the functional

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

a nonlinear Dirac operator is generated

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

where \( \nabla S \) is a gradient with respect to the standard Euclidean inner product on \( \mathbb{R}^r \). 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 \([11]\). Indeed it is a new class of multisymplectic Hamiltonian PDEs, the new property being the fact that the symplectic operators are a product of the induced metric times each of the generators of the Clifford algebra. It is this abstract class of Hamiltonian PDEs which feeds into the theory of solitary waves and their linear stability.

### 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. The generated Dirac PDO is a perfect model for the coupled mode equation and the massive Thirring model. 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,

\[
u = u_1 dt + u_2 dx \quad \text{and} \quad v := vt \wedge dx,
\]
where, to simplify notation, $v$ is both a form and a coordinate. The PDO in this case acting on $Z \in A(M)$ is

$$J_\partial Z = \begin{pmatrix} 0 & \delta & 0 \\ d & 0 & \delta \\ 0 & d & 0 \end{pmatrix} \begin{pmatrix} \phi \\ u \\ v \end{pmatrix}.$$ 

and it can be expressed coordinates as

$$J_\partial = J_1 \partial_t + J_2 \partial_x$$

with

$$J_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$ (3.15)

The pair $\{J_1, J_2\}$ generates the Clifford algebra $\mathcal{C}_\ell_{1,1}$,

$$J_i J_j + J_j J_i = -2 \mathcal{R}_{ij} I_4.$$ 

The Dirac property and the connection with the d’Alembertian is

$$J_\partial \circ J_\partial = (J_1 \partial_t + J_2 \partial_x)^2 = J_1^2 \partial_{tt} + (J_1 J_2 + J_2 J_1) \partial_{tx} + J_2^2 \partial_{xx} = - (\partial_{tt} - \partial_{xx}) \otimes I_4.$$ 

Introducing a scalar-valued function $S : A(M) \to \mathbb{R}$, a nonlinear Dirac equation is generated in the canonical form

$$J_\partial Z = \mathcal{R}^{(4)} \nabla S(Z), \quad Z \in A(M).$$ (3.16)

The induced metric in this case is

$$\mathcal{R}^{(4)} = \text{diag}(1, 1, -1, -1) = [+1] \oplus \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \oplus [-1].$$ (3.17)

This form follows by constructing the induced metric on each of the vector spaces $\Lambda^j$ and then concatenating. The space $\Lambda^4(\mathbb{R}^{1,1})$ 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

$$M := \mathcal{R}^{(4)} J_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad K := \mathcal{R}^{(4)} J_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$ (3.18)

The system

$$M Z_t + K Z_x = \nabla S(Z), \quad Z \in A(M),$$ (3.19)

is then in standard form for a multisymplectic Hamiltonian PDE, and the two operators $M$ and $K$ define independent symplectic vector spaces.
3.2 The coupled mode equation

The coupled-mode equation (CME) which appears in fluid mechanics \[13, 19, 20, 21\] and optics \[30, 14, 8\] can be characterised as a multisymplectic Dirac operator on \(A(\mathbb{R}^{1,1})\) in the form (3.16).

In the literature, the CME is represented in complex-amplitude form

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

(3.20)

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

\[
\mathbf{w} = (w_1, w_2) := (\phi, u_1) \quad \text{and} \quad \mathbf{v} = (v_1, v_2) := (u_2, v).
\]

The system (3.20) is transformed using

\[
\begin{align*}
A &= A_1 + \mathbf{i} A_2 = w_1 - v_2 + \mathbf{i}(w_2 - v_1), \\
B &= B_1 + \mathbf{i} B_2 = w_1 + v_2 + \mathbf{i}(w_2 + v_1).
\end{align*}
\]

In these coordinates the CME becomes

\[
\begin{pmatrix}
0 & -1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0
\end{pmatrix}
\begin{pmatrix}
w_1 \\
w_2 \\
v_1 \\
v_2
\end{pmatrix}
+ \begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & -1 \\
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
w_1 \\
w_2 \\
v_1 \\
v_2
\end{pmatrix}
= \mathcal{R}^{(4)} \begin{pmatrix}
\partial S/\partial w_1 \\
\partial S/\partial w_2 \\
\partial S/\partial v_1 \\
\partial S/\partial v_2
\end{pmatrix},
\]

(3.21)

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

\[
\mathbf{J}_1 Z_t + \mathbf{J}_2 Z_x = \mathcal{R}^{(4)} \nabla S(Z),
\]

(3.22)

using (3.15) with \(\mathcal{R}^{(4)} = \text{diag}(1,1,-1,-1)\).

A special case of (3.20) 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. \[30\]). An even more special case is the massive Thirring model (MTM) where \(\tau = \mu = 0\),

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

(3.23)

The transformed system for MTM is (3.21) with

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

(3.24)

3.3 Coupled second-order nonlinear wave equations

The pair of coupled second order nonlinear wave equations

\[
\phi_{tt} - \phi_{xx} + V_\phi = 0 \quad \text{and} \quad v_{tt} - v_{xx} - V_v = 0,
\]

(3.25)
where \( V(\phi, v) \) is a given smooth function, can also be transformed to the canonical form (3.21). 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 (3.25) are represented by (3.19).

4 Solitary wave solutions and linearization

The canonical 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 scalar-valued function, \( \nabla \) the gradient on \( \mathbb{R}^4 \), and \( M, K \) are \( 4 \times 4 \) skew-symmetric matrices. This is hypothesis (H1).

The form (4.1) includes the multisymplectic Dirac operators in §3.1 as a special case. The abstract form (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 [7, 9].

The only other hypothesis needed on this system is that \( J(c) := K + cM \)

(4.2)

is invertible for an open set of \( c \) values. This is hypothesis (H2). With this assumption \((\mathbb{R}^4, \Omega)\) is a symplectic vector space with

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

Since \( M \) is invertible, the PDE can be written in evolution form

\[
Z_t + CZ_x = M^{-1}\nabla S(Z), \quad Z(x, 0) = Z_0(x),
\]

with \( C = M^{-1}K \). When \([M, K] = 0\) 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 [29] and references therein). However, well-posedness is not required for the theory in this paper. Indeed, the example PDE to which the theory is applied in §7 is not well posed.

4.1 Solitary wave solutions

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

\[
T_sZ(x, t) = Z(x + s, t), \quad \forall s \in \mathbb{R},
\]
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

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

(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. The solitary wave is a solution of the ODE

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

(4.6)

This ODE can be characterized as a critical point problem. Let

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

(4.7)

The functional $I(\hat{Z})$ is called the momentum of the solitary wave as it is the conserved functional associated, via Noether’s Theorem, to the $x$–translation symmetry of (4.1). The operator $M$ appears in both $I(\hat{Z})$ and the governing equation (4.1) and this connection will be useful in connecting $dI/dc$ to the Evans function.

Solitary wave solutions correspond to critical points of $H(\hat{Z})$ restricted to level sets of the function $I(\hat{Z})$, with Lagrange necessary condition $\delta(H - cI) = 0$. The speed is then a Lagrange multiplier. Solitary waves come in one parameter families parameterized by $c$ and the family is non-degenerate when

$$\frac{dI}{dc} \neq 0.$$ 

(4.8)

It is assumed that there exists a solitary wave solution of the form (4.4)-(4.5) satisfying (4.6) and (4.8), and it is a smooth function of $\xi$ and $c$. This is hypothesis (H3).

The second variation is a linear operator

$$L(\xi,c) := D^2 H(\hat{Z}) - cD^2 I(\hat{Z}).$$

(4.9)

It is this operator that 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 (4.7) another representation of $L$ is

$$L = D^2 S(\hat{Z}) - K \frac{\partial}{\partial \xi} - cM \frac{\partial}{\partial \xi} = B(\xi,c) - J(c) \frac{\partial}{\partial \xi}.$$ 

(4.10)

with

$$B(\xi,c) = D^2 S(\hat{Z}).$$

(4.11)

This form for $L$ is also the linearization of the solitary wave equation (4.6) about $\hat{Z}$. 

15
The tangent vector to the solitary wave is in the kernel of $L$

$$L\hat{Z}_\xi = 0.$$  \hfill (4.12)

It is assumed that $\text{Ker}(L) \cap L^2(\mathbb{R}) = \text{span}\{\hat{Z}_\xi\}$. This is hypothesis (H4).

In the analysis of the Evans function an equation for $\hat{Z}_c$ will be needed. Differentiate (4.6) with respect to $c$

$$(K + cM)(\hat{Z}_c)_\xi + M\hat{Z}_\xi = B(\xi, c)\hat{Z}_c,$$

or

$$L\hat{Z}_c = M\hat{Z}_\xi.$$  \hfill (4.13)

Note the similarity with the second equation in the ODE case (2.9).

### 4.2 Linearization of the ODE about solitary waves

Written out, the linearization of the steady version (4.1) about the solitary wave solution is

$$J(c)Z_\xi = B(\xi, c)Z.$$  \hfill (4.14)

Due to the asymptotic 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,

$$Z(\xi) = \sum_{j=1}^{4} q_j \zeta_j e^{\mu_j \xi},$$

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

$$\Delta(\mu, 0; c) := \det \left[ B^\infty(c) - \mu J(c) \right] = 0.$$

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

$$\left[ B^\infty(c) - \mu_j(c, 0)J(c) \right] \zeta_j(c, 0) = 0, \quad j = 1, \ldots, 4.$$

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. It is assumed in addition that the four eigenvalues are simple. This is hypothesis H5.

Consistent with this splitting are solutions of the $\xi$-dependent equation (4.14)

$$E^u(\xi, 0) = \text{span}\{\hat{Z}_\xi, a^+\} \quad \text{and} \quad E^u(\xi, 0) = \text{span}\{\hat{Z}_\xi, a^-\},$$  \hfill (4.15)

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

$$\lim_{\xi \to +\infty} a^+(\xi, c) = 0 \quad \text{and} \quad \lim_{\xi \to -\infty} a^-\xi, c) = 0.$$  \hfill (4.16)
with the convergence exponential. In general $a^\pm$ are not bounded as $\xi \to \mp \infty$.

**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

$$\Omega(\hat{Z}_\xi, a^+) = 0 \quad \text{for all } \xi \in \mathbb{R}. \quad (4.17)$$

A direct calculation using (4.14) and the skew symmetry of $J(c)$ gives

$$\frac{d}{d\xi} \Omega(\hat{Z}_\xi, a^+) = 0.$$ 

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 (4.17). A similar proof confirms that $E^u$ is Lagrangian. ■

### 4.3 Transversality, the Lazutkin invariant, and orientation

A homoclinic orbit is said to be transversely constructed if $a^+$ and $a^-$ are linearly independent for all $\xi$ [12]; that is,

$$\Pi := \Omega(a^+, a^-) \neq 0. \quad (4.18)$$

This function is the Lazutkin invariant of a homoclinic orbit [16] [12].

To make the sign of $\Pi$ relevant, an orientation of the stable and unstable spaces is required. Represent $E^s$ and $E^u$ in by forms

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

These spaces are oriented as follows. Let

$$\alpha := \alpha_{11} \hat{Z}_\xi + \alpha_{12} a^+ \quad \text{and} \quad \beta := \alpha_{21} \hat{Z}_\xi + \alpha_{22} a^+,$$

be another basis for $E^s$. Then

$$\alpha \land \beta = \det \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \hat{Z}_\xi \land a^+.$$ 

We say that $E^s$ is **positively oriented** when

$$\det \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} > 0,$$

for the new basis. Assumption H6 is that both $E^s(\xi, 0)$ and $E^u(\xi, 0)$ are positively oriented. With this assumption the sign of the Lazutkin invariant (4.18) is independent of choice of basis of $E^s$ and $E^u$. 17
5 Linear stability and the Evans function

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

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

(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

\[ U_\xi = \mathbf{A}(\xi, \lambda) U, \quad U \in \mathbb{C}^4, \quad \lambda \in \Lambda, \]  

(5.2)

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

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

(5.3)

The asymptotic condition (4.5) assures that

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

(5.4)

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

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

(5.5)

with the dependence on \( c \) suppressed for brevity. This limit is assumed to exist for any fixed \( c \) and for all \( \lambda \in \Lambda \). The set \( \Lambda \) is defined based on the position of the eigenvalues of \( \mathbf{A}^\infty(\lambda) \). The spectrum of \( \mathbf{A}^\infty(\lambda) \) consists of four eigenvalues

\[ \sigma(\mathbf{A}^\infty(\lambda)) = \{ \mu_1(\lambda, c), \mu_2(\lambda, c), \mu_3(\lambda, c), \mu_4(\lambda, c) \}, \]  

and it is assumed that

\[ \text{Re}(\mu_1) \leq \text{Re}(\mu_2) < 0 < \text{Re}(\mu_3) \leq \text{Re}(\mu_4) \quad \forall \lambda \in \Lambda. \]  

(5.6)

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 \( \mathbf{A}^\infty(\lambda) \) satisfy the constraints (5.6). In this paper the focus is on \( \lambda \) near zero and proving the derivative formula (1.4). Hence, the set \( \Lambda \) is restricted further by assuming that the four spatial eigenvalues are simple for all \( \lambda \in \Lambda \) and satisfy (5.6). This is hypothesis (H7).

The continuous spectrum is defined by

\[ \sigma_c = \left\{ \lambda \in \mathbb{C} : \det[\mathbf{A}^\infty(\lambda) - i \kappa \mathbf{I}] = 0, \quad \kappa \in \mathbb{R} \right\}. \]  

(5.7)

There is no special assumption on the continuous spectrum. Normally, it would be assumed that the continuous spectrum is purely imaginary, \( \sigma_c \subset i\mathbb{R} \). However, our main interest is in the derivatives of the Evans function at \( \lambda = 0 \). Indeed in the example in §7 there exists continuous spectra in the unstable half plane.

The trace of \( \mathbf{A}(\xi, \lambda) \) satisfies

\[ \text{Tr}(\mathbf{A}(\xi, \lambda)) = \text{Tr}(\mathbf{B}(\xi, c) - \lambda \mathbf{M}) := \lambda \tau(c). \]  

(5.8)

for some scalar-valued function \( \tau(c) \). The first equality follows from the fact that \( \text{Tr}(\mathbf{B}(\xi, c)) \) is the trace of the product of a skew-symmetric with a symmetric matrix which is zero.
5.1 Constructing the Evans function

There are many equivalent ways of defining the Evans function (e.g. Chapters 8–10 in [22]). 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.

Associated with each of the simple eigenvalues (5.6) is an eigenvector

\[ A^\infty(\lambda) - \mu_j(\lambda)I \zeta_j(\lambda) = 0, \quad j = 1, \ldots, 4, \quad \forall \lambda \in \Lambda, \quad \tag{5.9} \]

suppressing the dependence on \( c \) as its importance is now secondary. Since \( \mu_j(\lambda) \) are simple and therefore analytic functions of \( \lambda \), the eigenvectors can also be constructed to be analytic for \( \lambda \in \Lambda \).

Using standard asymptotic theory for ODEs [22] there are four \((\zeta, \lambda)\)-dependent vectors satisfying

\[ (u_j)_\xi = A(\zeta, \lambda)u_j, \quad j = 1, \ldots, 4, \]

with the asymptotic properties

\[ \lim_{\xi \to +\infty} e^{-\mu_1(\lambda)\xi}u_1(\zeta, \lambda) = \zeta_1(\lambda), \quad \lim_{\xi \to +\infty} e^{-\mu_2(\lambda)\xi}u_2(\zeta, \lambda) = \zeta_2(\lambda), \]
\[ \lim_{\xi \to -\infty} e^{-\mu_3(\lambda)\xi}u_3(\zeta, \lambda) = \zeta_3(\lambda), \quad \lim_{\xi \to -\infty} e^{-\mu_4(\lambda)\xi}u_4(\zeta, \lambda) = \zeta_4(\lambda). \]

The Evans function is

\[ D(\lambda) = e^{-\tau(c)\xi}u_1(\zeta, \lambda) \wedge u_2(\zeta, \lambda) \wedge u_3(\zeta, \lambda) \wedge u_4(\zeta, \lambda), \quad \tag{5.10} \]

with \( \tau(c) \) defined in (5.8). The function \( D(\lambda) \) is independent of \( \xi \) and an analytic function of \( \lambda \) for all \( \lambda \in \Lambda \) [1, 22].

Here an equivalent definition in terms of individual vectors of (5.2) and its adjoint are used. The adjoint of (5.2) is

\[ W_\xi = -A(\zeta, \lambda)^H W, \quad W \in \mathbb{C}^4, \quad \lambda \in \Lambda. \quad \tag{5.11} \]

This adjoint equation (5.11) can be simplified, and connected more closely with (5.2) by pre-multiplying by \( J(c)^{-1} \) and using the special form of \( A(\zeta, \lambda) \) in (5.3),

\[ (J(c)^{-1}W)_\xi = -J(c)^{-1}A(\zeta, \lambda)^H W = J(c)^{-1}[B(\zeta, c) + \bar{\lambda}M](J(c)^{-1}W), \]

or

\[ (J(c)^{-1}W)_\xi = A(\zeta, -\bar{\lambda})(J(c)^{-1}W). \]

A natural definition of the solutions of the adjoint equation is then

\[ w(\zeta, \lambda) = J(c)^{-1}W(\zeta, \lambda). \]

The vector-valued functions \( w(\zeta, \lambda) \) are analytic and satisfy

\[ w_\xi = A(\zeta, -\lambda)w, \quad w \in \mathbb{C}^4, \quad \lambda \in \Lambda. \quad \tag{5.12} \]

Solutions of this equation paired with solutions of (5.2) are independent of \( \xi \), and this pairing can be expressed in terms of the symplectic form, \( \frac{d}{\xi} \Omega(w, U) = 0 \), and conjugation on one of the elements is not required.
There are four solutions of (5.12) 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 eigenvectors associated with the adjoint system at infinity
\[
[B^\infty + \lambda M \pm \mu_j J(c)] \eta_j = 0, \quad j = 1, 2, 3, 4,
\]
and they are normalized by \(\Omega(\eta_i, \zeta_j) = \delta_{ij}\).

Starting with the representation (5.10) and using the Hodge star operator [8, 9], an equivalent definition of the Evans function is
\[
D(\lambda) = \det \begin{bmatrix}
\Omega(w_3, u_3) & \Omega(w_3, u_4) \\
\Omega(w_4, u_3) & \Omega(w_4, u_4)
\end{bmatrix} \text{vol},
\] (5.13)
It is this representation that we will use in the proof of the derivative formula (1.4).

**Remark.** Two representations \(D^A(\lambda)\) and \(D^B(\lambda)\), of an Evans function, are said to be 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 \(C(\lambda)|_{\lambda \in \mathbb{R}}\) is required to be real and positive.

### 6 Derivatives of the Evans function

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

**Theorem 6.1.** The Evans function (5.13), 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 \(\lambda = 0\),
\[
D(0) = 0, \quad D'(0) = 0, \quad \text{and} \quad D''(0) = 2\Pi \frac{dI}{dc},
\] (6.1)
where \(\Pi\) is the transversality coefficient (1.5), and \(I(c)\) is the momentum, defined in (4.7), 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 (4.12). 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 (5.13) at \(\lambda = 0\),
\[
D(0) = \det \begin{bmatrix} 0 & 0 \\ C_3 C_4 \Pi \end{bmatrix} \text{vol},
\] (6.2)
where \(C_3\) and \(C_4\) are non-zero constants defined below. The zeros in the first column and row are confirmed by noting that \(u_3\) and \(w_3\) are in the kernel of \(L\) when \(\lambda = 0\), and so
\[
u_3(\xi, 0) = C_1 \tilde{Z}_\xi \quad \text{and} \quad w_3(\xi, 0) = C_2 \tilde{Z}_\xi,
\] (6.3)
for some real constants \( C_1 \) and \( C_2 \). For \( u_4 \), at \( \lambda = 0 \), we have that \( Lu_4 = 0 \) but is only required to decay as \( \xi \to -\infty \) whereas \( w_4 \) satisfies \( Lw_4 = 0 \) but is only required to decay as \( \xi \to +\infty \), giving

\[
    u_4(\xi, 0) = C_3a^- \quad \text{and} \quad w_4(\xi, 0) = C_4a^+ .
\]

where \( C_3 \) and \( C_4 \) are arbitrary constants. Combining these expressions

\[
    u_3(\xi, 0) \wedge u_4(\xi, 0) = C_1C_3\hat{\zeta} \wedge a^- \in E^a(\xi, 0)
\]

\[
    w_3(\xi, 0) \wedge w_4(\xi, 0) = C_2C_4\hat{\zeta} \wedge a^+ \in E^a(\xi, 0).
\]

Application of hypothesis \( H_6 \) then requires

\[
    C_1C_3 > 0 \quad \text{and} \quad C_2C_4 > 0 . \quad (6.4)
\]

Now use skew-symmetry of \( \Omega \) and the Lagrangian subspace property of the stable and unstable subspaces (Proposition 4.1) to conclude

\[
    \Omega(w_3, u_3)\big|_{\lambda=0} = C_1C_2\Omega(\hat{\zeta}, \hat{\zeta}) = 0
\]

\[
    \Omega(w_3, u_4)\big|_{\lambda=0} = C_1C_3\Omega(\hat{\zeta}, a^-) = 0
\]

\[
    \Omega(w_4, u_3)\big|_{\lambda=0} = C_2C_4\Omega(a^+, \hat{\zeta}) = 0
\]

\[
    \Omega(w_4, u_4)\big|_{\lambda=0} = C_3C_4\Omega(a^+, a^-) .
\]

Substitution into \( (5.13) \) 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

\[
    D(\lambda) = \left( d_1(\lambda)d_2(\lambda) - d_3(\lambda)d_4(\lambda) \right) \text{vol} \quad (6.5)
\]

where

\[
    d_1(\lambda) = \Omega(w_3, u_3) \quad d_2(\lambda) = \Omega(w_4, u_4) ,
\]

\[
    d_3(\lambda) = \Omega(w_3, u_4) , \quad d_4(\lambda) = \Omega(w_4, u_3) .
\]

It follows from \( (6.2) \) that

\[
    d_1(0) = d_3(0) = d_4(0) = 0 , \quad \text{and} \quad d_2(0) = C_3C_4\Pi . \quad (6.6)
\]

Computing the first derivative

\[
    D'(\lambda) = \left( d'_1(\lambda)d_2(\lambda) + d_1(\lambda)d'_2(\lambda) - d'_3(\lambda)d_4(\lambda) - d_3(\lambda)d'_4(\lambda) \right) \text{vol} .
\]

Evaluating at \( \lambda = 0 \) and using \( (6.6) \)

\[
    D'(0) = C_3C_4d'_1(0)\Pi . \quad (6.7)
\]

Now

\[
    d'_1(\lambda) = \Omega(\partial_\lambda w_3, u_3) + \Omega(w_3, \partial_\lambda u_3) . \quad (6.8)
\]

For \( \partial_\lambda u_3 \) and \( \partial_\lambda w_3 \), start with their defining equation,

\[
    \mathbf{J}(u_3)_\xi = [\mathbf{B} - \lambda\mathbf{M}]u_3 \quad \text{and} \quad \mathbf{J}(w_3)_\xi = [\mathbf{B} + \lambda\mathbf{M}]w_3 ;
\]
and differentiate with respect to \( \lambda \),

\[
J(u_3)_{\xi \lambda} = [B - \lambda M](u_3)_{\lambda} - Mu_3 \quad \text{and} \quad J(w_3)_{\xi \lambda} = [B + \lambda M](w_3)_{\lambda} + Mw_3.
\] (6.9)

Set \( \lambda = 0 \),

\[
L(u_3)_{\lambda} = Mu_3 \bigg|_{\lambda=0} = C_1 M \hat{Z}_{\xi}
\]

\[
L(w_3)_{\lambda} = -Mu_3 \bigg|_{\lambda=0} = -C_2 M \hat{Z}_{\xi},
\]

using (6.3). Now use equation (4.13), giving

\[
(u_3)_{\lambda} \big|_{\lambda=0} = C_1 \hat{Z}_c + C_5 \hat{Z}_\xi \quad \text{and} \quad (w_3)_{\lambda} \big|_{\lambda=0} = -C_2 \hat{Z}_c + C_6 \hat{Z}_\xi,
\] (6.10)

with \( C_5 \) and \( C_6 \) arbitrary constants.

Substitute the expressions (6.10) into (6.8) evaluated at \( \lambda = 0 \),

\[
d_1'(0) = [\Omega(\partial_\lambda w_3, u_3) + \Omega(w_3, \partial_\lambda u_3)] \bigg|_{\lambda=0}
\]

\[
= \left[ \Omega(-C_2 \hat{Z}_c + C_6 \hat{Z}_\xi, C_1 \hat{Z}_\xi) + \Omega(-C_2 \hat{Z}_c, C_1 \hat{Z}_c + C_5 \hat{Z}_\xi) \right]
\]

\[
= 2C_1 C_2 \Omega(\hat{Z}_\xi, \hat{Z}_c).
\]

This latter term is zero. To see this, first show that it is independent of \( \xi \),

\[
\frac{d}{d\xi} \Omega(\hat{Z}_c, \hat{Z}_\xi) = \langle J(\hat{Z}_c), \hat{Z}_\xi \rangle - \langle \hat{Z}_c, J(\hat{Z}_\xi) \rangle
\]

\[
= \langle B \hat{Z}_c, \hat{Z}_\xi \rangle + \langle M \hat{Z}_\xi, \hat{Z}_\xi \rangle - \langle \hat{Z}_c, B \hat{Z}_\xi \rangle
\]

\[
= 0,
\]

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 \( D'(0) = 0 \).

The second derivative is

\[
D''(\lambda) = \left( d_1''(\lambda) d_2(\lambda) + d_1'(\lambda) d_2'(\lambda) + d_1'(\lambda) d_2''(\lambda) + d_1''(\lambda) d_2'(\lambda) - d_1''(\lambda) d_4(\lambda) - d_3'(\lambda) d_3'(\lambda) - d_3'(\lambda) d_4'(\lambda) - d_3''(\lambda) d_3'(\lambda) \right) \text{vol}
\]

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

\[
D''(0) = (d_1''(0) d_2(0) + 2d_1'(0) d_2'(0) - 2d_3'(0) d_4'(0)) \text{vol}.
\] (6.11)

The second term is zero due to \( d_1'(0) = 0 \) as was shown above. A similar argument can be used to show that \( d_3'(0) = 0 \) as follows,

\[
d_3'(0) = C_2 C_3 \Omega(\hat{Z}_c, a^-) + C_2 \Omega(\hat{Z}_\xi, (u_4)_{\lambda=0}).
\] (6.12)

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 \( (u_4)_\lambda \big|_{\lambda = 0} \) be bounded as \( \xi \to -\infty \). But \( u_4 \big|_{\lambda = 0} \) goes to zero exponentially as \( \xi \to -\infty \) and \( \partial_\lambda u_4 \) will only add a polynomial in \( \xi \) to the exponential decay, resulting in the second term vanishing as well.

Hence the second derivative \([6.11]\) reduces to

\[
D''(0) = d_2(0) d_1'(0) \text{vol} = C_3 C_4 \Pi d_1''(0) \text{vol}. \tag{6.13}
\]

To compute \( d_1''(0) \) start with \( d_1'(\lambda) \) in \([6.8]\). Using \([6.9]\), we can write,

\[
\partial_\xi \langle J(w_3)_\lambda, u_3 \rangle = \langle [B + \lambda M] w_3, u_3 \rangle - \langle (w_3)_\lambda, [B - \lambda M] u_3 \rangle, \tag{6.14}
\]

and

\[
\partial_\xi \langle J w_3, (u_3)_\lambda \rangle = \langle [B + \lambda M] w_3, (u_3)_\lambda \rangle - \langle w_3, [B - \lambda M] (u_3)_\lambda \rangle + \langle w_3, M u_3 \rangle, \tag{6.15}
\]

which leaves us with:

\[
\partial_\xi \langle J(w_3)_\lambda, u_3 \rangle = \langle Mw_3, u_3 \rangle = -\partial_\xi \langle J w_3, (u_3)_\lambda \rangle. \tag{6.16}
\]

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:

\[
\left[ \langle J(w_3)_\lambda, u_3 \rangle \right]_{\xi = 0}^{\xi = R} = \int_0^R \langle Mw_3, u_3 \rangle \, d\xi, \tag{6.17}
\]

\[
\left[ -\langle J w_3, (u_3)_\lambda \rangle \right]_{\xi = -R}^{\xi = 0} = \int_{-R}^0 \langle Mw_3, u_3 \rangle \, d\xi. \tag{6.18}
\]

They can be combined to give:

\[
d_1'(\lambda) \big|_{\xi = 0} = -\int_{-R}^R \langle Mw_3, u_3 \rangle \, d\xi + \langle J(w_3)_\lambda, u_3 \rangle \big|_{\xi = R} - \langle J w_3, (u_3)_\lambda \rangle \big|_{\xi = -R}. \tag{6.19}
\]

(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

\[
d_1'(\lambda) = -\int_{-\infty}^{+\infty} \langle Mw_3, u_3 \rangle \, d\xi + \ell(\lambda), \tag{6.20}
\]

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 \( d_1''(\lambda) \):

\[
d_1''(\lambda) = -\int_{-\infty}^{+\infty} \langle M(w_3)_\lambda, u_3 \rangle + \langle Mw_3, (u_3)_\lambda \rangle \, d\xi + \ell'(\lambda). \tag{6.21}
\]

Now

\[
\ell'(\lambda) = \lim_{\xi \to +\infty} \left[ \langle (J(w_3)_\lambda), u_3 \rangle + \langle J(w_3)_\lambda, (u_3)_\lambda \rangle \right] \\
+ \lim_{\xi \to -\infty} \left[ \langle (J(w_3)_\lambda, (u_3)_\lambda \rangle + \langle J(w_3)_\lambda, (u_3)_\lambda \rangle \right].
\]
However, since
\[
\lim_{\xi \to +\infty} e^{\mu_3(\lambda)\xi} w_3 = \eta_3(\lambda)
\]
we can deduce that for \(\xi\) large and positive:
\[
w_3 \approx e^{-\mu_3\xi} \eta_3.
\]
This is turn implies that
\[
(w_3)_{\lambda\lambda} \approx p(\xi, \lambda)e^{-\mu_3\xi}.
\]
where \(p(\xi, \lambda)\) is a quadratic polynomial in \(\xi\). Since the exponential term will dominate the quadratic polynomial this tells us that
\[
\lim_{\xi \to +\infty} (w_3)_{\lambda\lambda} = 0.
\]
The same argument can be used to show that
\[
\lim_{\xi \to -\infty} (u_3)_{\lambda\lambda} = 0
\]
which means that
\[
\ell' (0) = \lim_{\xi \to +\infty} \left[ (J(w_3))_{\lambda\lambda}|_{\lambda=0} , C_1 \hat{Z}_\xi \right] + \lim_{\xi \to -\infty} \left[ (J(-C_2 \hat{Z}_\xi + C_6 \hat{Z}_\xi)) , C_1 C_2 \hat{Z}_\xi + C_5 \hat{Z}_\xi \right] + C_2 \langle \hat{Z}_\xi , (u_3)_{\lambda\lambda}|_{\lambda=0} \rangle
\]
\[
= 0.
\]
Therefore:
\[
d_1''(0) = - \int_{-\infty}^{+\infty} \langle M(-C_2 \hat{Z}_\xi + C_6 \hat{Z}_\xi), C_1 \hat{Z}_\xi \rangle + C_2 \langle M \hat{Z}_\xi, C_1 \hat{Z}_\xi + C_5 \hat{Z}_\xi \rangle \, d\xi
\]
\[
= 2C_1C_2 \int_{-\infty}^{+\infty} \langle M \hat{Z}_\xi, \hat{Z}_c \rangle \, d\xi
\]
\[
= 2C_1C_2 \frac{dI}{dc},
\]
(6.23)
since
\[
I = \frac{1}{2} \int_{-\infty}^{+\infty} \langle M \hat{Z}_\xi, \hat{Z} \rangle \, d\xi.
\]
Combining these results
\[
D''(0) = 2C_1C_2C_3C_4 \Pi \frac{dI}{dc}.
\]
Hypothesis \(\text{H6}\) via (6.4) gives that \(C_1C_2C_3C_4 > 0\). Scale \(D(\lambda)\) by dividing through by the positive constant \(C_1C_2C_3C_4\). This proves the Theorem.

Using Theorem 6.3 above and Theorem 10.1 in [12] an alternative formula for the second derivative in terms of the Maslov index is obtained,

**Corollary.** Under the above hypotheses, an alternative formula for \(D''(0)\) is,
\[
D''(0) = 2(-1)^{\text{Maslov}} \frac{dI}{dc}.
\]

7 Example: a coupled “wave equation”

To illustrate the theory it is applied to an example nonlinear “wave equation”

\[
\begin{align*}
\phi_{tt} + \phi_{xx} - 4\phi + 6\phi^2 - p(\phi - v) &= 0 \\
v_{tt} + v_{xx} - 4v + 6v^2 + p(\phi - v) &= 0.
\end{align*}
\] (7.1)

Wave equation is in quotes as the evolution equation is ill-posed. However, it is a useful example on two fronts. It shows that well-posedness is irrelevant in the computation of derivatives of the Evans function near the origin, and secondly, explicit calculations can be carried out illustrating by example the nature of both \(dI/dc\) and \(\Pi\).

In §3.3 it is shown that coupled wave equations can be put into the canonical form (3.19). The only difference here is that the PDO is constructed on the Riemannian manifold \(\mathbb{R}^2\) with the flat Euclidean metric, and so \(\mathcal{R}\) and its induced metrics are all represented by the identity. Hence the canonical multisymplectic formulation of (7.1) is

\[
MZ_t + KZ_x = \nabla S(Z), \quad Z \in \mathbb{R}^4,
\] (7.2)

with

\[
Z = \begin{pmatrix} \phi \\ u_1 \\ u_2 \\ v \end{pmatrix}, \quad M = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad K = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix},
\] (7.3)

where

\[
u_1 = \phi_t + v_x, \quad u_2 = -v_t + \phi_x,
\]

and

\[
S(Z) = 2(\phi^3 - \phi^2) + 2(v^3 - v^2) - \frac{1}{2}p(\phi - v)^2 + \frac{1}{2}(u_1^2 + u_2^2).
\] (7.4)

The symplectic operator \(J(c)\) is

\[
J(c) = (K + cM) = \begin{bmatrix} 0 & -c & -1 & 0 \\ c & 0 & 0 & 1 \\ 1 & 0 & 0 & -c \\ 0 & -1 & c & 0 \end{bmatrix},
\]

with \(\det(J(c)) = (1 + c^2)^2\) giving that \(J(c)\) is invertible for all \(c \in \mathbb{R}\).

The system has an exact solitary wave solution

\[
\phi(x, t) = \hat{\phi}(\xi) \quad \text{and} \quad v(x, t) = \hat{\phi}(\xi), \quad \xi = x + ct,
\]

with

\[
\hat{\phi}(\xi) = \text{sech}^2(\alpha \xi), \quad \alpha = \frac{1}{\sqrt{1 + c^2}}.
\]

The solitary wave solution exists for all \(c \in \mathbb{R}\). In terms of the \(Z\)-coordinates in (7.3) the solitary wave solution is

\[
\hat{Z}(\xi) = \begin{pmatrix} \hat{\phi} \\ (1 + c)\hat{\phi} \xi \\ (1 - c)\hat{\phi} \xi \\ \hat{\phi} \end{pmatrix}.
\] (7.5)
7.1 Linearization about the solitary wave

The linearization about the solitary wave solution is

\[ J(c) Z_\xi = [B(\xi, c) - \lambda M] Z, \quad Z \in \mathbb{C}^4, \]  

\[ (7.6) \]

with

\[ [B(\xi, c) - \lambda M] = \begin{bmatrix} 12\hat{\phi} - 4 - p & \lambda & 0 & p \\ -\lambda & 1 & 0 & 0 \\ 0 & 0 & 1 & \lambda \\ p & 0 & -\lambda & 12\hat{\phi} - 4 - p \end{bmatrix}. \]  

\[ (7.7) \]

The system at infinity is

\[ J(c) Z_\xi = [B^\infty(c) - \lambda M] Z, \quad Z \in \mathbb{C}^4, \]

with \([B^\infty(c) - \lambda M]\) the same as \((7.7)\) but with \(\hat{\phi}\) set to zero. The eigenvalues of the system at infinity are defined by \(\Delta(\mu, \lambda) = 0\) with

\[ \Delta(\mu, \lambda) = [B^\infty(c) - \lambda M - \mu J(c)] \]

\[ = (\varrho^2 + \mu^2)^2 - 2(4 + p)(\varrho^2 + \mu^2) + 16 + 8p, \]

where \(\varrho = \lambda + c\mu\). The continuous spectrum (defined in \((5.7)\)) is

\[ \sigma_c = \{ \lambda \in \mathbb{C} : \Delta(\imath \kappa, \lambda) = 0, \quad \kappa \in \mathbb{R} \} \]

\[ = \{ \lambda = -\imath \kappa \pm \sqrt{4 + p + \kappa^2 \pm p}, \quad \kappa \in \mathbb{R} \}. \]  

\[ (7.8) \]

There are four branches in \(\sigma_c\) and they are shown in the complex \(\lambda\)–plane in Figure 1 along with the double zero eigenvalue at \(\lambda = 0\) which is confirmed below using the theory in this paper. The fact that \(\text{Re}(\lambda) \to \pm \infty\) along the hyperbolae is a reflection of the ill-posedness of the initial-value problem for \((7.1)\).

Now set \(\lambda = 0\) and look at the spatial eigenvalues of the system at infinity. Setting \(\Delta(\mu, 0) = 0\) gives four spatial eigenvalues

\[ \mu_1 = -\sqrt{\frac{4 + 2p}{1 + c^2}}, \quad \mu_2 = -\sqrt{\frac{4}{1 + c^2}}, \quad \mu_3 = +\sqrt{\frac{4}{1 + c^2}}, \quad \mu_4 = +\sqrt{\frac{4 + 2p}{1 + c^2}}. \]
A schematic of their position is shown in Figure 2. In order to ensure that the four eigenvalues are simple, it is assumed that

\[ p > 0. \]  

(7.9)

With this assumption the example satisfies all the hypotheses in the theory, and the derivative formula in Theorem 6.3 is operational. Hence \( D(0) = D'(0) = 0 \), giving the double zero eigenvalue in Figure 1.

The aim here is to compute the two parts, \( II \) and \( \frac{dI}{dc} \), in the formula for \( D''(0) \). The easier of the two is \( I(c) \) and its derivative. Using the definition in (4.7), the momentum of the solitary wave is

\[ I = \frac{1}{2} \int_{-\infty}^{+\infty} \langle M \hat{Z}_\xi, \hat{Z} \rangle \, d\xi. \]

Substituting in \( M \) from (7.2) and \( \hat{Z} \) from (7.5), gives

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

after integrating by parts. Now

\[ \hat{\phi}_\xi = -2\alpha \text{sech}^2(\alpha \xi) \tanh(\alpha \xi) \]

and so

\[
\hat{\phi}_\xi^2 = 4\alpha^2 \text{sech}^4(\alpha \xi) \tanh^2(\alpha \xi) \\
= 4\alpha^2 \text{sech}^2(\alpha \xi) \left[ \tanh^2(\alpha \xi) - \tanh^4(\alpha \xi) \right]
\]

(7.10)
which allows us to evaluate the integral as

\[
I = 2c \int_{-\infty}^{+\infty} 4\alpha^2 \text{sech}^2(\alpha \xi) \left[ \tanh^2(\alpha \xi) - \tanh^4(\alpha \xi) \right] d\xi
\]

\[
= 8\alpha c \left[ \frac{1}{3} \tanh(\alpha \xi) - \frac{1}{5} \tanh^5(\alpha \xi) \right]_{\xi=-\infty}^{\xi=+\infty}
\]

\[
= \frac{32 c}{15 \sqrt{1 + c^2}}.
\]

(7.11)

Differentiating this w.r.t. \(c\) then gives us

\[
\frac{dI}{dc} = \frac{32}{15} (1 + c^2)^{-\frac{3}{2}} = \frac{32}{15} \alpha^3,
\]

which is strictly positive for all \(c \in \mathbb{R}\).

The transversality coefficient is

\[
\Pi = \langle J(c) a^+, a^- \rangle.
\]

(7.13)

Although the abstract existence of \(a^\pm\) is assured, their calculation is generally nontrivial, and in most applications they will need to be calculated numerically. The advantage of this example is that \(a^\pm\) can be calculated explicitly. Here the explicit calculation of the solutions of (7.6) are constructed for \(\lambda\) nonzero, and then \(a^\pm\) will be obtained \textit{a posteriori} by setting \(\lambda = 0\).

Express the solutions of (7.6) by \(Z = (\tilde{\phi}, \tilde{u}_1, \tilde{u}_2, \tilde{v})\). Then \(\tilde{u}_1\) and \(\tilde{u}_2\) can be obtained from (7.6) as

\[
\tilde{u}_1 = \lambda \tilde{\phi} + c \tilde{v}_\xi + \tilde{v}_\xi \quad \text{and} \quad \tilde{u}_2 = -\lambda \tilde{\phi} - c \tilde{v}_\xi + \tilde{\phi}_\xi.
\]

This allows us to rewrite the first and fourth components of (7.6) as the coupled pair of equations:

\[
\begin{align*}
(1 + c^2) \tilde{\phi}_{\xi\xi} + 2c\lambda \tilde{\phi}_\xi + (\chi + \lambda^2) \tilde{\phi} + p \tilde{v} &= 0, \\
(1 + c^2) \tilde{v}_{\xi\xi} + 2c\lambda \tilde{v}_\xi + (\chi + \lambda^2) \tilde{v} + p \tilde{\phi} &= 0,
\end{align*}
\]

(7.14)

where \(\chi := 12\tilde{\phi} - 4 - p\). If we now take the transformation

\[
\tilde{\phi} = e^{-\beta \xi}(\psi_1 + \psi_2), \quad \tilde{v} = e^{-\beta \xi}(\psi_2 - \psi_1), \quad \beta = \frac{c}{1 + c^2}
\]

then equations (7.14) will decouple to give

\[
\begin{align*}
(1 + c^2)(\psi_1)_{\xi\xi} + [\chi - p + \lambda^2(1 - \beta c)] \psi_1 &= 0, \\
(1 + c^2)(\psi_2)_{\xi\xi} + [\chi + p + \lambda^2(1 - \beta c)] \psi_2 &= 0,
\end{align*}
\]

(7.15)

which can be rearranged as

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

(7.16)

since

\[
1 - \beta c = 1 - \frac{c^2}{1 + c^2} = \frac{1 + c^2 - c^2}{1 + c^2} = \frac{1}{1 + c^2} = \alpha^2.
\]
Now set $\lambda = 0$ in the equations for $\psi_1$ and $\psi_2$. Firstly, it is easily shown that the hypothesis $\textbf{H4}$ is confirmed; that is $\text{Ker}(L) \cap L^2(\mathbb{R}) = \text{span}\{Z_\xi\}$. Secondly, we can see that $\mathbf{a}^\pm$ will be produced by the following two $\psi_1$ solutions (now denoted by $\psi^\pm$) of (7.16)

$$\psi^\pm = e^{\mp \alpha \gamma \xi} \left[ \pm \frac{2p\gamma}{15} + \left(1 + \frac{4p}{5}\right) \tanh(\alpha \xi) \pm \gamma \tanh^2(\alpha \xi) + \tanh^3(\alpha \xi) \right]$$

where $\gamma = \sqrt{4 + 2p}$. By reversing the transformations to express $\tilde{\phi}, \tilde{u}_1, \tilde{u}_2, \tilde{v}$ in terms of $\psi^\pm$ we find that

$$\mathbf{a}^+ = \begin{pmatrix} (c - 1) \psi_\xi^+ \\ (c + 1) \psi_\xi^+ \\ -\psi^+ \end{pmatrix} \quad \text{and} \quad \mathbf{a}^- = \begin{pmatrix} (c - 1) \psi_\xi^- \\ (c + 1) \psi_\xi^- \\ -\psi^- \end{pmatrix}.$$ 

Substitution into the formula for $\Pi$ then gives

$$\langle J \mathbf{a}^+, \mathbf{a}^- \rangle = -(1 + c^2) \psi_\xi^+ \psi^- + (c - 1)^2 \psi^+ \psi^- + (c + 1)^2 \psi^+ \psi^- - (1 + c^2) \psi_\xi^+ \psi^-$$

$$= (c^2 - 2c + 1 + c^2 + 2c + 1) \psi^+ \psi^- - 2(1 + c^2) \psi_\xi^+ \psi^-$$

$$= 2(1 + c^2) \left( \psi^+ \psi^- - \psi_\xi^+ \psi^- \right).$$ (7.17)

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

$$\psi_\xi^+ = \mp \alpha \gamma \psi^+ + \alpha e^{\mp \alpha \gamma \xi} \text{sech}^2(\alpha \xi) \left[ 1 + \frac{4p}{5} \pm 2\gamma \tanh(\alpha \xi) + 3 \tanh^2(\alpha \xi) \right]$$

we get

$$\langle J \mathbf{a}^+, \mathbf{a}^- \rangle = 2(1 + c^2) \alpha \left[ \frac{2p\gamma}{15} \left(1 + \frac{4p}{5} - \frac{2p\gamma^2}{15}\right) - \left(1 + \frac{4p}{5} - \frac{2p\gamma^2}{15}\right) \left(-\frac{2p\gamma}{15}\right) \right]$$

$$= \frac{8p\gamma}{15\alpha} \left(1 + \frac{4p}{5} - \frac{2p(4 + 2p)}{15}\right)$$

$$= \frac{8p\gamma}{225\alpha} \left(15 + 4p - 4p^2\right)$$

$$= \frac{8p}{225\alpha} \sqrt{4 + 2p(5 - 2p)(3 + 2p)}$$ (7.18)

which gives us the final result: $D(0) = 0$, $D'(0) = 0$, and

$$D''(0) = \frac{512\alpha^2 p}{3375} \sqrt{4 + 2p(5 - 2p)(3 + 2p)}. \quad (7.19)$$

With the assumption (7.9) of $p$ positive, the second derivative is positive for $0 < p < 5/2$ and negative for $p > 5/2$. Although this change of sign may indicate the existence of a $\lambda$-eigenvalue it does not correlate with stability or instability as the time evolution is ill posed.

8 Concluding remarks

The assumption of a four dimensional phase space in the steady problem (4.6) is sufficient to capture the essence of the theory, and avoids unnecessary complexity. When the phase space
has dimension 2n with \( n > 2 \), then section 4 would be similar with the major change arising in the construction of the stable and unstable spaces and the transversality coefficient. The \( 2 - 2 \) splitting would be replaced by an \( n - n \) splitting and (4.15) would be replaced by

\[
E^s(\xi, 0) = \text{span}\{\hat{Z}_\xi, a_1^+, \ldots, a_n^+\} \quad \text{and} \quad E^u(\xi, 0) = \text{span}\{\hat{Z}_\xi, a_1^-, \ldots, a_{n-1}^-\},
\]  
(8.1)

with \( \Pi \) replaced by the Lazutkin-Treschev invariant

\[
\Pi = \det \begin{bmatrix}
\Omega(a_1^+, a_1^-) & \cdots & \Omega(a_1^+, a_{n-1}^-) \\
\vdots & \ddots & \vdots \\
\Omega(a_{n-1}^+, a_1^-) & \cdots & \Omega(a_{n-1}^+, a_{n-1}^-)
\end{bmatrix},
\]  
(8.2)

(cf. Treschev [31] and Chardard & Bridges [12]). Hence, subject to the generalizations required in the Evans function construction, we expect that Theorem 6.1 generalizes with \( \Pi \) replaced by (8.2), although we do not want to underestimate the issues of detail that may arise.

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.

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

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

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

[7] 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).

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

[10] T.J. Bridges & G. Derks. *Constructing the symplectic Evans matrix using maximally analytic individual vectors*, Proc. Royal Soc. Edin. A 133 505–526 (2003).

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

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

[13] 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).

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

[15] F. Cooper, A. Khare, B. Mihaila, & A. Saxena. *Solitary waves in the nonlinear Dirac equation with arbitrary nonlinearity*, Phys. Rev. E 82 036604 (2010).

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

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

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

[19] 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).

[20] R. Grimshaw & P. Christodoulides. *Short-wave instability in a three-layer stratified shear flow*, Quart. J. Mech. Appl. Math. 54 375–388 (2001).

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

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

[23] J.H. Maddocks. *Stability and folds*, Arch. Rat. Mech. Anal. 99 301–328 (1987).

[24] 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).

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

[26] S. Morita. *Geometry of Differential Forms*, AMS: Providence (2001).
[27] S.J. Orfandis & R. Wang. *Soliton solutions of the massive Thirring model*, Phys. Lett. B 57 281–283 (1975).

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

[29] 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).

[30] 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).

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

[32] T.I. Vogel. *On constrained extrema*, Pac. J. Math. 176 557–561 (1996).