Transversality of homoclinic orbits, the Maslov index, and the symplectic Evans function

by Thomas J. Bridges[1] & Frédéric Chardard[2]

1. Department of Mathematics, University of Surrey, Guildford, Surrey GU2 7XH, England

2. Unité de Mathématiques Pures et Appliquées, École Normale Supérieure de Lyon, 46, allée d'Italie, 69007 Lyon, France

– Abstract –

Partial differential equations in one space dimension and time, which are gradient-like in time with Hamiltonian steady part, are considered. The interest is in the case where the steady equation has a homoclinic orbit, representing a solitary wave. Such homoclinic orbits have two important geometric invariants: a Maslov index and a Lazutkin invariant. A new relation between the two has been discovered and is moreover linked to transversal construction of homoclinic orbits: the sign of the Lazutkin invariant determines the parity of the Maslov index. A key tool is the geometry of Lagrangian planes. All this geometry feeds into the linearization about the homoclinic orbit in the time dependent system, which is studied using the Evans function. A new formula for the symplectification of the Evans function is presented, and it is proved that the the derivative of the Evans function is proportional to the Lazutkin invariant. A corollary is that the Evans function has a simple zero if and only if the homoclinic orbit of the steady problem is transversely constructed. Examples from the theory of gradient reaction-diffusion equations and pattern formation are presented.

— March 22, 2012—

Contents

1 Introduction

The starting point is partial differential equations in one space dimension and time where the time-independent part is a finite-dimensional Hamiltonian system. In particular, systems of the following form,

$$\mathbf{M}\mathbf{u}_t + \mathbf{J}\mathbf{u}_x = DH(\mathbf{u}, p)\,, \quad \mathbf{u} \in \mathbb{V}\,, \tag{1.1}$$

where $\mathbb{V}$ is a finite-dimensional normed vector space, $p \in \mathbb{R}$ is a parameter, and $H : \mathbb{V} \times \mathbb{R} \to \mathbb{R}$ is a smooth Hamiltonian function with $DH(\mathbf{u}, p)$ the derivative with respect to the first component. The matrix $\mathbf{J}$ is a symplectic operator associated with the symplectic form, denoted by $\mathbf{\Omega}$, and $\mathbf{M}$ is in general an arbitrary matrix acting on $\mathbb{V}$. However, in this paper two restrictions will be operational: the matrix $\mathbf{M}$ is assumed to be symmetric and the vector space $\mathbb{V}$ is taken to be four-dimensional. The former assumption gives the time dependence a gradient-like structure. The latter assumption can be relaxed: the theory presented here extends to dimension greater than four but the formulas can get unwieldy [4].

Examples of PDEs that can be represented in the form (1.1) are the Swift-Hohenberg equation

$$\phi_t + \phi_{xxxx} + p\phi_{xx} + \phi - \phi^2 = 0\,, \tag{1.2}$$

where p is a real parameter, which is widely used as a model in pattern formation, and coupled gradient reaction-diffusion equations

$$v_t = d_1\, v_{xx} + F_v(v, w) \quad \text{and} \quad w_t = d_2\, w_{xx} + F_w(v, w)\,, \tag{1.3}$$

where $F(v, w)$ is a given smooth function, and d_1, d_2 are positive constants, which is a model for coupled nerve fibers [3].

Suppose that the steady equation, $\mathbf{J}\mathbf{u}_x = DH(\mathbf{u}, p)$, has a homoclinic orbit, denoted $\widehat{\mathbf{u}}(x, p)$. This homoclinic orbit has two important characteristics: a Maslov index and a Lazutkin invariant. Since $\mathbb{V}$ has dimension four the stable and unstable subspaces, in the linearization of the steady system about the homoclinic orbit, are of the form $\mathrm{span}\{\widehat{\mathbf{u}}_x, \mathbf{a}^\pm\}$, for each x. The Lazutkin homoclinic invariant, *which is independent of* x, is defined by

$$\mathcal{T}(\widehat{\mathbf{u}}) = \mathbf{\Omega}(\mathbf{a}^-, \mathbf{a}^+)\,. \tag{1.4}$$

It was first used by Lazutkin [20] to study the distance between the stable and unstable separatrices in area preserving mappings. It has been a valuable tool to study the case where the distance between the stable and unstable manifolds is exponentially small (e.g. [17, 16, 18]) (the formula (1.4) is given explicitly in part C of §2.3 of [17] and an explicit example is given in [18]). In this paper three new results about this invariant are proved. Firstly, we give a new proof that a homoclinic orbit is transversely constructed if and only if $\mathcal{T}(\widehat{\mathbf{u}}) \neq 0$. Secondly we prove that it determines the parity of the Maslov index of the homoclinic orbit,

$$(-1)^{\mathsf{Maslov}} = \mathrm{sign}\big(\mathcal{T}(\widehat{\mathbf{u}})\big)\,, \tag{1.5}$$

where $\mathsf{Maslov} = I_{hom}(\widehat{\mathbf{u}}) + \frac{1}{2}$, and $I_{hom}(\widehat{\mathbf{u}})$ is the Maslov index of the homoclinic orbit. The addition of $\frac{1}{2}$ in the definition assures that Maslov is an integer.

Thirdly, the Evans function, constructed from the linearization about the homoclinic orbit in the time-dependent system (1.1), has a double zero eigenvalue if and only if $\mathcal{T}(\widehat{\mathbf{u}}) = 0$. All these properties are intimately connected with the fact that the Lazutkin invariant can be interpreted as an intersection index for codimension one intersection of two Lagrangian planes.

The use of the Maslov index to study of the linearization about homoclinic orbits, as models for solitary waves, was pioneered in the work of Jones [19] and Bose & Jones [3]. A numerical framework for computing the Maslov index of solitary waves was introduced in Chardard et al. [12, 13]. Other definitions of the Maslov index were proposed in Chardard [9] and Chen & Hu [14]. In this paper we use a definition for the Maslov index based on a theory of Souriau [25]. It is equivalent to the above definitions and it can be related much more easily to the Lazutkin invariant. In addition to the connection (1.5) we show how the Maslov index enters the theory of the Evans function.

The linearization of (1.1) about a homoclinic orbit, with a spectral ansatz and spectral parameter λ, can be put into standard form for the theory of the Evans function (e.g. Alexander et al. [1]). Let $\text{span}\{\mathbf{u}_1^+, \mathbf{u}_2^+\}$ be the $(x,\lambda)-$dependent stable subspace, and $\text{span}\{\mathbf{u}_3^+, \mathbf{u}_4^+\}$ be the $(x,\lambda)-$dependent unstable subspace in the linearization, then the Evans function is

$$D(\lambda)\mathsf{vol} = \mathbf{u}_1^+ \wedge \mathbf{u}_2^+ \wedge \mathbf{u}_3^- \wedge \mathbf{u}_4^- \,, \tag{1.6}$$

where vol is a volume form on $\mathbb{V}$. One of the main results of the paper is a proof of the formula

$$D'(0) = -\mathcal{T}(\widehat{\mathbf{u}}) \int_{-\infty}^{+\infty} \langle \mathbf{M}\widehat{\mathbf{u}}_x, \widehat{\mathbf{u}}_x \rangle \, \mathrm{d}x \,. \tag{1.7}$$

If the integral on the right-hand side is non-vanishing, then it is immediate that $D'(0) = 0$ if and only if the Lazutkin invariant vanishes. The proof that the Evans function has a simple zero when the homoclinic orbit is transversely constructed is a Hamiltonian version of a Theorem of Alexander & Jones [2] (see also §4 of [3]). There, transversality is obtained by lifting the phase space by one dimension by including a parameter. Here the dimension is reduced by one dimension due to the energy surface, and moreover the derivative $D'(0)$ in (1.7) is expressed in terms of a geometric invariant of the homoclinic orbit.

A key step in the proof is to reformulate the Evans function (1.6) in such a way that the symplectic structure becomes apparent. A new formula which will be used throughout is the following connection between four-forms on $\mathbb{V}$ and symplectic determinants. For any vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ in $\mathbb{V}$,

$$\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c} \wedge \mathbf{d} = \det \begin{bmatrix} \boldsymbol{\Omega}(\mathbf{a},\mathbf{c}) & \boldsymbol{\Omega}(\mathbf{a},\mathbf{d}) \\ \boldsymbol{\Omega}(\mathbf{b},\mathbf{c}) & \boldsymbol{\Omega}(\mathbf{b},\mathbf{d}) \end{bmatrix} \mathsf{vol} - \boldsymbol{\Omega}(\mathbf{a},\mathbf{b})\boldsymbol{\Omega}(\mathbf{c},\mathbf{d})\mathsf{vol} \,. \tag{1.8}$$

We have not seen this formula before. The key to proving it is the fact that the symplectic form and the volume form on $\mathbb{V}^*$ are related by $\mathsf{vol}^* = -\frac{1}{2}\mathbf{\Omega}\wedge\mathbf{\Omega}$, and a proof is given in Appendix A.

An outline of the paper is as follows. First in §2 it is shown how systems like (1.2) and (1.3) can be cast into the form (1.1), and establish some of the properties of the class of equations (1.1). A simplified ODE (ordinary differential equation) version of (1.1) is considered in §3 and it is shown how the formula (1.7) arises naturally.

The stable and unstable subspaces are paths of Lagrangian planes, and the background needed on the geometry of Lagrangian planes is given in §4 and Appendix B. Section 4 also includes a new proof of the necessary and sufficient condition for two Lagrangian planes to have a complete intersection, which is essential for understanding degeneracy of the Lazutkin invariant. In §5 transversal intersection and its implications are presented.

The construction, symplectification and differentiation of the Evans function are presented in §7, leading to a proof of the formula (1.7).

The longest proof in the paper is the proof of the connection between the Maslov index and the Lazutkin invariant (1.5). The proof uses the full power of homotopy equivalence of the Maslov index. First Souriau's construction of the Maslov index is reviewed in §9, and then the proof of (1.5) is given in §10.

One of the hidden features of Evans function analysis is the role of normalizations. In this paper three different normalizations are used, mainly to make the formulas tidy. Although essential, such normalizations do not make for interesting reading, and a summary is given in the Appendix C.

2 Gradient PDEs with Hamiltonian steady part

In this section, the examples (1.2) and (1.3) are formulated in the form (1.1), and the key properties of systems in the form (1.1) are identified. The assumptions on the matrix $\mathbf{M}:\mathbb{V}\to\mathbb{V}^*$ are

$$\mathbf{M}^T = \mathbf{M} \quad \text{and the eigenvalues of } \mathbf{M} \text{ are non-negative}\,. \tag{2.1}$$

The symplectic operator will be taken in standard form

$$\mathbf{J} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{2.2}$$

It is derived from the symplectic form after a choice of basis for $\mathbb{V}$ (cf. §4).

The Swift-Hohenberg equation (1.2) can be cast into the form (1.1) by taking

$$\mathbf{u} = (u_1,u_2,u_3,u_4) := (\phi, \phi_{xx}, -\phi_{xxx} - p\phi_x, -\phi_x)\,,$$

$\mathbf{J}$ in standard form (2.2), and

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} . \tag{2.3}$$

The Hamiltonian function is

$$H(\mathbf{u}, p) = \tfrac{1}{2} u_2^2 + \tfrac{1}{2} p u_4^2 - \tfrac{1}{2} u_1^2 - u_3 u_4 + \tfrac{1}{3} u_1^3 .$$

A second example is the pair of reaction diffusion equations (1.3). Systems of this type are considered in Bose & Jones [3]. The system (1.3) can be expressed in the form (1.1) by taking

$$\mathbf{u} = (u_1, u_2, u_3, u_4) := (v, w, v_x, w_x) ,$$

$\mathbf{J}$ in the standard form (2.2), and

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} .$$

The Hamiltonian function is

$$H(\mathbf{u}, d_1, d_2) = \frac{1}{2d_1} u_3^2 + \frac{1}{2d_2} u_4^2 + F(u_1, u_2) .$$

2.1 Gradient-like structure

We call PDEs of the form (1.1) "gradient-like PDEs" because there is a functional which is monotone on orbits. This designation is formal since a function space is not introduced. Define

$$\mathcal{F} := \tfrac{1}{2} \mathbf{\Omega}(\mathbf{u}_x, \mathbf{u}) - H(\mathbf{u}, p) \quad \text{and} \quad \mathcal{A} = \tfrac{1}{2} \mathbf{\Omega}(\mathbf{u}, \mathbf{u}_t) . \tag{2.4}$$

Note that $\mathcal{F}$ is the density for Hamilton's principle for steady solutions. Differentiating $\mathcal{F}$ and $\mathcal{A}$ gives

$$\mathcal{F}_t + \mathcal{A}_x = \langle \mathbf{M}\mathbf{u}_t + \mathbf{J}\mathbf{u}_x - DH(\mathbf{u}, p), \mathbf{u}_t \rangle - \langle \mathbf{M}\mathbf{u}_t, \mathbf{u}_t \rangle .$$

Suppose $\mathbf{u}$ is a solution of (1.1). Then with integration over x and appropriate boundary conditions on $\mathcal{A}$, the integral of $\mathcal{F}$, denoted $\overline{\mathcal{F}}$, is formally decreasing when evaluated on solutions of (1.1),

$$\overline{\mathcal{F}}_t = -\overline{\langle \mathbf{M}\mathbf{u}_t, \mathbf{u}_t \rangle} \le 0 .$$

The functional $\mathcal{F}$, being associated with Hamilton's principle, is indefinite in general. However, this gradient-like structure assures that the eigenvalue λ in the Evans function can be taken to be real, and it affects the formula (1.7). If $\mathbf{M}$ was skew-symmetric for example, then $D'(0) = 0$ in (1.7).

2.2 Cauchy-Riemann operators and Floer theory

Another interesting example is when $\mathbf{M}$ is the identity

$$\mathbf{u}_t + \mathbf{J}\mathbf{u}_x = DH(\mathbf{u}), \quad \mathbf{u} \in \mathbb{V}. \tag{2.5}$$

It is primarily of theoretical interest, as it is the form of the equation used in Morse-Floer theory [23, 14], and the left-hand side, $\mathbf{u}_t + \mathbf{J}\mathbf{u}_x$, is a Cauchy-Riemann operator. Since the Cauchy-Riemann operator is elliptic, this PDE is not an evolution equation. This case is not considered in the paper because the Evans function construction in the linearization would require modification: in this case the essential spectrum, in the linearization about a homoclinic orbit, is the entire real line.

3 Intermezzo: gradient ODE systems

Before proceeding to analyze the class of systems (1.1) is is useful to consider the case of gradient ODEs, as it provides the inspiration for the formula (1.7). Consider the system of gradient ODEs

$$\mathbf{M}\mathbf{u}_t = DH(\mathbf{u}, p), \quad \mathbf{u} \in \mathbb{V}, \tag{3.1}$$

where $H(\mathbf{u}, p)$ is a smooth function, and p is a parameter. The matrix $\mathbf{M}$ is assumed to have the properties (2.1). In this section the vector space $\mathbb{V}$ can have arbitrary finite dimension.

Suppose there exists a family of equilibrium solutions, $\widehat{\mathbf{u}}(p)$, of (3.1); that is, satisfying $DH(\widehat{\mathbf{u}}(p), p) = 0$. Let $\mathbf{L}(p) := D^2H(\widehat{\mathbf{u}}(p), p)$, and suppose there is a value of p, denoted p_0, at which $\mathbf{L}$ has a simple zero eigenvalue with eigenvector ξ,

$$\mathbf{L}(p_0)\xi = 0 \quad \text{with} \quad \|\xi\| = 1. \tag{3.2}$$

Look at the linearization of (3.1) about $\widehat{\mathbf{u}}(p)$,

$$\mathbf{M}\mathbf{v}_t = \mathbf{L}(p)\mathbf{v}.$$

With the spectral ansatz, $\mathbf{v}(t) \mapsto \mathrm{e}^{\lambda t}\mathbf{v}$, the exponent λ is an eigenvalue of

$$[\mathbf{L}(p) - \lambda\mathbf{M}]\mathbf{v} = 0.$$

The Evans function in this case is just the characteristic determinant

$$D(\lambda) = \det[\mathbf{L}(p) - \lambda\mathbf{M}].$$

At $p = p_0$ and $\lambda = 0$, $D(0) = \det[\mathbf{L}(p_0)] = 0$. Differentiating

$$D'(\lambda) = -\mathrm{Trace}\big([\mathbf{L}(p) - \lambda\mathbf{M}]^{\#}\mathbf{M}\big),$$

where the superscript $\#$ denotes adjugate. Hence at $\lambda = 0$ and $p = p_0$,

$$D'(0) = -\text{Trace}\big(\mathbf{L}(p_0)^{\#}\mathbf{M}\big)\,.$$

But

$$\mathbf{L}(p_0)^{\#} = \Pi\,\xi\,\xi^T\,, \tag{3.3}$$

where Π is the product of the nonzero eigenvalues of $\mathbf{L}(p)$. The formula (3.3) is proved as part of Theorem 3 on page 41 of Magnus & Neudecker [21]. Hence

$$D'(0) = -\text{Trace}\big(\mathbf{L}(p_0)^{\#}\mathbf{M}\big) = -\Pi\,\text{Trace}(\xi\xi^T\mathbf{M}\big) = -\Pi\,\langle\mathbf{M}\xi,\xi\rangle\,.$$

The close connection with the formula (1.7) is apparent. The formula (1.7) is a generalization of this case with the product of the nonzero eigenvalues replaced by the Lazutkin homoclinic invariant.

Since Π is the product of the nonzero eigenvalues of $\mathbf{L}(p_0)$, the sign of Π gives the parity of the Morse index, where here the Morse index is just the number of negative eigenvalues of $\mathbf{L}(p_0)$. Hence the ODE version of (1.5) is

$$(-1)^{\mathsf{Morse}} = \text{sign}\big(\Pi\big)\,.$$

4 Geometry of Lagrangian planes

Here and throughout $\mathbb{V}$ is a $4-$dimensional normed vector space. Let

$$\mathbb{V} = \text{span}\big\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\big\} \quad \text{and} \quad \mathbb{V}^* = \text{span}\big\{\mathbf{e}_1^*,\mathbf{e}_2^*,\mathbf{e}_3^*,\mathbf{e}_4^*\big\}\,, \tag{4.1}$$

be bases for $\mathbb{V}$ and the dual space $\mathbb{V}^*$, where $\mathbf{e}_j$ are not necessarily the standard unit vectors. The bases are normalized by $\langle\mathbf{e}_i^*,\mathbf{e}_j\rangle = \delta_{i,j}$, with pairing $\langle\cdot,\cdot\rangle\ :\ \mathbb{V}^*\times\mathbb{V}\to\mathbb{R}$.

Associated with $\mathbb{V}$ and $\mathbb{V}^*$ are the wedge spaces $\bigwedge^k(\mathbb{V})$ and $\bigwedge^k(\mathbb{V}^*)$ for $k = 1,2,3,4$. The convention here on the exterior algebra spaces follows Chapter 4 of Crampin & Pirani [15]. The induced pairing on the wedge spaces is denoted by

$$[\![\cdot,\cdot]\!]_k\ :\ \textstyle\bigwedge^k(\mathbb{V}^*)\times\bigwedge^k(\mathbb{V})\to\mathbb{R}\,,\quad k = 1,2,3,4\,,$$

with $[\![\cdot,\cdot]\!]_1 := \langle\cdot,\cdot\rangle$. The pair $(\mathbb{V},\mathbf{\Omega})$ with

$$\mathbf{\Omega} = \mathbf{e}_1^*\wedge\mathbf{e}_3^* + \mathbf{e}_2^*\wedge\mathbf{e}_4^* \tag{4.2}$$

is a symplectic vector space. The relation between the symplectic form $\mathbf{\Omega}$ and the symplectic operator (2.2), relative to the above basis, is

$$\langle\mathbf{a}\,\lrcorner\,\mathbf{\Omega},\mathbf{b}\rangle = [\![\mathbf{\Omega},\mathbf{a}\wedge\mathbf{b}]\!]_2 = \langle\mathbf{J}\mathbf{a},\mathbf{b}\rangle := \mathbf{\Omega}(\mathbf{a},\mathbf{b})\,,\quad \forall\ \mathbf{a},\mathbf{b}\in\mathbb{V}\,.$$

The first equality is the definition of the interior product, and the second equality follows by evaluating the expression on the bases for $\mathbb{V}$ and $\mathbb{V}^*$, giving (2.2).

On $\mathbb{V}$ and $\mathbb{V}^*$ take the following volume forms

$$\mathsf{vol} := \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3 \wedge \mathbf{e}_4 \quad \text{and} \quad \mathsf{vol}^* := \mathbf{e}_1^* \wedge \mathbf{e}_2^* \wedge \mathbf{e}_3^* \wedge \mathbf{e}_4^* \,.$$

In terms of the symplectic form,

$$\mathsf{vol}^* = -\tfrac{1}{2}\mathbf{\Omega} \wedge \mathbf{\Omega} \,. \tag{4.3}$$

Associated with $(\mathbb{V}, \mathbf{\Omega})$ is a *dual symplectic form*, denoted by $\mathbf{\Omega}^{\text{dual}}$, acting on elements in $\mathbb{V}^*$. The dual symplectic form is defined by

$$\mathbf{\Omega}(\mathbf{a}, \mathbf{b})\mathsf{vol} = \mathbf{\Omega}^{\text{dual}} \wedge \mathbf{a} \wedge \mathbf{b} \,, \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbb{V} \,. \tag{4.4}$$

In this case, a calculation, substituting the bases into (4.4), shows that

$$\mathbf{\Omega}^{\text{dual}} = \mathbf{e}_3 \wedge \mathbf{e}_1 + \mathbf{e}_4 \wedge \mathbf{e}_2 \,.$$

A two-dimensional subspace, $\text{span}\{\mathbf{a}, \mathbf{b}\}$, of $\mathbb{V}$ is a *Lagrangian subspace*, equivalently a *Lagrangian plane*, if

$$\mathbf{\Omega}(\mathbf{a}, \mathbf{b}) = 0 \,. \tag{4.5}$$

The manifold of Lagrangian planes in $\mathbb{V}$ is denoted by $\Lambda(2)$. Other representations of Lagrangian planes needed in the paper are summarized in Appendix B

4.1 Intersection indices

Consider pairs of *oriented* Lagrangian planes and define an intersection index. In what follows, we identify oriented subspaces of $\mathbb{V}$, say $\text{span}\{\mathbf{a}, \mathbf{b}\}$, with the corresponding elements in $\bigwedge^2(\mathbb{V})$, that is $\mathbb{R}_+^*(\mathbf{a}\wedge\mathbf{b})$, where $\mathbb{R}_+^*$ denotes multiplication by an arbitrary positive real number. Let $\mathbf{U}$ and $\mathbf{V}$ be two Lagrangian planes and define

$$d := \dim\big(\mathbf{U} \cap \mathbf{V}\big) \,.$$

The intersection index is denoted by $\mathbf{O}_d(\mathbf{U}, \mathbf{V})$.

Definition Suppose $d = 0$ and let $\mathbf{U} = \text{span}\{\mathbf{a}\wedge\mathbf{b}\}$ and $\mathbf{V} = \text{span}\{\mathbf{c}\wedge\mathbf{d}\}$ then the intersection index of this pair is defined as

$$\mathbf{O}_0(\mathbf{U}, \mathbf{V}) = \text{sign}(\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c} \wedge \mathbf{d}) \,. \tag{4.6}$$

Definition Suppose $d = 1$. Then there exists vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{V}$ such that

$$\mathbf{U} := \text{span}\{\mathbf{a} \wedge \mathbf{b}\} \quad \text{and} \quad \mathbf{V} := \text{span}\{\mathbf{a} \wedge \mathbf{c}\} \,.$$

The intersection index in this case is defined as

$$\mathbf{O}_1(\mathbf{U}, \mathbf{V}) = \text{sign}(\mathbf{\Omega}^{\text{dual}} \wedge \mathbf{b} \wedge \mathbf{c}) \,. \tag{4.7}$$

It is this latter intersection index which is closely associated with the Lazutkin invariant. Evaluating (4.7) on the tangent spaces of the stable and unstable manifolds gives

$$\begin{aligned} \mathbf{O}_1\big(\widehat{\mathbf{u}}_x(x)\wedge\mathbf{a}^-(x),\widehat{\mathbf{u}}_x(x)\wedge\mathbf{a}^+(x)\big) &= \mathrm{sign}\big(\mathbf{\Omega}^{\mathrm{dual}}\wedge\mathbf{a}^-\wedge\mathbf{a}^+\big)\\ &= \mathrm{sign}\big(\mathbf{\Omega}(\mathbf{a}^-,\mathbf{a}^+)\big) \quad \text{(using (4.4))}\\ &= \mathrm{sign}\big(\mathcal{T}(\widehat{\mathbf{u}})\big)\,. \end{aligned}$$

The above intersection indices are for the cases when the dimension $d = 0, 1$. We also need necessary and sufficient conditions for when the intersection is two-dimensional; that is, when the two Lagrangian planes have a complete intersection. It is used in the study of transversely-constructed homoclinic orbits.

Lemma 4.1 *Suppose* $\mathbf{a},\mathbf{b},\mathbf{c}\in\mathbb{V}$ *are such that* $\mathbf{a}\wedge\mathbf{b}\neq 0$ *and* $\mathbf{a}\wedge\mathbf{c}\neq 0$. *Suppose moreover that* $\mathrm{span}\{\mathbf{a},\mathbf{b}\}$ *and* $\mathrm{span}\{\mathbf{a},\mathbf{c}\}$ *are Lagrangian subspaces. Then*

$$\mathrm{span}\{\mathbf{a},\mathbf{b}\} = \mathrm{span}\{\mathbf{a},\mathbf{c}\} \quad \Leftrightarrow \quad \mathbf{\Omega}(\mathbf{b},\mathbf{c}) = 0\,.$$

Proof Let $\mathbf{r}$ and $\mathbf{s}$ be any vectors in $\mathbb{V}$ such that

$$\mathrm{span}\{\mathbf{a},\mathbf{b},\mathbf{r},\mathbf{s}\} = \mathbb{V}\,. \tag{4.8}$$

Then $\mathbf{c}$ can be represented as

$$\mathbf{c} = \alpha_1\mathbf{a} + \alpha_2\mathbf{b} + \alpha_3\mathbf{r} + \alpha_4\mathbf{s}\,,$$

for some constants $\alpha_1,\ldots,\alpha_4$. Now pair this expression with $\mathbf{Ja}$ and $\mathbf{Jb}$ and use the fact that $\mathrm{span}\{\mathbf{a},\mathbf{b}\}$ and $\mathrm{span}\{\mathbf{a},\mathbf{c}\}$ are Lagrangian,

$$\begin{aligned} 0 = \langle\mathbf{Ja},\mathbf{c}\rangle &= \alpha_1\langle\mathbf{Ja},\mathbf{a}\rangle + \alpha_2\langle\mathbf{Ja},\mathbf{b}\rangle + \alpha_3\langle\mathbf{Ja},\mathbf{r}\rangle + \alpha_4\langle\mathbf{Ja},\mathbf{s}\rangle\\ &= \alpha_3\langle\mathbf{Ja},\mathbf{r}\rangle + \alpha_4\langle\mathbf{Ja},\mathbf{s}\rangle\,, \end{aligned}$$

and

$$\begin{aligned} \langle\mathbf{Jb},\mathbf{c}\rangle &= \alpha_1\langle\mathbf{Jb},\mathbf{a}\rangle + \alpha_2\langle\mathbf{Jb},\mathbf{b}\rangle + \alpha_3\langle\mathbf{Jb},\mathbf{r}\rangle + \alpha_4\langle\mathbf{Jb},\mathbf{s}\rangle\\ &= \alpha_3\langle\mathbf{Jb},\mathbf{r}\rangle + \alpha_4\langle\mathbf{Jb},\mathbf{s}\rangle\,, \end{aligned}$$

or

$$\begin{pmatrix} \langle\mathbf{Ja},\mathbf{r}\rangle & \langle\mathbf{Ja},\mathbf{s}\rangle\\ \langle\mathbf{Jb},\mathbf{r}\rangle & \langle\mathbf{Jb},\mathbf{s}\rangle \end{pmatrix}\begin{pmatrix}\alpha_3\\ \alpha_4\end{pmatrix} = \begin{pmatrix} 0\\ \langle\mathbf{Jb},\mathbf{c}\rangle\end{pmatrix}.$$

But the coefficient matrix is invertible. This follows from (4.8) and the formula (1.8),

$$0\neq\mathbf{a}\wedge\mathbf{b}\wedge\mathbf{r}\wedge\mathbf{s} = \det\begin{bmatrix}\mathbf{\Omega}(\mathbf{a},\mathbf{r}) & \mathbf{\Omega}(\mathbf{a},\mathbf{s})\\ \mathbf{\Omega}(\mathbf{b},\mathbf{r}) & \mathbf{\Omega}(\mathbf{b},\mathbf{s})\end{bmatrix}\,\mathsf{vol}\,.$$

Hence $\alpha_3 = \alpha_4 = 0$ if and only if $\langle\mathbf{Jb},\mathbf{c}\rangle = 0$ and the Lemma is proved. □

5 Transversely constructed homoclinic orbits

Suppose there exists a homoclinic orbit, $\widehat{\mathbf{u}}(x,p)$, satisfying the steady part of (1.1)

$$\mathbf{J}\mathbf{u}_x = DH(\mathbf{u},p)\,, \quad \mathbf{u}\in\mathbb{V}\,, \tag{5.1}$$

with

$$\lim_{x\to\pm\infty}\widehat{\mathbf{u}}(x,p)=0 \quad\text{and}\quad 0<\int_{-\infty}^{+\infty}\left|\widehat{\mathbf{u}}(x,c)\right|^2\mathrm{d}x<+\infty\,. \tag{5.2}$$

The linearization about the trivial solution is assumed to be strictly hyperbolic.

The tangent vector to the homoclinic orbit is $\widehat{\mathbf{u}}_x$, and the orbit lies on an energy surface $H(\widehat{\mathbf{u}},p)=H(0,p)$. The stable and unstable manifolds of the origin also lie on the energy surface. Hence, there are two other tangent vectors in $\mathbb{V}$, denoted by $\mathbf{a}^-(x)$ and $\mathbf{a}^+(x)$ satisfying

$$\frac{d}{dx}\mathbf{a}^{\pm}=\mathbf{A}(x,0)\mathbf{a}^{\pm}\,, \quad\text{with}\quad \mathbf{A}(x,0):=\mathbf{J}^{-1}D^2H(\widehat{\mathbf{u}},p)\,, \tag{5.3}$$

and

$$E^s(x,0)=\mathrm{span}\{\widehat{\mathbf{u}}_x,\mathbf{a}^+\} \quad\text{and}\quad E^u(x,0)=\mathrm{span}\{\widehat{\mathbf{u}}_x,\mathbf{a}^-\}\,.$$

The notation with 0 in the second argument anticipates the extension to include λ dependence, and the explicit dependence on p is suppressed for brevity. The subspaces $E^{s,u}$ are $x-$dependent Lagrangian subspaces. This property is proved in §4 of [12].

Definition A homoclinic orbit is said to be "transversely constructed" if $\mathbf{a}^+$ and $\mathbf{a}^-$ are linearly independent for all x

$$\mathbf{a}^-(x)\wedge\mathbf{a}^+(x)\neq 0 \quad\text{for all } x\in\mathbb{R}\,. \tag{5.4}$$

Proposition 5.1 *If $\mathbf{a}^-(x_0)\wedge\mathbf{a}^+(x_0)\neq 0$ for some $x_0\in\mathbb{R}$ then the condition (5.4) is satisfied.*

Proof The proof follows from the fact that $\mathbf{a}^-(x)\wedge\mathbf{a}^+(x)$ satisfies an ODE

$$\frac{d}{dx}\mathbf{a}^-\wedge\mathbf{a}^+=\mathbf{A}^{(2)}(x,0)\mathbf{a}^-\wedge\mathbf{a}^+\,,$$

and the uniqueness of solutions of ODEs, where $\mathbf{A}^{(2)}(x,0)$ is the induced representation of $\mathbf{A}(x,0)$ on $\bigwedge^2(\mathbb{V})$. □

Definition The Lazutkin invariant of a homoclinic orbit is $\mathcal{T}(\widehat{\mathbf{u}})=\boldsymbol{\Omega}(\mathbf{a}^-,\mathbf{a}^+)$.

Theorem 5.2 *A homoclinic orbit is "transversely constructed" if and only if the Lazutkin homoclinic invariant is nonzero.*

Proof The Lazutkin invariant is independent of x. This follows since $\mathbf{a}^{\pm}$ are solutions of (5.3) and the symplectic form is independent of x when evaluated on any two solutions of (5.3).

Now suppose the Lazutkin invariant is zero $\mathbf{\Omega}(\mathbf{a}^-,\mathbf{a}^+)=0$. Then by Lemma 4.1, the stable and unstable subspaces have a two-dimensional intersection (for each x) and so the intersection is not transverse.

Conversely, suppose $\mathbf{a}^-(x_0)\wedge\mathbf{a}^+(x_0)=0$ for some x_0. Then it is zero for all x by Proposition 5.1. Hence $\mathbf{\Omega}^{\rm dual}\wedge\mathbf{a}^-\wedge\mathbf{a}^+=0$ and so

$$0=\mathbf{\Omega}^{\rm dual}\wedge\mathbf{a}^-\wedge\mathbf{a}^+=[\![\mathsf{vol}^*,\mathbf{\Omega}^{\rm dual}\wedge\mathbf{a}^-\wedge\mathbf{a}^+]\!]_4=[\![\mathbf{\Omega},\mathbf{a}^-\wedge\mathbf{a}^+]\!]_2=\mathbf{\Omega}(\mathbf{a}^-,\mathbf{a}^+)\,,$$

proving that the Lazutkin invariant is zero. □

Hence the Lazutkin invariant measures whether the codimension one intersection of the (Lagrangian) stable and unstable subspaces is non-degenerate. It is invariant under symplectic change of coordinates. However, in order to fix the sign of the Lazutkin invariant, and to define the sign of a homoclinic orbit, a normalization needs to be introduced.

Definition Suppose that the stable and unstable subspaces are normalized as follows

$$\lim_{x\to+\infty}\mathrm{e}^{2(\mu_1+\mu_2)x}\,\widehat{\mathbf{u}}_x(-x)\wedge\mathbf{a}^+(-x)\wedge\widehat{\mathbf{u}}_x(+x)\wedge\mathbf{a}^-(+x)=\mathsf{vol}\,. \tag{5.5}$$

Then the sign of the homoclinic orbit is

$$\mathrm{sign}(\mathcal{T})=\mathrm{sign}\left(\mathbf{\Omega}^{\rm dual}\wedge\mathbf{a}^+\wedge\mathbf{a}^-\right)=\mathrm{sign}\left(\Omega(\mathbf{a}^-,\mathbf{a}^+)\right)\,. \tag{5.6}$$

The exponents μ_1 and μ_2 are the decay rates at infinity for the stable subspace. They are defined in §7. It is not obvious that the solutions can be normalized so that the coefficient of the volume form on the right-hand side of (5.5) is unity. A proof that this normalization is possible is given in Appendix C.2.

5.1 Transversality on an energy surface

The transversality is not in $\mathbb{V}$ but on the level surface of the Hamiltonian function. A representation for the Lazutkin invariant on the energy surface can also be derived. Define a volume form on the energy surface, $\mathsf{vol}^{\rm Energy}$, by

$$DH\wedge\mathsf{vol}^{\rm Energy}=\mathsf{vol}^*=-\tfrac{1}{2}\mathbf{\Omega}\wedge\mathbf{\Omega}\,.$$

Contract both sides with $\widehat{\mathbf{u}}_x$

$$-\widehat{\mathbf{u}}_x\lrcorner\left(\tfrac{1}{2}\mathbf{\Omega}\wedge\mathbf{\Omega}\right)=\widehat{\mathbf{u}}_x\lrcorner\left(DH\wedge\mathsf{vol}^{\rm Energy}\right),$$

or, after using (A-3), $\widehat{\mathbf{u}}_x\lrcorner DH=0$ and $\widehat{\mathbf{u}}_x\lrcorner\mathbf{\Omega}=DH$,

$$DH\wedge\mathbf{\Omega}=DH\wedge\left(\widehat{\mathbf{u}}_x\lrcorner\mathsf{vol}^{\rm Energy}\right),$$

or

$$\widehat{\mathbf{u}}_x \lrcorner\, \mathsf{vol}^{\rm Energy} = \mathbf{\Omega}\,.$$

Contracting with $\mathbf{a}^+$ and $\mathbf{a}^-$ then gives an energy surface interpretation of the Lazutkin invariant

$$\mathfrak{I}(\widehat{\mathbf{u}}) = \mathbf{\Omega}(\mathbf{a}^-,\mathbf{a}^+) = \mathsf{vol}^{\rm Energy}(\widehat{\mathbf{u}}_x,\mathbf{a}^-,\mathbf{a}^+)\,.$$

6 Example: an explicit calculation of $\mathfrak{I}(\widehat{\mathbf{u}})$

Take $d_1 = d_2$ and

$$F(v,w) = -2(v^2+w^2) + 2(v^3+w^3) - \tfrac{1}{2}p(v-w)^2\,,$$

in (1.3). The resulting pair of gradient reaction-diffusion equations is

$$\begin{array}{rcl} v_t &=& v_{xx} - 4v + 6v^2 - p(v-w) \\ w_t &=& w_{xx} - 4w + 6w^2 + p(v-w)\,. \end{array} \tag{6.1}$$

This system was studied in §11 of [12] (with the parameter p here replaced by c there). The system (6.1) has an exact steady solution $v = w := \widehat{v}(x) = {\rm sech}^2(x)$. It is an example where the Maslov index and other geometric properties of the linearization about the steady solution can be explicitly computed. Here the Lazutkin invariant is calculated.

The tangent vector to the homoclinic orbit is

$$\widehat{\mathbf{u}}_x = -2{\rm sech}^2(x)\begin{pmatrix} \tanh(x) \\ \tanh(x) \\ 1-3\tanh^2(x) \\ 1-3\tanh^2(x) \end{pmatrix}\,,$$

and the complementary vectors $\mathbf{a}^\pm(x)$ are

$$\mathbf{a}^\pm(x) = \begin{pmatrix} -\sigma^\pm(x) \\ +\sigma^\pm(x) \\ -\sigma^\pm_x(x) \\ +\sigma^\pm_x(x) \end{pmatrix}\,,$$

where

$$\sigma^\pm(x) = {\rm e}^{\mp\sqrt{\kappa}x}(\mp a_0 + a_1\tanh(x) \mp a_2\tanh^2(x) + \tanh^3(x))\,,$$

with $\kappa = 4+2p$ and

$$a_0 = \frac{\sqrt{\kappa}}{15}(4-\kappa)\,,\quad a_1 = \frac{1}{5}(2\kappa-3)\,,\quad a_2 = -\sqrt{\kappa}\,.$$

Computing

$$\mathcal{T}(\widehat{\mathbf{u}})\mathsf{vol} = \mathbf{\Omega}^{\rm dual} \wedge \mathbf{a}^- \wedge \mathbf{a}^+ \,,$$

gives

$$\mathcal{T}(\widehat{\mathbf{u}}) = \frac{8p}{225}\sqrt{4+2p}(3+2p)(5-2p)\,.$$

Transversality of the construction of the homoclinic orbit is lost precisely when

$$p = -\frac{3}{2}\,, \quad p = 0\,, \quad p = \frac{5}{2}\,. \tag{6.2}$$

The above form for $\mathbf{a}^\pm$ is chosen so that the normalization (5.5) is operational. Therefore the formula (1.5) should hold. Indeed this can be checked directly. According to §11 of Part 1, the values of p in (6.2) are precisely the values where the Maslov index of the homoclinic orbit changes. The Maslov index is 2 for $0 < p < \frac{5}{2}$ and 1 for $p > \frac{5}{2}$. Hence confirming the relation

$$(-1)^{\mathsf{Maslov}} = {\rm sign}\big(\mathcal{T}(\widehat{\mathbf{u}})\big)\,.$$

This example also emphasizes that the Lazutkin invariant is not an invariant of the homoclinic orbit directly. It is a property of the intersection between the stable and unstable manifolds. Here the basic homoclinic orbit, and its tangent vector $\widehat{\mathbf{u}}_x$, are independent of the parameter p, but the complementary tangent vectors $\mathbf{a}^\pm$ are dependent on p and they determine when there is a loss of transversality.

7 The Evans function

Suppose that the Hamiltonian system (5.1) has a homoclinic orbit as in §5. Consider the linearization of the PDE (1.1) about the homoclinic orbit $\widehat{\mathbf{u}}$

$$\mathbf{M}\mathbf{u}_t + \mathbf{J}\mathbf{u}_x = \mathbf{B}(x,p)\mathbf{u}\,, \quad \mathbf{u} \in \mathbb{V}\,.$$

where $\mathbf{B}$ is the Hessian of H evaluated on the homoclinic orbit,

$$\mathbf{B}(x,p) = {\rm D}^2 H(\widehat{\mathbf{u}},p)\,.$$

Letting $\mathbf{u} = {\rm e}^{\lambda t}\widetilde{\mathbf{u}}$ results in the spectral problem, which will be formulated in the following way in preparation for the use of the Evans function theory

$$\mathbf{u}_x = \mathbf{A}(x,\lambda)\mathbf{u}\,, \quad \mathbf{u} \in \mathbb{V}\,, \tag{7.1}$$

with

$$\mathbf{A}(x,\lambda) = \mathbf{J}^{-1}[\mathbf{B}(x,p) - \lambda\mathbf{M}]\,. \tag{7.2}$$

The tilde over $\mathbf{u}$ has been dropped to simplify notation. The role of $\mathbf{u}(x,\lambda)$ versus $\mathbf{u}(x,t)$ will be clear from the context.

The "system at infinity", $\mathbf{A}_\infty(\lambda)$, that is used in the construction of the Evans function is defined by

$$\mathbf{J}\mathbf{A}_\infty(\lambda) = \big[\mathbf{B}_\infty - \lambda\mathbf{M}\big]\,, \tag{7.3}$$

with $\mathbf{B}_\infty = \lim_{x\to\pm\infty}\mathbf{B}(x,p)$, with the dependence on p suppressed.

The formal definition of an eigenvalue is: $\lambda\in\mathbb{C}$ is an eigenvalue of of (7.1) if there exists $\mathbf{u}(x,\lambda)$ such that

$$\int_{-\infty}^{+\infty} \|\mathbf{u}(x,\lambda)\|^2\,\mathrm{d}x < +\infty\,,$$

where $\|\cdot\|$ is a norm on $\mathbb{V}$.

In fact we will restrict attention to real λ, which can almost be proved in general. Suppose λ and $\mathbf{u}$ are complex,

$$\mathbf{u} = \mathbf{u}_r + \mathrm{i}\mathbf{u}_i \quad \text{and} \quad \lambda = \lambda_r + \mathrm{i}\lambda_i\,.$$

Substitute into (7.1), take real and imaginary parts, pair with $\mathbf{u}_i$ and $\mathbf{u}_r$ in turn, giving

$$\frac{d}{dx}\big(\mathbf{\Omega}(\mathbf{u}_r,\mathbf{u}_i)\big) = \lambda_i\big(\langle\mathbf{M}\mathbf{u}_r,\mathbf{u}_r\rangle + \langle\mathbf{M}\mathbf{u}_i,\mathbf{u}_i\rangle\big)\,.$$

Integrating over x and using $\|\mathbf{u}\|\to 0$ as $x\to\pm\infty$ gives

$$\lambda_i\int_{-\infty}^{+\infty}\big(\langle\mathbf{M}\mathbf{u}_r,\mathbf{u}_r\rangle + \langle\mathbf{M}\mathbf{u}_i,\mathbf{u}_i\rangle\big)\,\mathrm{d}x = 0\,.$$

If $\mathbf{M}$ is non-degenerate, $\lambda_i = 0$ and the argument is proved, but since $\mathbf{M}$ may have zero eigenvalues there may be exceptions. Here we will assume that the exceptions don't occur and take λ to be real throughout.

The essential spectrum is defined to be

$$\sigma_{ess} = \big\{\lambda\in\mathbb{R}\;\big|\;\det[\mathbf{B}_\infty - \mathrm{i}k\mathbf{J} - \lambda\mathbf{M}] = 0 \text{ with } k\in\mathbb{R}\big\}\,.$$

We will assume that

$$\sup_{\lambda}\sigma_{ess} < 0\,. \tag{7.4}$$

Now the Evans function can be constructed in the usual way. Denote the eigenvalues of $\mathbf{A}_\infty(\lambda)$ with negative real part by $\mu_1(\lambda)$ and $\mu_2(\lambda)$ and the eigenvalues with positive real part by $\mu_3(\lambda)$ and $\mu_4(\lambda)$, with eigenvectors

$$\big[\mathbf{B}_\infty - \lambda\mathbf{M}\big]\xi_i = \mu_i\mathbf{J}\xi_i\,,\quad i = 1,\ldots,4\,. \tag{7.5}$$

Due to the symplectic symmetry, $\mu_3 = -\mu_1$ and $\mu_4 = -\mu_2$. With the assumption of strict hyperbolicity, there are three possible configurations for the eigenvalues: (a) four hyperbolic real distinct eigenvalues, (b) two real eigenvalues each with multiplicity two, and (c) a hyperbolic complex quartet as shown qualitatively in Figure 1. To

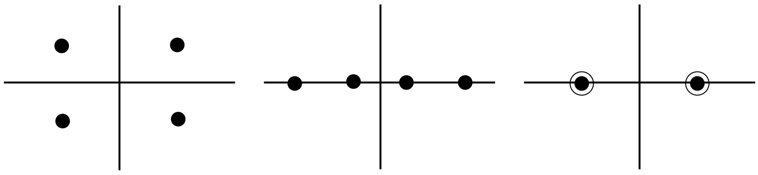

Figure 1: The 3 cases of hyperbolic eigenvalues on a $4-$dimensional space

simplify the exposition, assume that we are in case (a) and the $\mu_j(\lambda)$ and their associated eigenvectors are real. The complex case can be treated by splitting into real and imaginary parts. For the double eigenvalues generalized eigenvectors are used. In the latter case there is a loss of analyticity in the $\lambda-$plane near double eigenvalues, but this issue is well understood and so is not considered here [7].

Now define solutions of (7.1) that decay to zero as $x \to +\infty$ with the asymptotic properties

$$\lim_{x\to+\infty} \mathrm{e}^{-\mu_j(\lambda)x}\mathbf{u}_j^+(x,\lambda) = \xi_j(\lambda)\,, \quad j=1,2, \tag{7.6}$$

and solutions which decay as $x \to -\infty$,

$$\lim_{x\to-\infty} \mathrm{e}^{-\mu_j(\lambda)x}\mathbf{u}_j^-(x,\lambda) = \xi_j(\lambda)\,, \quad j=3,4\,. \tag{7.7}$$

Then the natural definition of the Evans function is

$$D(\lambda)\mathsf{vol} = \mathbf{u}_1^+(x,\lambda)\wedge\mathbf{u}_2^+(x,\lambda)\wedge\mathbf{u}_3^-(x,\lambda)\wedge\mathbf{u}_4^-(x,\lambda)\,. \tag{7.8}$$

It has the usual properties of an Evans function (cf. Alexander, Gardner & Jones [1]). In particular, $D(0)=0$ since $\widehat{\mathbf{u}}_x$ is a solution of (7.1) with $\lambda = 0$.

7.1 Symplectification of the Evans function

By working directly with the Evans function as a four-form (7.8) it is not immediately clear how to take advantage of the symplectic structure. The symplectic structure was brought out in the construction of the symplectic Evans function in Bridges & Derks [6]. However, there the solutions of the adjoint of (7.1) were required. Here we avoid use of the adjoint solutions by using the new formula (1.8). Indeed, since the stable and unstable subspaces are Lagrangian, the correction term vanishes and we find the following new formula for the Evans function

$$D(\lambda) = \det\begin{bmatrix}\mathbf{\Omega}(\mathbf{u}_3^-,\mathbf{u}_1^+) & \mathbf{\Omega}(\mathbf{u}_3^-,\mathbf{u}_2^+)\\ \mathbf{\Omega}(\mathbf{u}_4^-,\mathbf{u}_1^+) & \mathbf{\Omega}(\mathbf{u}_4^-,\mathbf{u}_2^+)\end{bmatrix}\,. \tag{7.9}$$

In this form, a symplectic proof that $D(0)=0$ can be given. In the limit $\lambda=0$,

$$\begin{array}{rcl}\mathbf{u}_1^+\big|_{\lambda=0} &=& \widehat{\mathbf{u}}_x \quad\text{and}\quad \mathbf{u}_2^+\big|_{\lambda=0} = \mathbf{a}^+\\ \mathbf{u}_3^-\big|_{\lambda=0} &=& \widehat{\mathbf{u}}_x \quad\text{and}\quad \mathbf{u}_4^-\big|_{\lambda=0} = \mathbf{a}^-\,.\end{array} \tag{7.10}$$

In general, in the $\lambda \to 0$ limit,

$$\text{span}\big\{\mathbf{u}_1^+(x,0), \mathbf{u}_2^+(x,0)\big\} = \text{span}\big\{\widehat{\mathbf{u}}_x, \mathbf{a}^+\big\}\,.$$

However, the Evans function can be suitably scaled so that the limits (7.10) are achieved (see Appendix C.3 for this argument).

With the limits (7.10), the Evans function at $\lambda = 0$ is

$$D(0) = \lim_{\lambda\to 0} \det \begin{bmatrix} \mathbf{\Omega}(\mathbf{u}_3^-, \mathbf{u}_1^+) & \mathbf{\Omega}(\mathbf{u}_3^-, \mathbf{u}_2^+) \\ \mathbf{\Omega}(\mathbf{u}_4^-, \mathbf{u}_1^+) & \mathbf{\Omega}(\mathbf{u}_4^-, \mathbf{u}_2^+) \end{bmatrix} = \det \begin{bmatrix} \mathbf{\Omega}(\widehat{\mathbf{u}}_x, \widehat{\mathbf{u}}_x) & \mathbf{\Omega}(\widehat{\mathbf{u}}_x, \mathbf{a}^+) \\ \mathbf{\Omega}(\mathbf{a}^-, \widehat{\mathbf{u}}_x) & \mathbf{\Omega}(\mathbf{a}^-, \mathbf{a}^+) \end{bmatrix}\,.$$

Now $\mathbf{\Omega}(\widehat{\mathbf{u}}_x, \widehat{\mathbf{u}}_x) = 0$ due to skew-symmetry, and $\mathbf{\Omega}(\widehat{\mathbf{u}}_x, \mathbf{a}^\pm) = 0$ due to the fact that the stable and unstable subspaces are Lagrangian. Hence

$$D(0) = \det \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{\Omega}(\mathbf{a}^-, \mathbf{a}^+) \end{bmatrix} = 0\,.$$

7.2 The derivative at $\lambda = 0$

Define the matrix in the definition of the symplectic Evans function as

$$\mathscr{D}(\lambda) = \begin{bmatrix} \mathbf{\Omega}(\mathbf{u}_3^-, \mathbf{u}_1^+) & \mathbf{\Omega}(\mathbf{u}_3^-, \mathbf{u}_2^+) \\ \mathbf{\Omega}(\mathbf{u}_4^-, \mathbf{u}_1^+) & \mathbf{\Omega}(\mathbf{u}_4^-, \mathbf{u}_2^+) \end{bmatrix}\,.$$

Then $D(\lambda) = \det[\mathscr{D}(\lambda)]$ and

$$\mathscr{D}(0) = \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{\Omega}(\mathbf{a}^-, \mathbf{a}^+) \end{bmatrix}\,.$$

Now use the standard formula for the derivative of a determinant

$$D'(\lambda) = \text{Trace}\big(\mathscr{D}(\lambda)^\# \mathscr{D}'(\lambda)\big)\,,$$

where the superscript $\#$ denotes adjugate. Evaluation at $\lambda = 0$ gives

$$D'(0) = \text{Trace}\big(\mathscr{D}(0)^\# \mathscr{D}'(0)\big)\,.$$

But

$$\mathscr{D}(0)^\# = \mathbf{\Omega}(\mathbf{a}^-, \mathbf{a}^+) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\,,$$

and so

$$D'(0) = \text{Trace}\Big(\mathscr{D}(0)^\# \mathscr{D}'(0)\Big) = \mathbf{\Omega}(\mathbf{a}^-, \mathbf{a}^+) \frac{d}{d\lambda} \mathbf{\Omega}(\mathbf{u}_3^-, \mathbf{u}_1^+)\Big|_{\lambda=0}\,.$$

It remains to prove the following

Proposition 7.1

$$\frac{d}{d\lambda}\boldsymbol{\Omega}(\mathbf{u}_3^-, \mathbf{u}_1^+)\bigg|_{\lambda=0} = -\int_{-\infty}^{+\infty} \langle \mathbf{M}\widehat{\mathbf{u}}_x, \widehat{\mathbf{u}}_x \rangle \, \mathrm{d}x \,. \tag{7.11}$$

This formula is very similar to $\lambda-$derivatives in previous work and a brief proof is given in Appendix D. This completes the proof of the formula (1.7). Moreover, we have proved the following result:

Theorem 7.2 *Suppose*

$$\int_{-\infty}^{+\infty} \langle \mathbf{M}\widehat{\mathbf{u}}_x, \widehat{\mathbf{u}}_x \rangle \, \mathrm{d}x > 0 \,.$$

Then $\lambda = 0$ is a simple eigenvalue of the Evans function if and only if the homoclinic orbit is transversely constructed.

This result is to be contrasted with the non-Hamiltonian case. For a class of parabolic reaction-diffusion equations, Alexander & Jones [2], prove that the Evans function has a simple zero if and only if the homoclinic orbit is transversely constructed (see Theorem 2.2 on page 59 of [2], and Theorem 4.1 on page 212 of [3]). In the Hamiltonian case the derivative $D'(0)$ is related to the symplectic invariant $\mathcal{T}(\widehat{\mathbf{u}})$.

8 Example: transversality for Swift-Hohenberg

Suppose that the Swift-Hohenberg equation (1.2) has a steady solitary wave, represented by a homoclinic orbit solution $\widehat{\phi}(x,p)$. Assume that it satisfies the basic properties

$$\lim_{x\to\pm\infty} \widehat{\phi}(x,p) = 0 \quad \text{and} \quad 0 < \int_{-\infty}^{+\infty} |\widehat{\phi}_x|^2 \, \mathrm{d}x < +\infty \,. \tag{8.1}$$

It could be a simple homoclinic orbit or a multi-pulse homoclinic orbit. Such solutions have been studied in [8, 11, 24]. The linearization about such solutions in the time-dependent equation, with in addition a spectral ansatz, leads to the spectral problem

$$\mathscr{L}\phi = \lambda\phi \,, \tag{8.2}$$

where

$$\mathscr{L}\phi := -\phi_{xxxx} - p\phi_{xx} - \phi + 2\widehat{\phi}\phi \,. \tag{8.3}$$

The theory of this paper leads to a new proof of Lemma 2.1(iii) in [24].

Lemma 8.1 (Sandstede [24]). *Any homoclinic orbit of the steady SH equation with $-2 < p < 2$ is transversely constructed if and only if zero is a simple eigenvalue of $\mathscr{L}$ in (8.2).*

Proof The spectral problem (8.2) can be recast into the form (7.1). The hypothesis on the essential spectrum is satisfied for $-2 < p < +2$, and with the properties (8.1) the formula (1.7) applies. Hence

$$D'(0) = -\mathfrak{I}(\widehat{\mathbf{u}}) \int_{-\infty}^{+\infty} \langle \mathbf{M}\widehat{\mathbf{u}}_x, \widehat{\mathbf{u}}_x \rangle \,\mathrm{d}x = -\mathfrak{I}(\widehat{\mathbf{u}}) \int_{-\infty}^{+\infty} \widehat{\phi}_x^2 \,\mathrm{d}x \,,$$

using the form of $\mathbf{M}$ in (2.3). Hypothesis (8.1) assures that the integral exists and is non-vanishing. Hence $D'(0) = 0$ if and only if $\mathfrak{I}(\widehat{\mathbf{u}})$. The proof is completed by applying Theorem 5.2. □.

The spectral problem here (8.2) is simple enough so that the Maslov index equals the Morse index of $\mathscr{L}$. Hence the formula (1.5) can be cast into a formula for the parity of the Morse index. The Morse index for a wide range of multi-pulse homoclinic orbits of the steady SH equation is computed in [11].

9 The Maslov index of homoclinic orbits

To prove the formula (1.5) we need a definition for the Maslov index of homoclinic orbits. Here we will use a formulation due to Souriau [25], following Chapter 1 of [10] where Souriau's formulation is used to show how the formulations of the Maslov index of Bose & Jones [3], Chardard [9], Chen & Hu [14] are related. In addition a new slightly modified definition is given here, which makes it easier to compare with the Lazutkin invariant.

Souriau's definition is formulated on the universal cover of the Lagrangian Grassmannian manifold. For simplicity we define it on the universal covering of the unitary group

$$\pi : \widetilde{U(2)} \to \Lambda(2) \,, \quad (\mathbf{U}, \kappa) \mapsto \text{the space spanned by } \mathbf{U} \,,$$

with

$$\widetilde{U(2)} = \left\{ (\mathbf{U}, \kappa) \;\Big|\; \mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} \,,\; \mathbf{U}_1 + \mathrm{i}\mathbf{U}_1 \in U(2) \,,\; \mathrm{e}^{-\mathrm{i}\frac{\kappa}{2}} = \det(\mathbf{U}_1 + \mathrm{i}\mathbf{U}_2) \right\} . \tag{9.1}$$

Let $\mathbf{U}$ and $\mathbf{V}$ be two Lagrangian planes in the unitary representation,

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} \quad \text{with} \quad \mathbf{U}_1 + \mathrm{i}\mathbf{U}_2 \in U(2) \,,$$

with a similar definition for $\mathbf{V}$, and define the mapping

$$\psi(\mathbf{U}, \mathbf{V}) = (\mathbf{U}_1 - \mathrm{i}\mathbf{U}_2)^{-1}(\mathbf{V}_1 - \mathrm{i}\mathbf{V}_2)(\mathbf{V}_1 + \mathrm{i}\mathbf{V}_2)^{-1}(\mathbf{U}_1 + \mathrm{i}\mathbf{U}_2) \,. \tag{9.2}$$

$\psi(\mathbf{U}, \mathbf{V})$ is a symmetric unitary matrix and its eigenvalues lie on the unit circle (cf. §1.1.2 of [10]). Hence the eigenvalues can be expressed in the form

$$\sigma\big(\psi(\mathbf{U}, \mathbf{V})\big) = \big\{ \mathrm{e}^{\mathrm{i}\alpha_1(\mathbf{U},\mathbf{V})}, \mathrm{e}^{\mathrm{i}\alpha_2(\mathbf{U},\mathbf{V})} \big\} \,, \quad 0 \le \alpha_1, \alpha_2 < 2\pi \,.$$

To lighten the notation, we will drop the arguments on α_1 and α_2 and their dependence will be clear from the context.

Let d be the intersection of the two Lagrangian planes spanned by the columns of $\mathbf{U}$ and $\mathbf{V}$.

Proposition 9.1 *The dimension of the eigenspace of ψ associated to the eigenvalue $+1$ is equal to d.*

Proof First prove that the dimension of the eigenspace is greater than d. The space

$$\{\gamma \in \mathbb{R}^{1\times 2} \mid \exists\, \beta \text{ such that } \gamma \mathbf{U} = \beta \mathbf{V}\}\,,$$

has dimension d. Let $\gamma = (\gamma_1, \gamma_2)$ and $\beta = (\beta_1, \beta_2)$ be such that $\gamma\mathbf{U} = \beta\mathbf{V}$. In the unitary representation

$$\gamma[\mathbf{U}_1 - \mathrm{i}\mathbf{U}_2] = \beta[\mathbf{V}_1 - \mathrm{i}\mathbf{V}_2]\,,$$

or $\gamma(\mathbf{U}_1 - \mathrm{i}\mathbf{U}_2)^{-1}(\mathbf{V}_1 - \mathrm{i}\mathbf{V}_2) = \beta$. Hence

$$\gamma\psi = \beta(\mathbf{U}_1 + \mathrm{i}\mathbf{U}_2)^{-1}(\mathbf{V}_1 + \mathrm{i}\mathbf{V}_2) = \gamma\,,$$

proving that ψ has eigenvalue 1 with left eigenvector γ. Therefore the eigenspace of ψ associated to 1 has a dimension greater to or equal to d.

Reciprocally, if γ is a left eigenvector of ψ associated to 1, then $\widetilde{\gamma}$ is also an eigenvector of ψ associated to 1. Hence, $\mathbb{X} = \{\gamma \in \mathbb{R}^{1\times 2} \mid \gamma\psi = \gamma\}$ has the same dimension as the eigenspace of ψ associated to the the eigenvalue $+1$.

Let $\gamma \in \mathbb{X}$ and let $\beta = \gamma(\mathbf{U}_1 - \mathrm{i}\mathbf{U}_2)^{-1}(\mathbf{V}_1 - \mathrm{i}\mathbf{V}_2)$, then $\gamma\mathbf{U} = \beta\mathbf{V}$ is in the set $\pi((\mathbf{U},\kappa)) \cap \pi((\mathbf{V},\tau))$. Therefore the eigenspace of ψ associated to the eigenvalue $+1$ has a dimension equal to d. □

According to Souriau's formula (cf. pages 126–128 of [25]), the Maslov index of this pair of elements is defined in the following way.

Definition Let $(\mathbf{U},\kappa)$ and $(\mathbf{V},\tau)$ be in $\widetilde{U(2)}$. The Maslov index of this pair of elements is defined by:

$$m\big((\mathbf{U},\kappa),(\mathbf{V},\tau)\big) = \frac{\tau-\kappa}{2\pi} - \frac{\alpha_1+\alpha_2}{2\pi} + \frac{1}{2}d$$

where

$$d := \dim\big(\pi((\mathbf{U},\kappa)) \cap \pi((\mathbf{V},\tau))\big)\,.$$

Now, the definition for the Maslov index of a homoclinic orbit, based on Souriau's definition above, is as follows. Let $\widehat{\mathbf{u}}$ be an homoclinic orbit. Let $(\mathbf{U}^+,\kappa^+)$ and $(\mathbf{U}^-,\kappa^-)$ in $\widetilde{U(2)}$ be such that:

$$\begin{cases} \text{span } \mathbf{U}^+(x) = \text{span}\{\widehat{\mathbf{u}}_x(x), \mathbf{a}^+(x)\} \\ \text{span } \mathbf{U}^-(x) = \text{span}\{\widehat{\mathbf{u}}_x(x), \mathbf{a}^-(x)\} \\ \kappa^+, \kappa^- \text{ are continuous} \end{cases}.$$

Then, the Maslov index of the homoclinic is defined by:

$$\begin{aligned} I_{hom}(\widehat{\mathbf{u}}) &= m\big((\mathbf{U}^-(x),\kappa^-(x)),(\mathbf{U}^+(x),\kappa^+(x))\big) \\ &\quad - \lim_{y\to+\infty} m\big((\mathbf{U}^-(-y),\kappa^-(-y)),(\mathbf{U}^+(y),\kappa^+(y))\big)\,, \end{aligned} \tag{9.3}$$

and the definition is independent of x.

10 Transversality and parity of the Maslov index

The purpose of this section is to prove the following connection between the parity of the Maslov index and transversality.

Theorem 10.1 *Suppose* $\widehat{\mathbf{u}}$ *is a transversely constructed homoclinic orbit with Maslov index* $I_{\rm hom}(\widehat{\mathbf{u}})$. *Then*

$$(-1)^{\mathsf{Maslov}} = \mathrm{sign}\big(\mathcal{T}(\widehat{\mathbf{u}})\big)\,, \quad \mathsf{Maslov} := I_{\rm hom}(\widehat{\mathbf{u}}) + \frac{1}{2}\,.$$

The key point to prove the relationship between the Maslov index and the intersection index for the tangent spaces of the stable and unstable subspaces lies in the following lemma:

Lemma 10.2 *Let* $(\mathbf{U},\kappa),(\mathbf{V},\tau)\in\widetilde{U(2)}$ *such that* $d=\dim\big(\pi((\mathbf{U},\kappa))\cap\pi((\mathbf{V},\tau))\big)$ *with* $d=0,1$, *and let* $\mathbf{U}^\wedge$ *and* $\mathbf{V}^\wedge$ *be the corresponding 2-forms. Then*

$$\mathbf{O}_d(\mathbf{U}^\wedge,\mathbf{V}^\wedge) = (-1)^{m+\frac{1}{2}d}\,, \quad m := m((\mathbf{U},\kappa),(\mathbf{V},\tau))\,.$$

Proof First consider the case $d=1$. Define the function

$$g(\mathbf{U},\mathbf{V}) = (-1)^{m+\frac{1}{2}}\,\mathbf{O}_1(\mathbf{U}^\wedge,\mathbf{V}^\wedge)\,,$$

where κ,τ are taken such that $(\mathbf{U},\kappa),(\mathbf{V},\tau)\in\widetilde{U(2)}$. We need to prove that $g(\mathbf{U},\mathbf{V})=1$.

The function g is well defined, because if $(\mathbf{U},\kappa)$ and $\mathbf{U},\mu)$ are in $\widetilde{U(2)}$, then $\kappa\in\mu+4\pi\mathbb{Z}$ and therefore $m((\mathbf{U},\kappa),(\mathbf{V},\tau))-m((\mathbf{U},\mu),(\mathbf{V},\tau))\in 2\mathbb{Z}$, which proves that $(-1)^{m+\frac{1}{2}}\mathbf{O}_1(\mathbf{U}^\wedge,\mathbf{V}^\wedge)$ does not depend on the particular choice of κ and τ.

Now let $\mathbf{U}$ and $\mathbf{V}$ be unitary matrix representations of the Lagrangian planes, with

$$\mathbf{U} = \begin{pmatrix}\mathbf{U}_1\\ \mathbf{U}_2\end{pmatrix} \quad\text{and}\quad \mathbf{V} = \begin{pmatrix}\mathbf{V}_1\\ \mathbf{V}_2\end{pmatrix}\,.$$

By considering $\mathbf{I},0,(\mathbf{U}_1+\mathrm{i}\mathbf{U}_2)^{-1}(\mathbf{V}_1+\mathrm{i}\mathbf{V}_2),\tau-\kappa$, instead of $\mathbf{U}_1+\mathrm{i}\mathbf{U}_2,\kappa,\mathbf{V}_1+\mathrm{i}\mathbf{V}_2,\tau$, it is possible to assume that $\mathbf{U}_1+\mathrm{i}\mathbf{U}_2=\mathbf{I}$, and $\kappa=0$.

Use a homotopy to reduce the calculation of $\mathbf{O}_1$ and g to a simpler problem. Since $d=1$, $\mathbf{V}$ has exactly one eigenvalue at ± 1 (since ψ has an eigenvalue $+1$ and ψ is quadratic in $\mathbf{V}$). Since $\mathbf{V}_1+\mathrm{i}\mathbf{V}_2$ is unitary, it can be put into the following form:

$$\mathbf{V}_1+\mathrm{i}\mathbf{V}_2=\mathbf{T}^H\begin{pmatrix}(-1)^{r_2} & 0\\ 0 & \mathrm{i}(-1)^{r_4}e^{\mathrm{i}\alpha}\end{pmatrix}\mathbf{T}$$

where $\mathbf{T}$ is also a unitary matrix and $\alpha\in(-\frac{\pi}{2},\frac{\pi}{2})$. Since the unitary group $U(2)$ is connected, there exists $\mathbf{T}:[0,1]\to U(2)$ such that $\mathbf{T}(0)=\mathbf{T}$ and $\mathbf{T}(1)=\mathbf{I}$. Hence, the following path has the desired properties:

$$(\mathbf{V}_1+\mathrm{i}\mathbf{V}_2)(t)=\mathbf{T}(t)^H\begin{pmatrix}(-1)^{r_2} & 0\\ 0 & \mathrm{i}(-1)^{r_4}\mathrm{e}^{\mathrm{i}\alpha(1-t)}\end{pmatrix}\mathbf{T}(t).$$

Now $(\mathbf{V}_1+\mathrm{i}\mathbf{V}_2)(t)$ has exactly one eigenvalue at ± 1,

$$(\mathbf{V}_1+\mathrm{i}\mathbf{V}_2)(0)=\mathbf{V}_1+\mathrm{i}\mathbf{V}_2\quad\text{and}\quad(\mathbf{V}_1+\mathrm{i}\mathbf{V}_2)(1)=\begin{pmatrix}(-1)^{r_2} & 0\\ 0 & \mathrm{i}(-1)^{r_4}\end{pmatrix},$$

and τ is such that $\mathrm{e}^{-\frac{i}{2}\tau(t)}=\det(\mathbf{V}_1(t)+\mathrm{i}\mathbf{V}_2(t))$. We have $\tau(1)\in 2\pi(r_2+r_4)-\pi+4\pi\mathbb{Z}$ and the eigenvalues of ψ at $t=1$ are

$$\sigma(\psi(\mathbf{U},\mathbf{V}(1)))=\sigma\big((\mathbf{V}_1(1)-\mathrm{i}\mathbf{V}_2(1))(\mathbf{V}_1(1)+\mathrm{i}\mathbf{V}_2(1))^{-1}\big)=\big\{1,\mathrm{e}^{\mathrm{i}\pi}\big\}.$$

Let $\mathbf{V}(t)^\wedge$ be the $2-$form corresponding to $\mathbf{V}(t)$. The functions $t\mapsto\mathbf{O}_1(\mathbf{U}^\wedge,\mathbf{V}(t)^\wedge)$ and $t\mapsto g(\mathbf{U},\mathbf{V}(t))$ are continuous with values in $\{\pm1\}$ and are therefore constant. In particular, $\mathbf{O}_1(\mathbf{U},\mathbf{V})=\mathbf{O}_1(\mathbf{U},\mathbf{V}(1)^\wedge)$ and $g(\mathbf{U},\mathbf{V})=g(\mathbf{U},\mathbf{V}(1)^\wedge)$. Therefore, the proof of the lemma is complete if we can prove that $g=1$ at $t=1$. Now, in the exterior algebra representation $\mathbf{U}^\wedge=\mathbf{e}_1\wedge\mathbf{e}_2$ and $\mathbf{V}(1)^\wedge=(-1)^{r_2}\mathbf{e}_1\wedge(-1)^{r_4}\mathbf{e}_4$. Therefore, the intersection index, for the case $d=1$ at $t=1$, is

$$\begin{aligned}\mathbf{O}_1(\mathbf{U}^\wedge,\mathbf{V}^\wedge)&=\mathbf{O}_1(\mathbf{U}^\wedge,\mathbf{V}(1)^\wedge)\\&=(-1)^{r_2+r_4}\,\mathrm{sign}(\boldsymbol{\Omega}^{\mathrm{dual}}\wedge\mathbf{e}_2\wedge\mathbf{e}_4)\\&=(-1)^{r_2+r_4}\,\mathrm{sign}(\mathbf{e}_3\wedge\mathbf{e}_1\wedge\mathbf{e}_2\wedge\mathbf{e}_4)\\&=(-1)^{r_2+r_4},\end{aligned}\tag{10.1}$$

and

$$\begin{aligned}g(\mathbf{U},\mathbf{V})&=g(\mathbf{U},\mathbf{V}(1))\\&=(-1)^{m+\frac{1}{2}}\mathbf{O}_1(\mathbf{U}^\wedge,\mathbf{V}(1)^\wedge)\\&=(-1)^{m+\frac{1}{2}}(-1)^{r_2+r_4}\\&=(-1)^{r_4+r_2-\frac{1}{2}+\frac{1}{2}}(-1)^{r_2+r_4}\\&=1,\end{aligned}$$

using the fact that $\kappa = \alpha_1 = 0$ and $\alpha_2 = \pi$ and so

$$\begin{aligned} m((\mathbf{U},\kappa),(\mathbf{V},\tau)) &= m((\mathbf{I},0),(\mathbf{V}(1),\tau(1))) \\ &= \tfrac{\tau(1)}{2\pi} - \tfrac{\pi}{2\pi} + \tfrac{1}{2} \\ &= \tfrac{\tau(1)}{2\pi} \\ &= r_2 + r_4 - \tfrac{1}{2}\,. \end{aligned}$$

This completes the proof in the case $d = 1$.

Now consider the case $d = 0$. Define the function

$$g(\mathbf{U},\mathbf{V}) = (-1)^m\, \mathbf{O}_0(\mathbf{U}^\wedge, \mathbf{V}^\wedge)\,,$$

where κ, τ are taken such that $(\mathbf{U},\kappa),(\mathbf{V},\tau) \in \widetilde{U(2)}$.

By the same argument as in case $d = 1$, one can assume without generality loss that $\mathbf{U}_1 + \mathrm{i}\mathbf{U}_2 = \mathbf{I}$ and $\kappa = 0$.

In this case $\mathbf{V}_1 + \mathrm{i}\mathbf{V}_2$ doesn't have any eigenvalues at ± 1. Hence there exists a continuous path $\mathbf{V}_1 + \mathrm{i}\mathbf{V}_2 : [0,1] \to U(2)$ such that $\mathbf{V}_1(t) + \mathrm{i}\mathbf{V}_2(t)$ has no eigenvalues at ± 1, with

$$(\mathbf{V}_1 + \mathrm{i}\mathbf{V}_2)(0) = \mathbf{V}_1 + \mathrm{i}\mathbf{V}_2 \quad \text{and} \quad (\mathbf{V}_1 + \mathrm{i}\mathbf{V}_2)(1) = \begin{pmatrix} \mathrm{i}(-1)^k & 0 \\ 0 & \mathrm{i}(-1)^l \end{pmatrix}\,,$$

and τ is such that $e^{-\frac{i}{2}\tau(t)} = \det(\mathbf{V}_1 + \mathrm{i}\mathbf{V}_2(t))$. We have $\tau(1) \in 2\pi + 2(k+l)\pi + 4\pi\mathbb{Z}$ and the eigenvalues of ψ are

$$\sigma\big(\psi(\mathbf{U},\mathbf{V})\big) = \sigma\big((\mathbf{V}_1(1) - \mathrm{i}\mathbf{V}_2(1))(\mathbf{V}_1(1) + \mathrm{i}\mathbf{V}_2(1))^{-1}\big) = \big\{\mathrm{e}^{\mathrm{i}\pi}, \mathrm{e}^{\mathrm{i}\pi}\big\}\,.$$

Let $\mathbf{V}(t)^\wedge$ be the 2-form corresponding to $\mathbf{V}(t)$. As in the case $d = 1$, the functions $t \mapsto \mathbf{O}_0(\mathbf{U}^\wedge, \mathbf{V}(t)^\wedge)$ and $t \mapsto g(\mathbf{U}, \mathbf{V}(t))$ are continuous with values in $\{\pm 1\}$ and are therefore constant. In particular, $\mathbf{O}_0(\mathbf{U}^\wedge, \mathbf{V}^\wedge) = \mathbf{O}_0(\mathbf{U}^\wedge, \mathbf{V}(1)^\wedge)$ and $g(\mathbf{U},\mathbf{V}) = g(\mathbf{U},\mathbf{V}(1))$.

Since $\mathbf{U} = \mathbf{I}$, the exterior algebra representation is $\mathbf{U}^\wedge = \mathbf{e}_1 \wedge \mathbf{e}_2$; similarly $\mathbf{V}(1)^\wedge = (-1)^k\mathbf{e}_3 \wedge (-1)^l\mathbf{e}_4$. Therefore, the intersection index is

$$\begin{aligned} \mathbf{O}_0(\mathbf{U},\mathbf{V}) &= \mathbf{O}_0(\mathbf{U}^\wedge, \mathbf{V}(1)^\wedge) \\ &= \operatorname{sign}(\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge (-1)^k\mathbf{e}_3 \wedge (-1)^l\mathbf{e}_4) \\ &= (-1)^{k+l}\,, \end{aligned}$$

and

$$\begin{aligned} g(\mathbf{U},\mathbf{V}) &:= g(\mathbf{U},\mathbf{V}(1)) \\ &= (-1)^m\mathbf{O}_0(\mathbf{U}^\wedge, \mathbf{V}(1)^\wedge) \\ &= (-1)^m(-1)^{k+l} \\ &= (-1)^{k+l}(-1)^{k+l} \\ &= 1\,, \end{aligned}$$

since $m = k + l$ in this case. □

10.1 The Lazutkin invariant and the Maslov index

The Maslov index, in the Souriau representation, for a homoclinic orbit is defined in (9.3). Use Lemma 10.2 above to conclude the proof of Theorem 10.1,

$$\begin{aligned}(-1)^{I_{hom}(\widehat{\mathbf{u}})+\frac{1}{2}} &= \mathbf{O}_1\big((\widehat{\mathbf{u}}_x \wedge \mathbf{a}^-(x)), (\widehat{\mathbf{u}}_x \wedge \mathbf{a}^+(x))\big) \\ &\qquad - \lim_{y\to+\infty} \mathbf{O}_1\big((\widehat{\mathbf{u}}_x \wedge \mathbf{a}^-(-y)), (\widehat{\mathbf{u}}_x \wedge \mathbf{a}^+(y)\big)\,.\end{aligned} \tag{10.1}$$

But the right-hand side is just the sign of $\mathcal{T}(\widehat{\mathbf{u}})$. Hence, we conclude that

$$(-1)^{I_{hom}(\widehat{\mathbf{u}})+\frac{1}{2}} = \operatorname{sign}\big(\mathcal{T}(\widehat{\mathbf{u}})\big)\,.$$

□

11 Concluding remarks

The generalization of this theory to spaces with $\mathbb{V}$ of dimension greater than 4 is given in [4]. The key feature needed is a generalization of the Lazutkin invariant and such a generalization has been given by Treshchev [26]. An additional new direction that is now possible is a new proof of the sufficient condition for instability in [5, 6] in the case where the PDE is multi-symplectic (the matrix $\mathbf{M}$ is skew-symmetric). The proof in [5, 6] required additional symmetry and that hypothesis can now be replaced by the property of Lagrangian subspaces.

Acknowledgements

This paper emerged from open questions in [11] and began to take form at a MIRaW Day on the Maslov index at the University of Warwick in November 2009, organized by Chris Jones and Frédéric Dias. Helpful comments from the participants are gratefully acknowledged. The authors are particularly grateful to Vassili Gelfreich for helpful discussions on the Lazutkin invariant.

— Appendix —

A Symplectifying four-vectors

The purpose of this appendix is to prove the following

Proposition A.1 *Let* $\mathbf{a}$*,* $\mathbf{b}$*,* $\mathbf{c}$ *and* $\mathbf{d}$ *be any four vectors in* $\mathbb{V}$*. Then*

$$\mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d} = \det\begin{bmatrix}\mathbf{\Omega}(\mathbf{a},\mathbf{c}) & \mathbf{\Omega}(\mathbf{a},\mathbf{d})\\ \mathbf{\Omega}(\mathbf{b},\mathbf{c}) & \mathbf{\Omega}(\mathbf{b},\mathbf{d})\end{bmatrix} \mathit{vol} - \mathbf{\Omega}(\mathbf{a},\mathbf{b})\mathbf{\Omega}(\mathbf{c},\mathbf{d})\mathit{vol}\,. \qquad \text{(A-1)}$$

Corollary A.2 *If either* $\mathrm{span}\{\mathbf{a},\mathbf{b}\}$ *or* $\mathrm{span}\{\mathbf{c},\mathbf{d}\}$ *is a Lagrangian subspace. Then the formula reduces to*

$$\mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d} = \det\begin{bmatrix}\mathbf{\Omega}(\mathbf{a},\mathbf{c}) & \mathbf{\Omega}(\mathbf{a},\mathbf{d})\\ \mathbf{\Omega}(\mathbf{b},\mathbf{c}) & \mathbf{\Omega}(\mathbf{b},\mathbf{d})\end{bmatrix} \mathit{vol}\,. \qquad \text{(A-2)}$$

Proof The corollary follows from the proposition using the definition of Lagrangian subspace. For the proposition, if the vectors are dependent then the left and right hand sides vanish. So restrict to the case where the vectors are linearly independent.

The four-form is proportional to the volume form,

$$\mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d} = [\![\mathsf{vol}^*, \mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d}]\!]_4\,\mathsf{vol}\,.$$

The key then is to use the symplectic form representation of the volume form (4.3)

$$\begin{aligned}\mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d} &= -\tfrac{1}{2}[\![\mathbf{\Omega}\wedge\mathbf{\Omega}, \mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d}]\!]_4\,\mathsf{vol}\\ &= -\tfrac{1}{2}[\![(\mathbf{a}\wedge\mathbf{b}) \lrcorner (\mathbf{\Omega}\wedge\mathbf{\Omega}), \mathbf{c}\wedge\mathbf{d}]\!]_2\,.\end{aligned}$$

We will need the following two formulas

$$\mathbf{a}\lrcorner\big(\mathbf{B}\wedge\mathbf{C}\big) = (\mathbf{a}\lrcorner\mathbf{B})\wedge\mathbf{C} + (-1)^k\mathbf{B}\wedge(\mathbf{a}\lrcorner\mathbf{C}) \quad \text{and} \quad \mathbf{B}\wedge\mathbf{C} = (-1)^{k\ell}\mathbf{C}\wedge\mathbf{B}\,, \qquad \text{(A-3)}$$

for any $\mathbf{a}\in\mathbb{V}$, $\mathbf{B}\in\bigwedge^k(\mathbb{V}^*)$, and $\mathbf{C}\in\bigwedge^\ell(\mathbb{V}^*)$, which are proved in [15]. From these formulas it follows that

$$\mathbf{a}\lrcorner(\mathbf{\Omega}\wedge\mathbf{\Omega}) = 2(\mathbf{a}\lrcorner\mathbf{\Omega})\wedge\mathbf{\Omega}\,,$$

and

$$\begin{aligned}(\mathbf{a}\wedge\mathbf{b})\lrcorner(\mathbf{\Omega}\wedge\mathbf{\Omega}) &= \mathbf{b}\lrcorner\big(\mathbf{a}\lrcorner(\mathbf{\Omega}\wedge\mathbf{\Omega})\big)\\ &= 2\mathbf{\Omega}(\mathbf{a},\mathbf{b})\,\mathbf{\Omega} - 2(\mathbf{a}\lrcorner)\mathbf{\Omega}\wedge(\mathbf{b}\lrcorner\mathbf{\Omega})\\ &= 2\mathbf{\Omega}(\mathbf{a},\mathbf{b})\,\mathbf{\Omega} - 2\mathbf{J}\mathbf{a}\wedge\mathbf{J}\mathbf{b}\,,\end{aligned}$$

using $\mathbf{a} \lrcorner \boldsymbol{\Omega} = \mathbf{J}\mathbf{a}$, and so

$$(\mathbf{a}\wedge\mathbf{b}) \lrcorner \mathsf{vol}^* = -\tfrac{1}{2}(\mathbf{a}\wedge\mathbf{b}) \lrcorner (\boldsymbol{\Omega}\wedge\boldsymbol{\Omega}) = \mathbf{J}\mathbf{a}\wedge\mathbf{J}\mathbf{b} - \boldsymbol{\Omega}(\mathbf{a},\mathbf{b})\,\boldsymbol{\Omega}\,.$$

Pair this equation with $\mathbf{c}\wedge\mathbf{d}$

$$[\![(\mathbf{a}\wedge\mathbf{b}) \lrcorner \mathsf{vol}^*, \mathbf{c}\wedge\mathbf{d}]\!]_2 = [\![\mathbf{J}\mathbf{a}\wedge\mathbf{J}\mathbf{b} - \boldsymbol{\Omega}(\mathbf{a},\mathbf{b})\boldsymbol{\Omega}, \mathbf{c}\wedge\mathbf{d}]\!]_2\,.$$

or

$$[\![\mathsf{vol}^*, \mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d}]\!]_4 = [\![\mathbf{J}\mathbf{a}\wedge\mathbf{J}\mathbf{b}, \mathbf{c}\wedge\mathbf{d}]\!]_2 - \boldsymbol{\Omega}(\mathbf{a},\mathbf{b})\boldsymbol{\Omega}(\mathbf{c},\mathbf{d})\,,$$

the proof of the formula (A-1) is complete by noting that

$$[\![\mathbf{J}\mathbf{a}\wedge\mathbf{J}\mathbf{b}, \mathbf{c}\wedge\mathbf{d}]\!]_2 = \det\begin{bmatrix}\boldsymbol{\Omega}(\mathbf{a},\mathbf{c}) & \boldsymbol{\Omega}(\mathbf{a},\mathbf{d})\\ \boldsymbol{\Omega}(\mathbf{b},\mathbf{c}) & \boldsymbol{\Omega}(\mathbf{b},\mathbf{d})\end{bmatrix}\,.$$

□

B Lagrangian planes

In this appendix the discussion of Lagrangian planes in §4 is expanded. In the paper, three other representations of Lagrangian planes are needed. The second representation is in terms of exterior algebra

$$\mathrm{span}\{\mathbf{a}\wedge\mathbf{b}\}\quad\text{with}\quad [\![\boldsymbol{\Omega}, \mathbf{a}\wedge\mathbf{b}]\!]_2 = 0\,, \tag{B-1}$$

equivalently,

$$\boldsymbol{\Omega}^{\mathrm{dual}}\wedge\mathbf{a}\wedge\mathbf{b} = 0\,.$$

The third representation is in terms of 4×2 matrices of rank 2,

$$\mathbf{Z} = \begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}\quad\text{with}\quad \mathbf{Y}^T\mathbf{X} = \mathbf{X}^T\mathbf{Y}\,, \tag{B-2}$$

where $\mathbf{X}, \mathbf{Y}$ are 2×2 matrices. This representation is called a *Lagrangian frame.*

The fourth representation is in terms of 2×2 unitary matrices, $U(2)$. The identity (B-2) implies that

$$(\mathbf{X} - \mathrm{i}\mathbf{Y})^T(\mathbf{X} + \mathrm{i}\mathbf{Y}) = \mathbf{X}^T\mathbf{X} + \mathbf{Y}^T\mathbf{Y}\,.$$

Hence if $\mathbf{R} = \sqrt{\mathbf{X}^T\mathbf{X} + \mathbf{Y}^T\mathbf{Y}}$, which is well defined and symmetric since the argument is positive definite, then

$$\mathbf{Q} = \mathbf{Q}_1 + \mathrm{i}\mathbf{Q}_2 := (\mathbf{X} + \mathrm{i}\mathbf{Y})\mathbf{R}^{-1}\,,$$

is unitary. This provides a unitary representation of a Lagrangian plane. The determinant of a unitary matrix lies on the unit circle. Hence this representation can be used to define the *Maslov angle* of a Lagrangian plane

$$\mathrm{e}^{-\mathrm{i}\frac{\kappa}{2}} = \det[\mathbf{Q}_1 + \mathrm{i}\mathbf{Q}_2]\,, \tag{B-3}$$

and the polar representation $\det(\mathbf{X} + \mathrm{i}\mathbf{Y}) = \det(\mathbf{R})\mathrm{e}^{-\mathrm{i}\frac{\kappa}{2}}$.

C Normalizations

There are three normalizations used in the paper, and they are summarized in this appendix. The first is the normalization of the eigenvectors $\xi_1, \ldots, \xi_4$. The second is normalization of the tangent vectors $\mathbf{a}^\pm$ in order to impose the property (5.5), and the third is the normalization (7.10).

C.1 Normalization of the eigenvectors

Eigenvectors are usually normalized by using the adjoint eigenvectors. Here the eigenvectors are normalized by imposing symplecticity.

There are three qualitative configurations of the eigenvalues in the hyperbolic case as shown in Figure 1. In the case of real and distinct eigenvalues ξ_1 and ξ_2 are the eigenvectors corresponding to μ_1 and μ_2. In the case $\mu_1 = \mu_2$, ξ_1 is the eigenvector and ξ_2 is the generalized eigenvector. In the case where μ_1 and μ_2 are complex with $\mu_2 = \overline{\mu_1}$, then ξ_1 and ξ_2 are the real and imaginary parts of the complex eigenvector associated with μ_1. In all three cases $\text{span}\{\xi_1, \xi_2\}$ is a Lagrangian subspace. A similar formulation applies to ξ_3 and ξ_4.

Now, define

$$\mathbf{K} = \begin{bmatrix} \mathbf{\Omega}(\xi_1, \xi_3) & \mathbf{\Omega}(\xi_1, \xi_4) \\ \mathbf{\Omega}(\xi_2, \xi_3) & \mathbf{\Omega}(\xi_2, \xi_4) \end{bmatrix} .$$

This matrix is invertible since the eigenvectors are linearly independent and from the formula (1.8)

$$\xi_1 \wedge \xi_2 \wedge \xi_3 \wedge \xi_4 = \det\big[\mathbf{K}\big] \, \mathsf{vol} \, .$$

Let

$$\mathbf{\Phi} = \big[\xi_1 \,\big|\, \xi_2 \,\big|\, \xi_3 \,\big|\, \xi_4\big] \, ,$$

then

$$\mathbf{\Phi}^T \mathbf{J} \mathbf{\Phi} = \begin{bmatrix} \mathbf{0} & -\mathbf{K} \\ \mathbf{K} & \mathbf{0} \end{bmatrix} .$$

Therefore take the equivalent subspace for $\text{span}\{\xi_3, \xi_4\}$,

$$\big[\xi_3 | \xi_4\big] = \big[\widehat{\xi}_3 | \widehat{\xi}_4\big] \mathbf{K}^T \, , \tag{C-1}$$

and redefine $\mathbf{\Phi}$ to

$$\widehat{\mathbf{\Phi}} = \Big[\xi_1 \,\Big|\, \xi_2 \,\Big|\, \widehat{\xi}_3 \,\Big|\, \widehat{\xi}_4\Big] \, .$$

Then the transformation $\widehat{\mathbf{\Phi}}$ is symplectic

$$\widehat{\mathbf{\Phi}}^T \mathbf{J} \widehat{\mathbf{\Phi}} = \mathbf{J} \, ,$$

and

$$\xi_1 \wedge \xi_2 \wedge \widehat{\xi}_3 \wedge \widehat{\xi}_4 = \mathsf{vol} \, .$$

C.2 Normalizing the tangent vectors $\mathbf{a}^{\pm}$

In the linearization with $\lambda = 0$, the stable subspace is $\text{span}\{\widehat{\mathbf{u}}_x, \mathbf{a}^+\}$ and the unstable subspace is $\text{span}\{\widehat{\mathbf{u}}_x, \mathbf{a}^-\}$, and they satisfy

$$\text{span}\{\widehat{\mathbf{u}}_x, \mathbf{a}^+\} \to \text{span}\{\xi_1, \xi_2\} \quad \text{as} \quad x \to +\infty\,,$$

and

$$\text{span}\{\widehat{\mathbf{u}}_x, \mathbf{a}^-\} \to \text{span}\{\xi_3, \xi_4\} \quad \text{as} \quad x \to -\infty\,.$$

Hence there exists constants $C^{\pm}$ with the property that

$$\mathrm{e}^{-(\mu_1+\mu_2)x}\, \widehat{\mathbf{u}}_x(x) \wedge \mathbf{a}^+(x) \to C^+\xi_1 \wedge \xi_2 \quad \text{as} \quad x \to +\infty\,, \tag{C-2}$$

and

$$\mathrm{e}^{-(\mu_3+\mu_4)x}\, \widehat{\mathbf{u}}_x(x) \wedge \mathbf{a}^-(x) \to C^-\xi_3 \wedge \xi_4 \quad \text{as} \quad x \to -\infty\,.$$

Now using the fact that $\mu_3 + \mu_4 = -\mu_1 - \mu_2$ and replacing ξ_3, ξ_4 by their scaled versions (C-1), the latter expression can be replaced by

$$\mathrm{e}^{-(\mu_1+\mu_2)x}\, \widehat{\mathbf{u}}_x(-x) \wedge \mathbf{a}^-(-x) \to \widehat{C^-}\widehat{\xi}_3 \wedge \widehat{\xi}_4 \quad \text{as} \quad x \to +\infty\,. \tag{C-3}$$

Combining (C-2) and (C-3) then gives

$$\mathrm{e}^{-2(\mu_1+\mu_2)x}\widehat{\mathbf{u}}_x(x) \wedge \mathbf{a}^+(x) \wedge \widehat{\mathbf{u}}_x(-x) \wedge \mathbf{a}^-(-x) \to C^+C^-\, \xi_1 \wedge \xi_2 \wedge \xi_3 \wedge \xi_4\,,$$

or

$$\mathrm{e}^{-2(\mu_1+\mu_2)x}\widehat{\mathbf{u}}_x(x) \wedge \mathbf{a}^+(x) \wedge \widehat{\mathbf{u}}_x(-x) \wedge \mathbf{a}^-(-x) \to C^+\widehat{C^-}\, \mathsf{vol}\,.$$

Now it is clear how to scale $\mathbf{a}^{\pm}(x)$: define

$$\widehat{\mathbf{a}}^+(x) := \frac{1}{C^+}\mathbf{a}^+(x) \quad \text{and} \quad \widehat{\mathbf{a}}^-(x) := \frac{1}{\widehat{C^-}}\mathbf{a}^-(x)\,.$$

Then

$$\lim_{x\to+\infty} \mathrm{e}^{-2(\mu_1+\mu_2)x}\widehat{\mathbf{u}}_x(x) \wedge \widehat{\mathbf{a}}^+(x) \wedge \widehat{\mathbf{u}}_x(-x) \wedge \widehat{\mathbf{a}}^-(-x) = \mathsf{vol}\,, \tag{C-4}$$

which is what is required in (5.5).

C.3 The Evans function in the limit $\lambda \to 0$

Define 4×2 matrices

$$\mathbf{U}^+(x,\lambda) = \left[\mathbf{u}_1^+(x,\lambda) \,\middle|\, \mathbf{u}_2^+(x,\lambda)\right] \quad \text{and} \quad \mathbf{U}^-(x,\lambda) = \left[\mathbf{u}_3^-(x,\lambda) \,\middle|\, \mathbf{u}_4^-(x,\lambda)\right].$$

Define

$$\Lambda^+(x,\lambda) = \begin{bmatrix} \mathrm{e}^{\mu_1(\lambda)x} & 0 \\ 0 & \mathrm{e}^{\mu_2(\lambda)x} \end{bmatrix} \quad \text{and} \quad \Lambda^-(x,\lambda) = \begin{bmatrix} \mathrm{e}^{\mu_3(\lambda)x} & 0 \\ 0 & \mathrm{e}^{\mu_4(\lambda)x} \end{bmatrix}.$$

Then, the asymptotic x limits with λ fixed are

$$\lim_{x\to+\infty} \mathbf{U}^+(x,\lambda)\Lambda^+(x,\lambda)^{-1} = \Xi^+(\lambda) \quad \text{and} \quad \lim_{x\to-\infty} \mathbf{U}^-(x,\lambda)\Lambda^-(x,\lambda)^{-1} = \Xi^-(\lambda)\,,$$

where

$$\Xi^+(\lambda) = \left[\xi_1(\lambda) \,\middle|\, \xi_2(\lambda)\right] \quad \text{and} \quad \Xi^-(\lambda) = \left[\xi_3(\lambda) \,\middle|\, \xi_4(\lambda)\right].$$

The limit $\lambda \to 0$ with x fixed gives

$$\mathbf{U}^\pm(x,0) = \left[\widehat{\mathbf{u}}_x|\mathbf{a}^\pm\right]\mathbf{c}^\pm\,,$$

for some invertible 2×2 matrices

$$\mathbf{c}^\pm = \begin{bmatrix} c_{11}^\pm & c_{12}^\pm \\ c_{21}^\pm & c_{22}^\pm \end{bmatrix}.$$

Let $\mathbf{T}^\pm = {\mathbf{c}^\pm}^{-1}$ and define

$$\widetilde{\mathbf{U}}^+(x,\lambda) = \left[\widetilde{\mathbf{u}}_1^+(x,\lambda) \,\middle|\, \widetilde{\mathbf{u}}_2^+(x,\lambda)\right] = \left[\mathbf{u}_1^+(x,\lambda) \,\middle|\, \mathbf{u}_2^+(x,\lambda)\right]\mathbf{T}^+\,,$$

or

$$\mathbf{U}^+(x,\lambda) = \widetilde{\mathbf{U}}^+(x,\lambda)(\mathbf{T}^+)^{-1}\,,$$

with a similar definition for $\widetilde{\mathbf{U}}^-(x,\lambda)$.

Now, the basis vectors for the stable subspace satisfy

$$\lim_{x\to+\infty} \mathbf{U}^+(x,\lambda)\Lambda^+(x,\lambda)^{-1} = \Xi^+(\lambda)\,,$$

equivalently

$$\lim_{x\to+\infty} \widetilde{\mathbf{U}}^+(x,\lambda)\Lambda^+(x,\lambda)^{-1}\mathbf{T}^+ = \Xi^+(\lambda)\mathbf{T}^+\,.$$

Now define

$$\widetilde{\Lambda}^+(x,\lambda) = {\mathbf{T}^+}^{-1}\Lambda^+(x,\lambda)\mathbf{T}^+ \quad \text{and} \quad \widetilde{\Xi}^+(\lambda) = \Xi^+(\lambda)\mathbf{T}^+\,.$$

Then $\mathbf{T}^+\widetilde{\Lambda}^+(x,\lambda)^{-1} = \Lambda^+(x,\lambda)^{-1}\mathbf{T}^+$ and

$$\lim_{x\to+\infty} \mathbf{U}^+(x,\lambda)\mathbf{T}^+\widetilde{\Lambda}^+(x,\lambda)^{-1} = \widetilde{\Xi}^+(\lambda)\,,$$

or

$$\lim_{x\to+\infty} \widetilde{\mathbf{U}}^+(x,\lambda)\widetilde{\Lambda}^+(x,\lambda)^{-1} = \widetilde{\Xi}^+(\lambda)\,,$$

and

$$\left[\widetilde{\mathbf{u}}_1^+(x,0) \,\middle|\, \widetilde{\mathbf{u}}_2^+(x,0)\right] = \left[\widehat{\mathbf{u}}_x \,\middle|\, \widehat{\mathbf{a}}^+\right]. \tag{C-5}$$

Similarly

$$\lim_{x\to-\infty} \widetilde{\mathbf{U}}^-(x,\lambda)\widetilde{\Lambda}^-(x,\lambda)^{-1} = \widetilde{\Xi}^-(\lambda)\,,$$

and

$$\left[\widetilde{\mathbf{u}}_3^-(x,0) \,\middle|\, \widetilde{\mathbf{u}}_4^-(x,0)\right] = \left[\widehat{\mathbf{u}}_x \,\middle|\, \widehat{\mathbf{a}}^-\right]. \tag{C-6}$$

Hence the normalized vectors now satisfy the limits (7.10).

D $\lambda-$derivative of $\mathbf{\Omega}(\mathbf{u}_3^-, \mathbf{u}_1^+)$

The purpose of this appendix is to prove Proposition 7.1. $\lambda-$derivatives are widely used in Evans function analysis so we can be brief. Derivatives of the form used here were first given by PEGO & WEINSTEIN [22], with symplectic versions presented in BRIDGES & DERKS [5, 6].

Differentiating (7.1) with respect to λ

$$\mathbf{J}\left(\frac{\partial \mathbf{u}}{\partial \lambda}\right)_x = [\mathbf{B}(x,p) - \lambda \mathbf{M}]\left(\frac{\partial \mathbf{u}}{\partial \lambda}\right) - \mathbf{M}\mathbf{u}\,.$$

Now,

$$\begin{aligned}
\tfrac{\partial}{\partial x}\mathbf{\Omega}\left(\tfrac{\partial}{\partial \lambda}\mathbf{u}_3^-, \mathbf{u}_1^+\right) &= \mathbf{\Omega}\big((\mathbf{u}_3^-)_{x\lambda}, \mathbf{u}_1^+\big) + \mathbf{\Omega}\big((\mathbf{u}_3^-)_\lambda, (\mathbf{u}_1^+)_x\big) \\
&= \langle \mathbf{J}(\mathbf{u}_3^-)_{\lambda x}, \mathbf{u}_1^+\rangle - \langle (\mathbf{u}_3^-)_\lambda, \mathbf{J}(\mathbf{u}_1^+)_x\rangle \\
&= \langle [\mathbf{B} - \lambda\mathbf{M}](\mathbf{u}_3^-)_\lambda - \mathbf{M}\mathbf{u}_3^-, \mathbf{u}_1^+\rangle - \langle [\mathbf{B} - \lambda\mathbf{M}](\mathbf{u}_3^-)_\lambda, \mathbf{u}_1^+\rangle \\
&= -\langle \mathbf{M}\mathbf{u}_3^-, \mathbf{u}_1^+\rangle\,.
\end{aligned}$$

Similarly

$$\frac{\partial}{\partial x}\mathbf{\Omega}\left(\mathbf{u}_3^-, \frac{\partial}{\partial \lambda}\mathbf{u}_1^+\right) = +\langle \mathbf{M}\mathbf{u}_3^-, \mathbf{u}_1^+\rangle\,.$$

Integrating over x

$$\mathbf{\Omega}((\mathbf{u}_3^-)_\lambda, \mathbf{u}_1^+)\Big|_{-R}^{x_0} = -\int_{-R}^{x_0}\langle \mathbf{M}\mathbf{u}_3^-, \mathbf{u}_1^+\rangle \quad \text{and} \quad -\mathbf{\Omega}(\mathbf{u}_3^-, (\mathbf{u}_1^+)_\lambda)\Big|_{x_0}^{S} = -\int_{x_0}^{S}\langle \mathbf{M}\mathbf{u}_3^-, \mathbf{u}_1^+\rangle\,.$$

Adding and using the $x-$asymptotics of $\mathbf{u}_1^+, \mathbf{u}_3^-$ as $R, S \to \infty$ gives

$$\big[\mathbf{\Omega}((\mathbf{u}_3^-)_\lambda, \mathbf{u}_1^+) + \mathbf{\Omega}(\mathbf{u}_3^-, (\mathbf{u}_1^+)_\lambda)\big]\Big|_{x=x_0} = -\int_{-\infty}^{+\infty}\langle \mathbf{M}\mathbf{u}_3^-, \mathbf{u}_1^+\rangle\,\mathrm{d}x\,.$$

Evaluating at $\lambda = 0$, using the limits (7.10), and noting that the choice $x = x_0$ is arbitrary, then gives

$$\partial_\lambda \mathbf{\Omega}(\mathbf{u}_3^-, \mathbf{u}_1^+)\Big|_{\lambda=0} = -\int_{-\infty}^{+\infty}\langle \mathbf{M}\widehat{\mathbf{u}}_x, \widehat{\mathbf{u}}_x\rangle\,\mathrm{d}x\,,$$

completing the proof of Proposition 7.1. □

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] A. Bose & C.K.R.T. Jones. *Stability of the in-phase travelling wave solution in a pair of coupled nerve fibers*, Indiana Univ. Math. J. **44**, 189–220 (1995).

[4] T.J. Bridges & F. Chardard. *Symplectic transversality of homoclinic orbits in dimension $2n$, $n > 2$, and its implications*, Draft manuscript.

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

[6] 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).

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

[8] B. Buffoni, A.R. Champneys & J.F. Toland. *Bifurcation and coalescence of a plethora of multi-modal homoclinic orbits in a Hamiltonian system*, J. Dyn. Diff. Eqns. **8** 221–281 (1996).

[9] F. Chardard. *Maslov index for solitary waves obtained as a limit of the Maslov index for periodic waves*, C. R. Acad. Sci. Paris, Ser. I **345** 689–694 (2007).

[10] F. Chardard. *Stabilité des Ondes Solitaires*, PhD Thesis, École Normale Supérieure de Cachan, France (2009).

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

[12] F. Chardard, F. Dias & T.J. Bridges. *On the Maslov index of solitary waves. Part 1: Hamiltonian systems on a four-dimensional phase space*, Physica D **238** 1841–1867 (2009).

[13] F. Chardard, F. Dias & T.J. Bridges. *On the Maslov index of solitary waves. Part 2: phase space with dimension greater than four*, Physica D **240** 1334–1344 (2011).

[14] C.-N. Chen & X. Hu. *Maslov index for homoclinic orbits of Hamiltonian systems*, Ann. Inst. H. Poincaré Anal. Non Linéaire **24** (2007), no. 4, 589–603.

[15] M. Crampin & F.A.E. Pirani. *Applicable Differential Geometry*, London Math. Soc. Lecture Notes **59**, Cambridge University Press (1986).

[16] V.G. Gelfreich. *Numerics and exponential smallness*, Chapter 6 in volume 2 of the Handbook of Dynamical Systems, 265–312 (2002).

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

[18] V.G. Gelfreich & D.K. Sharomov. *Examples of Hamiltonian systems with transversal homoclinic orbits*, Phys. Lett. A **197** 139–146 (1995).

[19] C.K.R.T. Jones. *Instability of standing waves for non-linear Schrödinger-type equations*, Ergodic Theory and Dynamical Systems **8***, 119–138 (1988).

[20] V.F. Lazutkin. *Splitting of separatrices for the Chirikov standard map*, VINITI 6372-84, 45 pages, (1984).

[21] J.R. Magnus & H. Neudecker. *Matrix Differential Calculus*, John Wiley & Sons: Chichester (1988).

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

[23] J.W. Robbin & D.A. Salamon. *The Maslov index for paths*, Topology **32** 827–844 (1993).

[24] B. Sandstede. *Instability of localized buckling modes in a one-dimensional strut model*, Phil. Trans. R. Soc. Lond. A **355** 2083–2097 (1997).

[25] J.-M. Souriau. *Construction explicite de l'indice de Maslov. Applications*, Group theoretical methods in physics (Fourth Internat. Colloq., Nijmegen, 1975), Springer Lecture Notes in Physics **50** 117–148 (1976).

[26] D.V. Treshchev. *Separatrix splitting from the point of view of symplectic geometry*, Matematicheskie Zametki **61** 890–906 (1997). (English translation: Mathematical Notes **61** 744–757 (1997)).