Under consideration for publication in *Nonlinearity*

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

Frédéric Chardard[1], **Thomas J. Bridges**[2]

[1] Département de Mathématiques, Institut Camille Jordan (UMR 5208), Université Jean Monnet, 23, rue du docteur Paul Michelon, 42023 Saint-Étienne, France

[2] Department of Mathematics, University of Surrey, Guildford GU2 7XH England

E-mail: Frederic.Chardard@univ-st-etienne.fr, T.Bridges@surrey.ac.uk

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-Treschev 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-Treschev 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 derivative of the Evans function is proportional to the Lazutkin-Treschev 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.

23 September 2014

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 argument. The matrix $\mathbf{J}$ is a symplectic operator associated with the symplectic form, denoted by $\mathbf{\Omega}$, and $\mathbf{M}$ is assumed to be symmetric and the vector space $\mathbb{V}$ is taken to be $2n-$dimensional with $n \geq 2$.

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

$$\mathbf{v}_t = \mathbf{D}\mathbf{v}_{xx} + DF(\mathbf{v})\,, \quad \mathbf{v} = (v_1, \ldots, v_n)\,, \tag{1.3}$$

where $F(\mathbf{v})$ is a given smooth function, and $\mathbf{D}$ is a diagonal positive definite matrix. The case $n = 2$ 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 to a hyperbolic equilibrium, denoted $\widehat{\mathbf{u}}(x, p)$. This homoclinic orbit has two important characteristics: a Maslov index and a Lazutkin-Treschev invariant. The latter is defined as follows. Since $\mathbb{V}$ has dimension $2n$ the stable (+) and unstable (−) subspaces, in the linearization of the steady system about the homoclinic orbit, are of the form

$$E^{s,u} = \mathrm{span}\{\widehat{\mathbf{u}}_x, \mathbf{a}_1^{\pm}, \ldots, \mathbf{a}_{n-1}^{\pm}\}\,,$$

for each x (with p fixed). The Lazutkin-Treschev invariant, *which is independent of* x, is defined by

$$\mathcal{T}(\widehat{\mathbf{u}}) = \det \begin{bmatrix} \mathbf{\Omega}(\mathbf{a}_1^-, \mathbf{a}_1^+) & \cdots & \mathbf{\Omega}(\mathbf{a}_1^-, \mathbf{a}_{n-1}^+) \\ \vdots & \ddots & \vdots \\ \mathbf{\Omega}(\mathbf{a}_{n-1}^-, \mathbf{a}_1^+) & \cdots & \mathbf{\Omega}(\mathbf{a}_{n-1}^-, \mathbf{a}_{n-1}^+) \end{bmatrix}. \tag{1.4}$$

It was discovered by Treschev [29], and generalizes the case $n = 2$ introduced by Lazutkin [25]. In the case $n = 2$, it has been a valuable tool to study the case where the distance between the stable and unstable manifolds is exponentially small (e.g. [19, 18, 20]) (the formula (1.4) in the case $n = 2$ is given explicitly in part C of §2.3 of [19] and an explicit example is given in [20]). In the case $n > 2$ it has been used to study the intersection between the stable and unstable manifolds associated with hyperbolic tori, and for perturbation of invariant manifolds [29].

In this paper three new results about this invariant are proved. Firstly, we give a new proof that a homoclinic orbit to a hyperbolic equilibrium is transversely constructed if and only if $\mathcal{T}(\widehat{\mathbf{u}}) \neq 0$. Secondly we prove that when $\mathbf{a}_1^{\pm}, \ldots, \mathbf{a}_{n-1}^{\pm}$ are suitably normalized, 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 Maslov is the Maslov index of the homoclinic orbit.

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-Treschev invariant can be interpreted as an 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 [23] 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 [8] and CHEN & HU [14]. In this paper we use a definition for the Maslov index based on a theory of SOURIAU [28]. It is equivalent to the above definitions and it can be related much more easily to the Lazutkin-Treschev 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 $\mathrm{span}\{\mathbf{u}_1^+,\dots,\mathbf{u}_n^+\}$ be the $(x,\lambda)-$dependent stable subspace, and $\mathrm{span}\{\mathbf{u}_1^-,\dots,\mathbf{u}_n^-\}$ be the $(x,\lambda)-$dependent unstable subspace in the linearization, then the Evans function is

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

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

$$D'(0) = -\mathfrak{I}(\widehat{\mathbf{u}}) \int_{-\infty}^{+\infty} \langle \mathbf{M}\widehat{\mathbf{u}}_x, \widehat{\mathbf{u}}_x\rangle \,\mathrm{d}x \,. \tag{1.7}$$

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-Treschev invariant vanishes. For the examples (1.2) and (1.3) the integral on the right-hand side of (1.7) is the $H^1(\mathbb{R})$ norm of ϕ and $\mathbf{v}$ respectively.

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 symplectic 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. Towards this end, a formula which will be used throughout is the following connection between $2n-$forms on $\mathbb{V}$ and symplectic determinants. For a pair of $n-$dimensional subspaces in $\mathbb{V}$,

$$A = \mathrm{span}\{\mathbf{a}_1,\dots,\mathbf{a}_n\} \quad \text{and} \quad B = \mathrm{span}\{\mathbf{b}_1,\dots,\mathbf{b}_n\}\,,$$

the formula connecting the exterior algebra representation and the symplectic determinant representation is

$$\mathbf{a}_1\wedge\cdots\wedge\mathbf{a}_n\wedge\mathbf{b}_1\wedge\cdots\wedge\mathbf{b}_n = \det\begin{bmatrix}\mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_1) & \cdots & \mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_n)\\ \vdots & \ddots & \vdots\\ \mathbf{\Omega}(\mathbf{a}_n,\mathbf{b}_1) & \cdots & \mathbf{\Omega}(\mathbf{a}_n,\mathbf{b}_n)\end{bmatrix}\mathsf{vol} + \Upsilon\mathsf{vol}\,, \tag{1.8}$$

where Υ is to be defined. The formula simplifies dramatically when either A or B is Lagrangian since $\Upsilon = 0$ if A or B is Lagrangian. We have not seen this formula before. A proof is given in §4.

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 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 for the intersection theory of Lagrangian planes is given in §4. Section 4 also includes a new proof of the necessary and sufficient condition for two Lagrangian planes to have a two-dimensional intersection, which is essential for understanding degeneracy of the Lazutkin-Treschev invariant. In §6 transversal construction of homoclinic orbits and its implications are presented.

The construction, symplectification and differentiation of the Evans function are presented in §8, leading to a proof of the formula (1.7), with an application in §7.

The longest proof in the paper is the proof of the connection between the Maslov index and the Lazutkin-Treschev invariant (1.5). The Souriau definition of the Maslov index is introduced in §10, and then applied to homoclinic orbits. The proof of (1.5) is then given in §11. In §12 the details of an example, with calculation of $\mathfrak{T}$ and the Maslov index, for a system with $\dim(\mathbb{V}) = 6$ is given.

The technical report [10] gives detailed calculations for the case $n = 2$.

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 matrix representation of the symplectic operator will be taken in standard form

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

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) with $n = 2$, 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

$$H(\mathbf{u},p) = \tfrac{1}{2}u_2^2 + \tfrac{1}{2}pu_4^2 - \tfrac{1}{2}u_1^2 - u_3u_4 + \tfrac{1}{3}u_1^3\,.$$

A second example is the system of reaction diffusion equations (1.3). Systems of this type are considered in Bose & Jones [3] in the case $n = 2$ and an example with $n = 3$ is considered in §12. The system (1.3) can be expressed in the form (1.1) by taking

$$\mathbf{u} = (\mathbf{v},\mathbf{p}) := (\mathbf{v},\mathbf{D}\mathbf{v}_x)\,,$$

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

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

The Hamiltonian function is

$$H(\mathbf{v},\mathbf{p}) = \frac{1}{2}\mathbf{p}\cdot\mathbf{D}^{-1}\mathbf{p} + F(\mathbf{v})\,.$$

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. 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 indicates that the eigenvalue λ in the Evans function can be taken to be real, and it affects the formula (1.7). When $\mathbf{M}$ is 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 [17, 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, with $\mathbf{M}$ having the property (2.1),

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

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 $p = p_0$,

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

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 [26]. Hence

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

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-Treschev 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

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

4. Intersection of Lagrangian planes

Here and throughout $\mathbb{V}$ is a $2n-$dimensional normed vector space. Let

$$\mathbb{V} = \mathrm{span}\big\{\mathbf{e}_1,\ldots,\mathbf{e}_{2n}\big\} \quad \text{and} \quad \mathbb{V}^* = \mathrm{span}\big\{\mathbf{e}_1^*,\ldots,\mathbf{e}_{2n}^*\big\}\,, \tag{4.1}$$

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 $k=1,\ldots,2n$. The convention here on the exterior algebra spaces follows Chapter 4 of Crampin & Pirani [16]. The induced pairing on the wedge spaces is denoted by

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

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

$$\mathbf{\Omega} = \mathbf{e}_1^*\wedge\mathbf{e}_{n+1}^* + \cdots + \mathbf{e}_n^*\wedge\mathbf{e}_{2n}^* \tag{4.2}$$

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

$$\mathsf{vol} := \mathbf{e}_1\wedge\cdots\wedge\mathbf{e}_{2n} \quad \text{and} \quad \mathsf{vol}^* := \mathbf{e}_1^*\wedge\cdots\wedge\mathbf{e}_{2n}^*\,.$$

An $n-$dimensional subspace, $\mathrm{span}\{\mathbf{a}_1,\ldots,\mathbf{a}_n\}$, of $\mathbb{V}$ is a *Lagrangian subspace*, equivalently a *Lagrangian plane*, if

$$\mathbf{\Omega}(\mathbf{a}_i,\mathbf{a}_j) = 0\,,\quad \forall\ i,j=1,\ldots,n\,. \tag{4.3}$$

The manifold of Lagrangian planes in $\mathbb{V}$ is denoted by $\Lambda(n)$.

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

$$\mathbf{\Omega}^{\mathrm{dual}} = \mathbf{\Omega}^{n-1}\ \lrcorner\,\mathsf{vol}\,, \tag{4.4}$$

where $\mathbf{\Omega}^k = \underbrace{\mathbf{\Omega}\wedge\cdots\wedge\mathbf{\Omega}}_{k \text{ times}}$ denotes the k-th exterior power of Ω. To give an explicit expression for the dual symplectic form defined in (4.4), first related the volume form to powers of the symplectic form,

$$\mathbf{\Omega}^n = (-1)^{\frac{n(n-1)}{2}}\, n!\, \mathsf{vol}^*\,. \tag{4.5}$$

This formula can be seen as a consequence of section 5. Substitution then shows that

$$\mathbf{\Omega}^{\rm dual} = (-1)^{\frac{n(n-1)}{2}}\,(n-1)!\,\sum_{j=1}^{n} \mathbf{e}_j \wedge \mathbf{e}_{j+n}\,. \tag{4.6}$$

In the case $n=2$, $\mathbf{\Omega}\wedge\mathbf{\Omega} = -2\,\mathsf{vol}^*$, and the dual symplectic form is defined by

$$\mathbf{\Omega}(\mathbf{a},\mathbf{b})\mathsf{vol} = \mathbf{\Omega}^{\rm dual}\wedge\mathbf{a}\wedge\mathbf{b}\,,\quad \forall\ \mathbf{a},\mathbf{b}\in\mathbb{V}\,. \tag{4.7}$$

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

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

4.1. Intersection index

Consider pairs of oriented Lagrangian planes. In what follows, we identify oriented subspaces of $\mathbb{V}$, say $\text{span}\{\mathbf{a}_1,\ldots,\mathbf{a}_n\}$, with the corresponding elements $\text{span}\{\mathbf{a}_1\wedge\cdots\wedge\mathbf{a}_n\}$ in $\bigwedge^n(\mathbb{V})$. Let U and V be two Lagrangian planes and define

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

Now suppose U and V have a $d-$dimensional intersection, and denote the intersection index by $\mathbf{O}_d(U,V)$. Then there exists vectors

$$\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_d,\mathbf{a}_1,\ldots,\mathbf{a}_{n-d},\mathbf{b}_1,\ldots,\mathbf{b}_{n-d}\in\mathbb{V}\,,$$

such that

$$\begin{aligned} U &:= \text{span}\{\boldsymbol{\xi}_1\wedge\cdots\wedge\boldsymbol{\xi}_d\wedge\mathbf{a}_1\wedge\cdots\wedge\mathbf{a}_{n-d}\},\\ V &:= \text{span}\{\boldsymbol{\xi}_1\wedge\cdots\wedge\boldsymbol{\xi}_d\wedge\mathbf{b}_1\wedge\cdots\wedge\mathbf{b}_{n-d}\}\,, \end{aligned}$$

and

$$\mathbf{a}_1\wedge\cdots\wedge\mathbf{a}_{n-d}\wedge\mathbf{b}_1\wedge\cdots\wedge\mathbf{b}_{n-d}\neq 0.$$

Definition 4.1 *The orientation index of the pair (U,V) is defined as:*

$$\mathbf{O}_d(U,V) = \text{sign}\det\begin{bmatrix} \mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_1) & \ldots & \mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_{n-d})\\ \vdots & & \vdots\\ \mathbf{\Omega}(\mathbf{a}_{n-d},\mathbf{b}_1) & \ldots & \mathbf{\Omega}(\mathbf{a}_{n-d},\mathbf{b}_{n-d}) \end{bmatrix}. \tag{4.8}$$

Two special cases, $d=1,2$, are of great interest and will be treated in more detail.

4.2. $d=1$ intersection

The $d=1$ intersection is transversal if

$$\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{n-1} \wedge \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_{n-1} \neq 0\,.$$

An equivalent definition is

$$\mathbf{\Omega}^{\rm dual} \wedge \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{n-1} \wedge \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_{n-1} \neq 0\,.$$

This follows since, using (4.4),

$$\begin{aligned}\mathbf{\Omega}^{\rm dual} \wedge \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{n-1} \wedge \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_{n-1} \quad &= \\ &= [\![\mathsf{vol}^*, \mathbf{\Omega}^{\rm dual} \wedge \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{n-1} \wedge \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_{n-1}]\!]_{2n} \mathsf{vol} \\ &= [\![\mathbf{\Omega}^{\rm dual} \lrcorner \, \mathsf{vol}^*, \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{n-1} \wedge \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_{n-1}]\!]_{2n-2} \mathsf{vol} \\ &= [\![\mathbf{\Omega}^{n-1}, \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{n-1} \wedge \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_{n-1}]\!]_{2n-2} \mathsf{vol} \\ &= \tau \mathsf{vol}\,,\end{aligned}$$

with

$$\tau = \det \begin{bmatrix} \mathbf{\Omega}(\mathbf{a}_1, \mathbf{b}_1) & \cdots & \mathbf{\Omega}(\mathbf{a}_1, \mathbf{b}_{n-1}) \\ \vdots & \ddots & \vdots \\ \mathbf{\Omega}(\mathbf{a}_{n-1}, \mathbf{b}_1) & \cdots & \mathbf{\Omega}(\mathbf{a}_{n-1}, \mathbf{b}_{n-1}) \end{bmatrix}. \tag{4.9}$$

Hence, an equivalent definition of the intersection index in this case is

$$\mathbf{O}_1(U,V) = {\rm sign}(\mathbf{\Omega}^{\rm dual} \wedge \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{n-1} \wedge \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_{n-1})\,. \tag{4.10}$$

In the case $n=2$, this formula simplifies. With $d=1$ and $n=2$, there exists vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{V}$ such that

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

The intersection index in this case is defined as

$$\mathbf{O}_1(U,V) = {\rm sign}(\mathbf{\Omega}(\mathbf{b}, \mathbf{c})) = {\rm sign}(\mathbf{\Omega}^{\rm dual} \wedge \mathbf{b} \wedge \mathbf{c})\,. \tag{4.11}$$

4.3. $d=2$ intersection

Loss of transversality in the $d=1$ intersection leads to investigation of the $d=2$ intersection. This $d=2$ intersection will be useful for the study of non transversely-constructed homoclinic orbits.

Lemma. *Let A, B be $n-$dimensional Lagrangian subspaces defined by*

$$A = {\rm span}\{\xi, \mathbf{a}_1, \ldots, \mathbf{a}_{n-1}\} \quad \textit{and} \quad B = {\rm span}\{\xi, \mathbf{b}_1, \ldots, \mathbf{b}_{n-1}\}\,, \tag{4.12}$$

with

$$\xi \wedge \mathbf{a}_1 \wedge \ldots \wedge \mathbf{a}_{n-1} \neq 0 \quad \textit{and} \quad \xi \wedge \mathbf{b}_1 \wedge \ldots \wedge \mathbf{b}_{n-1} \neq 0\,.$$

Then $\dim(A\bigcap B) \geq 2$ *if and only if* $\tau = 0$ *where* τ *is defined in (4.9).*

Proof. Suppose $\dim(A\bigcap B) \geq 2$. Then there exists constants $\beta_1,\ldots,\beta_{n-1}$ (not all zero) such that

$$\beta_1\mathbf{b}_1 + \cdots + \beta_{n-1}\mathbf{b}_{n-1} \in A \tag{4.13}$$

and so

$$\mathbf{\Omega}(\mathbf{a}_j, \beta_1\mathbf{b}_1 + \cdots + \beta_{n-1}\mathbf{b}_{n-1}) = 0 \quad \text{for} \quad j = 1,\ldots,n-1\,, \tag{4.14}$$

or

$$\begin{bmatrix} \mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_1) & \cdots & \mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_{n-1}) \\ \vdots & \ddots & \vdots \\ \mathbf{\Omega}(\mathbf{a}_{n-1},\mathbf{b}_1) & \cdots & \mathbf{\Omega}(\mathbf{a}_{n-1},\mathbf{b}_{n-1}) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4.15}$$

Since $\beta \neq 0$ it follows that the determinant must vanish giving $\tau = 0$.

Now suppose $\tau = 0$. The above argument can just be reversed: there exists constants $\beta_1,\ldots,\beta_{n-1}$ such that (4.15) holds, which in turn implies (4.14) holds which then implies (4.13). Hence $\tau = 0$ implies that the dimension of the intersection is two or greater. □

The case $n = 2$ is of special interest. 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 $\text{span}\{\mathbf{a},\mathbf{b}\}$ and $\text{span}\{\mathbf{a},\mathbf{c}\}$ are Lagrangian subspaces. Then

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

5. Volume form to symplectic determinants

In this section a proof of the formula (1.8) is given.

Let S_n be the set of permutations; that is mappings $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$ that are one-to-one. For $\sigma \in S_n$, let

$$\varepsilon(\sigma) := \prod_{1\leq i<j\leq n} \text{sign}\left(\sigma(j) - \sigma(i)\right),$$

be the parity of the permutation σ. The following lemma on permutations is required.

Lemma 5.1 *Define the mapping*

$$f : \begin{cases} S_n \times S_n \times \{0,1\}^n \to \{\sigma \in S_{2n} \ s.t.\ \forall k\ \{\sigma(2k-1),\sigma(2k)\} \not\subseteq \{1,\ldots,n\}\} \\ (\mu,\psi,u) \mapsto \sigma : \begin{cases} \{1,\ldots,2n\} \to \{1,\ldots,2n\} \\ 2k-1+u_k \mapsto \mu(k) \\ 2k-u_k \mapsto \psi(k)+n \end{cases} \end{cases}.$$

Then f *is a one-to-one correspondence and* $\varepsilon(f(\mu,\psi,u)) = (-1)^{\frac{n(n-1)}{2}}\varepsilon(\mu\psi)\prod_{i=1}^{n}(-1)^{u_i}$.

Proof: It is clear that f is well-defined and one-to-one. To prove that f is an onto mapping, let

$$\sigma \in S_{2n} \quad \text{such that} \quad \forall k\ \{\sigma(2k),\sigma(2k+1)\} \not\subseteq \{1,\ldots,n\}\,.$$

Then, for all k, the cardinality of $\{1,\dots,n\}\cap\{\sigma(2k-1),\sigma(2k)\}$ is one. Let u_k be such that $\sigma(2k-1+u_k)\in\{1,\dots,n\}$. Then define $\mu(k),\psi(k)$ as

$$\mu(k)=\sigma(2k-1+u_k) \quad\text{and}\quad \psi(k)=\sigma(2k-u_k)-n\,.$$

Then $\sigma=f(\mu,\psi,u)$, which proves the surjectivity.

It remains to prove the assertion on the parity of the permutation. Let (μ,ψ,u) be in $S_n\times S_n\times\{0,1\}^n$. Then, it is clear that $\varepsilon(f(\mu,\psi,u))=\prod_{i=1}^n(-1)^{u_i}f(\mu,\psi,0)$.

So let σ be $\sigma=f(\mu,\psi,0)$. Then:

$$\varepsilon(\sigma)=\operatorname{sign}\left(\prod_{2k<2l}\sigma(2l)-\sigma(2k)\prod_{2k<2l-1}\sigma(2l-1)-\sigma(2k)\right.$$
$$\left.\prod_{2k-1<2l}\sigma(2l)-\sigma(2k-1)\prod_{2k-1<2l-1}\sigma(2l-1)-\sigma(2k-1)\right)\,. \quad (5.1)$$

Hence,

$$\varepsilon(\sigma)=\operatorname{sign}\left(\prod_{k<l}\psi(l)-\psi(k)\prod_{k<l}\mu(l)-\psi(k)-n\prod_{k\leq l}\psi(l)+n-\mu(k)\prod_{k<l}\mu(l)-\mu(k)\right)$$

Therefore $\varepsilon(\sigma)=\varepsilon(\psi)(-1)^{\frac{n(n-1)}{2}}1^{\frac{n(n+1)}{2}}\varepsilon(\mu)$. This proves the lemma. □

5.1. *The n^{th} exterior power of the symplectic form*

For $1\leq k\leq r$, let ρ_k be an i_k-form and $t_k=\sum_{j=1}^k i_j$.

According to page 116 of [24], the wedge product of these forms is equal to:

$$\left(\bigwedge_{k=1}^r\rho_k\right)(\mathbf{a}_1,\dots,\mathbf{a}_{t_r})=\frac{1}{i_1!i_2!\dots i_r!}\sum_{\sigma\in S_n}\varepsilon(\sigma)\prod_{k=1}^r\rho_k(\mathbf{a}_{\sigma(t_{k-1}+1)},\dots,\mathbf{a}_{\sigma(t_k)})\,.$$

Let $\Delta:=\mathbf{\Omega}\wedge\dots\wedge\mathbf{\Omega}(\mathbf{a}_1,\dots,\mathbf{a}_{2n})$. Then we have:

$$\begin{aligned}
\Delta &= \tfrac{1}{2^n}\textstyle\sum_{\sigma\in S_{2n}}\varepsilon(\sigma)\mathbf{\Omega}(\mathbf{a}_{\sigma(1)},\mathbf{a}_{\sigma(2)})\dots\mathbf{\Omega}(\mathbf{a}_{\sigma(2n-1)},\mathbf{a}_{\sigma(2n)})\\
&= \tfrac{1}{2^n}\textstyle\sum_{\sigma\in S_{2n},\forall k\ \{\sigma(2k),\sigma(2k+1)\}\nsubseteq\{1,\dots,n\}}\varepsilon(\sigma)\mathbf{\Omega}(\mathbf{a}_{\sigma(1)},\mathbf{a}_{\sigma(2)})\dots\mathbf{\Omega}(\mathbf{a}_{\sigma(2n-1)},\mathbf{a}_{\sigma(2n)})\\
&\quad+\tfrac{1}{2^n}\textstyle\sum_{\sigma\in S_{2n},\exists k\ \{\sigma(2k),\sigma(2k+1)\}\subseteq\{1,\dots,n\}}\varepsilon(\sigma)\mathbf{\Omega}(\mathbf{a}_{\sigma(1)},\mathbf{a}_{\sigma(2)})\dots\mathbf{\Omega}(\mathbf{a}_{\sigma(2n-1)},\mathbf{a}_{\sigma(2n)})\\
&= \tfrac{2^n}{2^n}\textstyle\sum_{\mu,\psi\in S_n}(-1)^{\frac{n(n-1)}{2}}\varepsilon(\mu)\varepsilon(\psi)\mathbf{\Omega}(\mathbf{a}_{\mu(1)},\mathbf{a}_{n+\psi(1)})\dots\mathbf{\Omega}(\mathbf{a}_{\mu(n)},\mathbf{a}_{n+\psi(n)})\\
&\quad+\tfrac{1}{2^n}\textstyle\sum_{\sigma\in S_{2n},\exists k\ \{\sigma(2k),\sigma(2k+1)\}\subseteq\{1,\dots,n\}}\varepsilon(\sigma)\mathbf{\Omega}(\mathbf{a}_{\sigma(1)},\mathbf{a}_{\sigma(2)})\dots\mathbf{\Omega}(\mathbf{a}_{\sigma(2n-1)},\mathbf{a}_{\sigma(2n)})\\
&= (-1)^{\frac{n(n-1)}{2}}n!\det\begin{bmatrix}\mathbf{\Omega}(\mathbf{a}_1,\mathbf{a}_{n+1}) & \dots & \mathbf{\Omega}(\mathbf{a}_1,\mathbf{a}_{n+k}) & \dots & \mathbf{\Omega}(\mathbf{a}_1,\mathbf{a}_{2n})\\ \vdots & \ddots & \vdots & & \vdots\\ \mathbf{\Omega}(\mathbf{a}_k,\mathbf{a}_{n+1}) & \dots & \mathbf{\Omega}(\mathbf{a}_k,\mathbf{a}_{n+k}) & \dots & \mathbf{\Omega}(\mathbf{a}_k,\mathbf{a}_{2n})\\ \vdots & & \vdots & \ddots & \vdots\\ \mathbf{\Omega}(\mathbf{a}_n,\mathbf{a}_{n+1}) & \dots & \mathbf{\Omega}(\mathbf{a}_n,\mathbf{a}_{n+k}) & \dots & \mathbf{\Omega}(\mathbf{a}_n,\mathbf{a}_{2n})\end{bmatrix}+(-1)^{\frac{n(n-1)}{2}}n!\Upsilon\,,
\end{aligned}$$

with

$$\Upsilon = \frac{(-1)^{\frac{n(n-1)}{2}}}{n!2^n} \sum_{\sigma\in S_{2n},\exists k\ \{\sigma(2k-1),\sigma(2k)\}\subseteq\{1,\dots,n\}} \varepsilon(\sigma)\mathbf{\Omega}(\mathbf{a}_{\sigma(1)},\mathbf{a}_{\sigma(2)})\dots\mathbf{\Omega}(\mathbf{a}_{\sigma(2n-1)},\mathbf{a}_{\sigma(2n)})\,.$$

This proves the formula (1.8), noting that the left-hand side is related to the volume form (see equation (4.5)). A special case is as follows. Let $\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n,\mathbf{f}_1,\mathbf{f}_2,\dots,\mathbf{f}_n$ be a basis for $\mathbb{V}$ normalized such that

$$\mathbf{\Omega}(\mathbf{e}_i,\mathbf{e}_j)=0\,,\quad \mathbf{\Omega}(\mathbf{e}_i,\mathbf{f}_j)=\delta_{ij}\quad\text{and}\quad \mathbf{\Omega}(\mathbf{f}_i,\mathbf{f}_j)=0\,,$$

then

$$\mathbf{\Omega}\wedge\cdots\wedge\mathbf{\Omega}(\mathbf{e}_1,\dots,\mathbf{e}_n,\mathbf{f}_1,\dots,\mathbf{f}_n)=(-1)^{\frac{n(n-1)}{2}}n! \tag{5.2}$$

□

5.2. The formula in the cases $\dim(\mathbb{V})=4,6$

The formula simplifies in low dimension. In the case $n=2$:

Proposition 5.2 *Let* $\mathbf{a}$*,* $\mathbf{b}$*,* $\mathbf{c}$ *and* $\mathbf{d}$ *be any four vectors in* $\mathbb{V}$*. Then*

$$\mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d}=\det\begin{bmatrix}\mathbf{\Omega}(\mathbf{a},\mathbf{c}) & \mathbf{\Omega}(\mathbf{a},\mathbf{d})\\ \mathbf{\Omega}(\mathbf{b},\mathbf{c}) & \mathbf{\Omega}(\mathbf{b},\mathbf{d})\end{bmatrix}\mathsf{vol}-\mathbf{\Omega}(\mathbf{a},\mathbf{b})\mathbf{\Omega}(\mathbf{c},\mathbf{d})\mathsf{vol}\,. \tag{5.3}$$

Corollary 5.3 *If either* $\mathrm{span}\{\mathbf{a},\mathbf{b}\}$ *or* $\mathrm{span}\{\mathbf{c},\mathbf{d}\}$ *is a Lagrangian subspace. Then the formula reduces to*

$$\mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d}=\det\begin{bmatrix}\mathbf{\Omega}(\mathbf{a},\mathbf{c}) & \mathbf{\Omega}(\mathbf{a},\mathbf{d})\\ \mathbf{\Omega}(\mathbf{b},\mathbf{c}) & \mathbf{\Omega}(\mathbf{b},\mathbf{d})\end{bmatrix}\mathsf{vol}\,. \tag{5.4}$$

The case with $n>2$ of greatest use in applications is the case $n=3$. An explicit formula for this case is given in the following.

Proposition 5.4 *Suppose* $\dim(\mathbb{V})=6$ *and let* A,B *be two three-dimensional subspaces defined by*

$$A=\mathrm{span}\{\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3\}\quad\textit{and}\quad B=\mathrm{span}\{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3\}\,.$$

Then

$$\mathbf{a}_1\wedge\mathbf{a}_2\wedge\mathbf{a}_3\wedge\mathbf{b}_1\wedge\mathbf{b}_2\wedge\mathbf{b}_3=\det\begin{bmatrix}\mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_1) & \mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_2) & \mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_3)\\ \mathbf{\Omega}(\mathbf{a}_2,\mathbf{b}_1) & \mathbf{\Omega}(\mathbf{a}_2,\mathbf{b}_2) & \mathbf{\Omega}(\mathbf{a}_2,\mathbf{b}_3)\\ \mathbf{\Omega}(\mathbf{a}_3,\mathbf{b}_1) & \mathbf{\Omega}(\mathbf{a}_3,\mathbf{b}_2) & \mathbf{\Omega}(\mathbf{a}_3,\mathbf{b}_3)\end{bmatrix}\mathsf{vol}+\Upsilon\,\mathsf{vol}\,,$$

where

$$\begin{aligned}\Upsilon \;=\;& -\mathbf{\Omega}(\mathbf{a}_1,\mathbf{a}_2)\left[\mathbf{\Omega}(\mathbf{a}_3,\mathbf{b}_1)\mathbf{\Omega}(\mathbf{b}_2,\mathbf{b}_3)-\det\begin{pmatrix}\mathbf{\Omega}(\mathbf{a}_3,\mathbf{b}_2) & \mathbf{\Omega}(\mathbf{a}_3,\mathbf{b}_3)\\ \mathbf{\Omega}(\mathbf{b}_1,\mathbf{b}_2) & \mathbf{\Omega}(\mathbf{b}_1,\mathbf{b}_3)\end{pmatrix}\right]\\ &+\mathbf{\Omega}(\mathbf{a}_1,\mathbf{a}_3)\left[\mathbf{\Omega}(\mathbf{a}_2,\mathbf{b}_1)\mathbf{\Omega}(\mathbf{b}_2,\mathbf{b}_3)-\det\begin{pmatrix}\mathbf{\Omega}(\mathbf{a}_2,\mathbf{b}_2) & \mathbf{\Omega}(\mathbf{a}_2,\mathbf{b}_3)\\ \mathbf{\Omega}(\mathbf{b}_1,\mathbf{b}_2) & \mathbf{\Omega}(\mathbf{b}_1,\mathbf{b}_3)\end{pmatrix}\right]\\ &-\mathbf{\Omega}(\mathbf{a}_2,\mathbf{a}_3)\left[\mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_1)\mathbf{\Omega}(\mathbf{b}_2,\mathbf{b}_3)-\det\begin{pmatrix}\mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_2) & \mathbf{\Omega}(\mathbf{a}_1,\mathbf{b}_3)\\ \mathbf{\Omega}(\mathbf{b}_1,\mathbf{b}_2) & \mathbf{\Omega}(\mathbf{b}_1,\mathbf{b}_3)\end{pmatrix}\right]\,.\end{aligned}$$

Corollary. *If either A or B is a Lagrangian plane then $\Upsilon = 0$ and the formula reduces to*

$$\mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \mathbf{a}_3 \wedge \mathbf{b}_1 \wedge \mathbf{b}_2 \wedge \mathbf{b}_3 = \det \begin{bmatrix} \boldsymbol{\Omega}(\mathbf{a}_1,\mathbf{b}_1) & \boldsymbol{\Omega}(\mathbf{a}_1,\mathbf{b}_2) & \boldsymbol{\Omega}(\mathbf{a}_1,\mathbf{b}_3) \\ \boldsymbol{\Omega}(\mathbf{a}_2,\mathbf{b}_1) & \boldsymbol{\Omega}(\mathbf{a}_2,\mathbf{b}_2) & \boldsymbol{\Omega}(\mathbf{a}_2,\mathbf{b}_3) \\ \boldsymbol{\Omega}(\mathbf{a}_3,\mathbf{b}_1) & \boldsymbol{\Omega}(\mathbf{a}_3,\mathbf{b}_2) & \boldsymbol{\Omega}(\mathbf{a}_3,\mathbf{b}_3) \end{bmatrix} \mathsf{vol}\,.$$

Proofs follow by evaluating the formula for Υ in §5.1.

6. Transversely constructed homoclinic orbits

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

$$\mathbf{J}\mathbf{u}_x = DH(\mathbf{u},p)\,, \quad \mathbf{u} \in \mathbb{V}\,, \tag{6.1}$$

with

$$\lim_{x\to\pm\infty} \widehat{\mathbf{u}}(x,p) = 0 \quad \text{and} \quad 0 < \int_{-\infty}^{+\infty} \left|\widehat{\mathbf{u}}(x,p)\right|^2 \mathrm{d}x < +\infty\,. \tag{6.2}$$

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 $2(n-1)$ other tangent vectors in $\mathbb{V}$, denoted by $\mathbf{a}_j^-(x)$ and $\mathbf{a}_j^+(x)$ for $j = 1,\ldots,n-1$ satisfying

$$\frac{d}{dx}\mathbf{a}_j^\pm = \mathbf{A}(x,0)\mathbf{a}_j^\pm\,, \quad \text{with} \quad \mathbf{A}(x,0) := \mathbf{J}^{-1}D^2H(\widehat{\mathbf{u}},p)\,, \tag{6.3}$$

and

$$\begin{aligned} E^s(x,0) &= \mathrm{span}\{\widehat{\mathbf{u}}_x, \mathbf{a}_1^+,\ldots,\mathbf{a}_{n-1}^+\}; \quad \mathbf{a}_j^+ \to 0 \text{ as } x \to +\infty\,, \\ E^u(x,0) &= \mathrm{span}\{\widehat{\mathbf{u}}_x, \mathbf{a}_1^-,\ldots,\mathbf{a}_{n-1}^-\}; \quad \mathbf{a}_j^- \to 0 \text{ as } x \to -\infty\,. \end{aligned}$$

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 6.1 *For the homoclinic orbit $\widehat{\mathbf{u}}(x,p)$, define*

$$\Xi(x,p) := \mathbf{a}_1^-(x,p) \wedge \cdots \wedge \mathbf{a}_{n-1}^-(x,p) \wedge \mathbf{a}_1^+(x,p) \wedge \cdots \wedge \mathbf{a}_{n-1}^+(x,p)\,. \tag{6.4}$$

The homoclinic orbit is said to be "transversely constructed" if $\Xi(x,p) \neq 0$ for all $x \in \mathbb{R}$.

Proposition 6.2 *If $\Xi(x_0,p) = 0$ ($\Xi(x_0,p) \neq 0$) for some $x_0 \in \mathbb{R}$ then $\Xi(x,p) = 0$ ($\Xi(x,p) \neq 0$) for all $x \in \mathbb{R}$.*

Proof. The proof follows from the fact that $\Xi(x,p)$ satisfies an ordinary differential equation

$$\frac{d}{dx}\Xi(x,p) = \mathbf{A}^{(2n-2)}(x,0)\Xi(x,p)\,,$$

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

Definition 6.3 *The Lazutkin-Treschev invariant associated with a homoclinic orbit is*

$$\mathcal{T}(\widehat{\mathbf{u}}) = \det \begin{bmatrix} \mathbf{\Omega}(\mathbf{a}_1^-, \mathbf{a}_1^+) & \cdots & \mathbf{\Omega}(\mathbf{a}_1^-, \mathbf{a}_{n-1}^+) \\ \vdots & \ddots & \vdots \\ \mathbf{\Omega}(\mathbf{a}_{n-1}^-, \mathbf{a}_1^+) & \cdots & \mathbf{\Omega}(\mathbf{a}_{n-1}^-, \mathbf{a}_{n-1}^+) \end{bmatrix} \tag{6.5}$$

Theorem 6.4 *A homoclinic orbit is "transversely constructed" if and only if the Lazutkin-Treschev homoclinic invariant is nonzero.*

Proof. The Lazutkin-Treschev invariant is independent of x. This follows since $\mathbf{a}_j^\pm$, $j = 1, \ldots, n-1$, are solutions of (6.3) and the symplectic form is independent of x when evaluated on any two solutions of (6.3).

Now suppose the Lazutkin-Treschev invariant is zero. Then by the Lemma in §4.1, the stable and unstable subspaces have a (at least) two-dimensional intersection (for each x) and so the intersection is not transverse.

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

$$\begin{aligned} 0 &= \mathbf{\Omega}^{\rm dual} \wedge \Xi(x,p) \\ &= [\![\mathsf{vol}^*, \mathbf{\Omega}^{\rm dual} \wedge \Xi,]\!]_{2n} \mathsf{vol} \\ &= [\![\mathbf{\Omega}^{\rm dual} \lrcorner\, \mathsf{vol}^*, \Xi]\!]_{2n-2} \mathsf{vol} \\ &= [\![\mathbf{\Omega} \wedge \cdots \wedge \mathbf{\Omega}, \Xi]\!]_{2n-2} \mathsf{vol} \\ &= \tau \mathsf{vol}\,, \end{aligned}$$

proving that the Lazutkin-Treschev invariant is zero. □

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

Definition 6.5 *Suppose that the stable and unstable subspaces are normalized as follows*

$$\lim_{x \to +\infty} \mathrm{e}^{2(\mu_1^+ + \cdots + \mu_n^+)x}\, \mathcal{E}^+(x) \wedge \mathcal{E}^-(x) = \rho\, \mathsf{vol}\,, \rho > 0\,, \tag{6.6}$$

where

$$\mathcal{E}^\pm(x) = \widehat{\mathbf{u}}_x(\pm x) \wedge \mathbf{a}_1^\pm(\pm x) \wedge \cdots \wedge \mathbf{a}_{n-1}^\pm(\pm x)\,,$$

and $\mu_1^+, \ldots, \mu_n^+$ are the eigenvalues of the linearization at infinity with negative real parts (cf. §8). Then the sign of the homoclinic orbit is defined to be $\mathrm{sign}(\mathcal{T})$.

According to Alexander, Gardner & Jones [1], $\lim_{x\to\infty} \mathrm{e}^{(\mu_1^+ + \cdots + \mu_n^+)x} \mathcal{E}^+(x)$ and $\lim_{x\to\infty} \mathrm{e}^{(\mu_1^+ + \cdots + \mu_n^+)x} \mathcal{E}^-(x)$ are well-defined, nonzero and span respectively the unstable and the stable of $\mathbf{J}^{-1} D^2 H(0,p)$. As a consequence, $\lim_{x\to\infty} \mathrm{e}^{(\mu_1^+ + \cdots + \mu_n^+)x} \mathcal{E}^+(x) \wedge \lim_{x\to\infty} \mathrm{e}^{(\mu_1^+ + \cdots + \mu_n^+)x} \mathcal{E}^-(x) \neq 0$ since the stable space and the unstable space are transverse. It is sufficient to divide one of the $\mathbf{a}_i^\pm$ by a constant to have the suitable normalization.

7. Example: an explicit calculation of $\mathcal{T}(\widehat{\mathbf{u}})$

Take $n = 2$ and $\mathbf{D} = \mathbf{I}$ and

$$F(\mathbf{v}) = -2(v_1^2 + v_2^2) + 2(v_1^3 + v_2^3) - \tfrac{1}{2}p(v_1 - v_2)^2\,,$$

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

$$\begin{aligned} \frac{\partial v_1}{\partial t} &= \frac{\partial^2 v_1}{\partial x^2} - 4v_1 + 6v_1^2 - p(v_1 - v_2) \\ \frac{\partial v_2}{\partial t} &= \frac{\partial^2 v_2}{\partial x^2} - 4v_2 + 6v_2^2 + p(v_1 - v_2)\,. \end{aligned} \tag{7.1}$$

This system was studied in §11 of [12] (with the parameter p here replaced by c there). The system (7.1) has an exact steady solution $v_1 = v_2 := \widehat{v}(x) = \mathrm{sech}^2(x)$. 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-Treschev invariant is calculated.

The tangent vector to the homoclinic orbit is

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

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

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

where

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

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

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

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

$$p = -\frac{3}{2}\,,\quad p = 0\,,\quad p = \frac{5}{2}\,. \tag{7.2}$$

The above form for $\mathbf{a}^\pm$ is chosen so that the normalization (6.6) is operational. Therefore the formula (1.5) should hold. Indeed this can be checked directly. According to §11 of [12], the values of p in (7.2) 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 proved in section 11:

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

This example also reminds that the Lazutkin-Treschev 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.

8. The symplectic Evans function

Suppose that the Hamiltonian system (6.1) has a homoclinic orbit as in §6. 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,

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

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

$$\mathbf{u}_x = \mathbf{A}(x,\lambda)\mathbf{u}\,, \quad \mathbf{u} \in \mathbb{V}\,, \tag{8.1}$$

with

$$\mathbf{A}(x,\lambda) = \mathbf{J}^{-1}[\mathbf{B}(x,p) - \lambda\mathbf{M}]\,. \tag{8.2}$$

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

$$\mathbf{J}\mathbf{A}_\infty(\lambda) = \big[\mathbf{B}_\infty - \lambda\mathbf{M}\big]\,, \tag{8.3}$$

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 (8.1) if there exists a solution $\mathbf{u}(x,\lambda)$ to (8.1) 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 our attention to real λ, which can almost be proved in general. Suppose λ and $\mathbf{u}$ are complex and let us decompose them into their real and imaginary parts:

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

Substitute into (8.1), 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

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

Since it is assumed that the linearization about the trivial solution is hyperbolic, it is clear that $0 \notin \sigma_{ess}$. Since σ_{ess} is a closed set, there exists an open interval containing 0 with no essential spectrum in it.

Now the Evans function can be constructed in the usual way for $\lambda \notin \sigma_{ess}$. Denote the eigenvalues of $\mathbf{A}_\infty(\lambda)$ with negative real part by $\mu_1^+(\lambda), \ldots, \mu_n^+(\lambda)$ and the eigenvalues with positive real part by $\mu_1^-(\lambda), \ldots, \mu_n^-(\lambda)$, with eigenvectors

$$\big[\mathbf{B}_\infty - \lambda\mathbf{M}\big]\xi_j^\pm = \mu_j^\pm \mathbf{J}\xi_j^\pm\,, \quad i = 1, \ldots, n\,. \tag{8.4}$$

With the assumption of strict hyperbolicity, the eigenvalues can be simple, strictly real and nonzero, simple and complex with non-zero real part, or non-simple. 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 [6].

Now define solutions of (8.1) that decay to zero as $x \to +\infty$ with the asymptotic properties

$$\lim_{x\to+\infty} \mathrm{e}^{-\mu_j^+(\lambda)x}\mathbf{u}_j^+(x,\lambda) = \xi_j^+(\lambda)\,, \quad j = 1, \ldots, n, \tag{8.5}$$

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

$$\lim_{x\to-\infty} \mathrm{e}^{-\mu_j^-(\lambda)x}\mathbf{u}_j^-(x,\lambda) = \xi_j^-(\lambda)\,, \quad j = 1, \ldots, n\,. \tag{8.6}$$

Then the natural definition of the Evans function is

$$D(\lambda)\mathsf{vol} = \mathbf{u}_1^+(x,\lambda) \wedge \cdots \wedge \mathbf{u}_n^+(x,\lambda) \wedge \mathbf{u}_1^-(x,\lambda) \wedge \cdots \wedge \mathbf{u}_n^-(x,\lambda)\,. \tag{8.7}$$

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 (8.1) with $\lambda = 0$.

8.1. Symplectification of the Evans function

By working directly with the Evans function as a $2n-$form (8.7) it is not immediately clear how to take advantage of the symplectic structure. Since the stable and unstable subspaces are Lagrangian, the correction term Υ in (1.8) vanishes and so application of (1.8) to (8.7) gives the following formula for the Evans function

$$\mathrm{D}(\lambda) = \det[\boldsymbol{\Sigma}(\lambda)]\,, \quad \boldsymbol{\Sigma}(\lambda) = \begin{bmatrix} \boldsymbol{\Omega}(\mathbf{u}_1^-,\mathbf{u}_1^+) & \cdots & \boldsymbol{\Omega}(\mathbf{u}_1^-,\mathbf{u}_n^+) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Omega}(\mathbf{u}_n^-,\mathbf{u}_1^+) & \cdots & \boldsymbol{\Omega}(\mathbf{u}_n^-,\mathbf{u}_n^+) \end{bmatrix} . \tag{8.8}$$

With this formula a symplectic proof that $D(0) = 0$ can be given. Taking the limit $\lambda \to 0$, the Evans function reduces to

$$\mathrm{D}(0) = \det \begin{bmatrix} \boldsymbol{\Omega}(\widehat{\mathbf{u}}_x,\widehat{\mathbf{u}}_x) & \boldsymbol{\Omega}(\widehat{\mathbf{u}}_x,\mathbf{a}_1^+) & \cdots & \boldsymbol{\Omega}(\widehat{\mathbf{u}}_x,\mathbf{a}_{n-1}^+) \\ \boldsymbol{\Omega}(\mathbf{a}_1^-,\widehat{\mathbf{u}}_x) & \boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_1^+) & \cdots & \boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_{n-1}^+) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Omega}(\mathbf{a}_{n-1}^-,\widehat{\mathbf{u}}_x) & \boldsymbol{\Omega}(\mathbf{a}_{n-1}^-,\mathbf{a}_1^+) & \cdots & \boldsymbol{\Omega}(\mathbf{a}_{n-1}^-,\mathbf{a}_{n-1}^+) \end{bmatrix} .$$

Now, $\boldsymbol{\Omega}(\widehat{\mathbf{u}}_x,\widehat{\mathbf{u}}_x) = 0$ by skew symmetry, and

$$\boldsymbol{\Omega}(\widehat{\mathbf{u}}_x,\mathbf{a}_j^+) = \boldsymbol{\Omega}(\mathbf{a}_j^-,\widehat{\mathbf{u}}_x) = 0\,, \quad j = 1,\ldots,n-1\,,$$

since the stable and unstable subspaces are Lagrangian subspaces. Hence

$$\mathrm{D}(0) = \det[\boldsymbol{\Sigma}(0)] = 0\,, \quad \text{since } \boldsymbol{\Sigma}(0) = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_1^+) & \cdots & \boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_{n-1}^+) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \boldsymbol{\Omega}(\mathbf{a}_{n-1}^-,\mathbf{a}_1^+) & \cdots & \boldsymbol{\Omega}(\mathbf{a}_{n-1}^-,\mathbf{a}_{n-1}^+) \end{bmatrix} .$$

8.2. The derivative at $\lambda = 0$

The derivative of $\mathrm{D}(\lambda)$ is

$$\mathrm{D}'(\lambda) = \mathrm{Trace}\left(\boldsymbol{\Sigma}(\lambda)^{\#}\boldsymbol{\Sigma}'(\lambda)\right)\,, \tag{8.9}$$

where $\boldsymbol{\Sigma}(\lambda)^{\#}$ is the adjugate. Since we are only interested in the first derivative, take the limit as $\lambda \to 0$. This limit takes the remarkably simple form

$$\lim_{\lambda\to 0} \boldsymbol{\Sigma}(\lambda)^{\#} = \begin{bmatrix} \mathcal{T}(\widehat{\mathbf{u}}) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} .$$

Substitute into the expression for the derivative (8.9)

$$\mathrm{D}'(0) = \mathcal{T}(\widehat{\mathbf{u}})\mathrm{Trace}\left(\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \boldsymbol{\Sigma}'(0) \right) , \tag{8.10}$$

or

$$\mathrm{D}'(0) = \mathcal{T}(\widehat{\mathbf{u}})\frac{d}{d\lambda}\mathbf{\Omega}(\mathbf{u}_1^-,\mathbf{u}_1^+)\Big|_{\lambda=0}\,. \tag{8.11}$$

Using results on $\lambda-$derivatives [5, 6] it follows that

$$\frac{d}{d\lambda}\mathbf{\Omega}\left(\mathbf{u}_1^-,\mathbf{u}_1^+\right)\Big|_{x=0} = \int_{-\infty}^{+\infty}\left\langle \mathbf{u}_1^-,\mathbf{B}_\lambda\mathbf{u}_1^+\right\rangle \,\mathrm{d}x\,.$$

Evaluation at $\lambda = 0$ and substitution into (8.11) then gives

$$\mathrm{D}'(0) = -\mathcal{T}(\widehat{\mathbf{u}})\,\left[\int_{-\infty}^{+\infty}\langle \widehat{\mathbf{u}}_x, \mathbf{M}\widehat{\mathbf{u}}_x\rangle\,\mathrm{d}x\right]\,, \tag{8.12}$$

proving the following Theorem.

Theorem. *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 $\mathcal{T}(\widehat{\mathbf{u}})$.

9. 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

$$\lim_{x\to\pm\infty}\widehat{\phi}(x,p) = 0 \quad \text{and} \quad 0 < \int_{-\infty}^{+\infty}|\widehat{\phi}_x|^2\,\mathrm{d}x < +\infty\,. \tag{9.1}$$

It could be a simple homoclinic orbit or a multi-pulse homoclinic orbit. Such solutions have been widely studied (e.g. see [11, 27] and references therein). The linearization about such solutions in the time-dependent equation, with in addition a spectral ansatz, leads to the spectral problem

$$\mathscr{L}\phi = \lambda\phi\,, \tag{9.2}$$

where

$$\mathscr{L}\phi := -\phi_{xxxx} - p\phi_{xx} - \phi + 2\widehat{\phi}\phi\,. \tag{9.3}$$

The theory of this paper leads to a new proof of Lemma 2.1(iii) in [27].

Lemma 9.1 (Sandstede [27]). *Any homoclinic orbit of the steady SH equation satisfying (9.1) with $-2 < p < 2$ is transversely constructed if and only if zero is a simple eigenvalue of $\mathscr{L}$ in (9.2).*

Proof The spectral problem (9.2) can be recast into the form (8.1). The hypothesis on the essential spectrum is satisfied for $-2 < p < +2$, and with the properties (9.1) the formula (1.7) applies. Hence

$$D'(0) = -\mathfrak{T}(\widehat{\mathbf{u}}) \int_{-\infty}^{+\infty} \langle \mathbf{M}\widehat{\mathbf{u}}_x, \widehat{\mathbf{u}}_x \rangle \, \mathrm{d}x = -\mathfrak{T}(\widehat{\mathbf{u}}) \int_{-\infty}^{+\infty} \widehat{\phi}_x^2 \, \mathrm{d}x \,,$$

using the form of $\mathbf{M}$ in (2.3). Hypothesis (9.1) assures that the integral exists and is non-vanishing. Hence $D'(0) = 0$ if and only if $\mathfrak{T}(\widehat{\mathbf{u}}) \neq 0$. The proof is completed by applying Theorem 6.4. □.

The spectral problem here (9.2) 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].

10. The Maslov index à la Souriau

To prove the formula (1.5) we need a definition for the Maslov index of homoclinic orbits. There are a range of definitions in the literature. The predominant definition is to take a path of Lagrangian subspaces $L(x)$ (in this case the path of stable subspaces in the linearization about the homoclinic orbit) and count the number of intersections with a reference Lagrangian subspace. It can also be based on a triple (L_1, L_2, L_3) of Lagrangian subspaces. A third approach is to take a pair $(\widetilde{L}_1, \widetilde{L}_2)$ of elements in the universal cover of $\Lambda(n)$. It is proved in Cappell et al. [7] that all three approaches to defining the Maslov index are equivalent. In previous work [11, 12, 13], the first approach was used. Here a variant of the third formulation due to Souriau [28] is used, which makes it easier to compare with the Lazutkin-Treschev invariant. A comparison with other definitions can be found in [22].

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

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

with

$$\widetilde{U(n)} = \left\{ (\mathbf{U}, \kappa) \;\middle|\; \mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} \,,\ \mathbf{U}_1 + \mathrm{i}\mathbf{U}_1 \in U(n) \,,\ \mathrm{e}^{-\mathrm{i}\frac{\kappa}{2}} = \det(\mathbf{U}_1 + \mathrm{i}\mathbf{U}_2) \right\} . \quad (10.1)$$

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

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

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

$$\psi(\mathbf{U},\mathbf{V}) = (\mathbf{U}_1 - \mathrm{i}\mathbf{U}_2)^{-1}(\mathbf{V}_1 - \mathrm{i}\mathbf{V}_2)(\mathbf{V}_1 + \mathrm{i}\mathbf{V}_2)^{-1}(\mathbf{U}_1 + \mathrm{i}\mathbf{U}_2)\,. \tag{10.2}$$

$\psi(\mathbf{U},\mathbf{V})$ is a symmetric unitary matrix (cf. §1.1.2 of [9]):

$$\psi^H = \psi^{-1} \quad \text{and} \quad \psi^T = \psi \quad \Rightarrow \quad \overline{\psi}\psi = \mathbf{I}\,. \tag{10.3}$$

Hence its eigenvalues are on the unit circle and can be expressed in the form:

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

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

Proposition 10.1 *Let* $\mathbf{D} = \begin{bmatrix} \mathrm{e}^{\frac{1}{2}\mathrm{i}\alpha_1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{\frac{1}{2}\mathrm{i}\alpha_n} \end{bmatrix}$. *There exists* $n\times n$ *orthogonal matrices* $\mathbf{R}$ *and* $\mathbf{S}$ *such that:*

$$\mathbf{R}^{-1}\psi(\mathbf{U},\mathbf{V})\mathbf{R} = \mathbf{D}^2, \quad (\mathbf{U}_1\mathbf{R} + \mathrm{i}\mathbf{U}_2\mathbf{R})\overline{\mathbf{D}} = (\mathbf{V}_1\mathbf{S} + \mathrm{i}\mathbf{V}_2\mathbf{S})\,.$$

Proof As mentioned in [21], by using equation (10.3), it is clear that ψ and $\overline{\psi}$ commute with each other. Therefore $\Re\psi$ and $\Im\psi$ also commute with each other. As a consequence, there exists $\mathbf{R} \in O(n)$ such that $\mathbf{R}^{-1}\Re\psi\mathbf{R}$ and $\mathbf{R}^{-1}\Im\psi\mathbf{R}$ are diagonal. By choosing an appropriate permutations of columns, one can choose $\mathbf{R}$ such that:

$$\mathbf{R}^{-1}\psi(\mathbf{U},\mathbf{V})\mathbf{R} = \mathbf{D}^2 = \mathbf{D}\overline{\mathbf{D}}^{-1}$$

If we denote $\mathbf{S} = (\mathbf{V}_1 + \mathrm{i}\mathbf{V}_2)^{-1}(\mathbf{U}_1 + \mathrm{i}\mathbf{U}_2)\mathbf{R}\overline{\mathbf{D}}$, then the previous equality implies that:

$$\overline{\mathbf{S}} = \mathbf{S}$$

Hence, $\mathbf{S}$ is a real orthogonal matrix and we have:

$$(\mathbf{U}_1\mathbf{R} + \mathrm{i}\mathbf{U}_2\mathbf{R})\overline{\mathbf{D}} = (\mathbf{V}_1\mathbf{S} + \mathrm{i}\mathbf{V}_2\mathbf{S})\,.$$

Proposition 10.2 *Let* $\mathbf{r}_1,\ldots,\mathbf{r}_n$ *be the columns of* $\mathbf{UR}$, $\mathbf{r}_{n+i} = \mathbf{J}\mathbf{r}_i$, *and* $\mathbf{s}_1,\ldots,\mathbf{s}_n$ *be the columns of* $\mathbf{VS}$. *Then:*

$$\mathbf{s}_i = \cos(\frac{\alpha_i}{2})\mathbf{r}_i - \sin(\frac{\alpha_i}{2})\mathbf{r}_{n+i}$$

Let d *be the multiplicity of* 1 *as an eigenvalue of* ψ. *Then the intersection of the spaces spanned by* $\mathbf{U}$ *and* $\mathbf{V}$ *(which are the same as the one spanned by* $\mathbf{UR}$ *and* $\mathbf{VS}$, *but not necessarily with the same orientation) is the space spanned by* $\mathbf{r}_1,\ldots,\mathbf{r}_d$

Proof First, we have:

$$\mathbf{UR}(\Re\overline{\mathbf{D}}) + \mathbf{JUR}(\Im\overline{\mathbf{D}}) = \mathbf{VS}$$

Then it is clear that $\mathbf{s}_i = \cos(\frac{\alpha_i}{2})\mathbf{r}_i - \sin(\frac{\alpha_i}{2})\mathbf{r}_{n+i}$. Besides:

$$\mathbf{\Omega}(\mathbf{r}_i, \mathbf{s}_j) = \langle \mathbf{J}\mathbf{r}_i, \cos(\frac{\alpha_j}{2})\mathbf{r}_j - \sin(\frac{\alpha_j}{2})\mathbf{J}\mathbf{s}_j \rangle = -\sin(\frac{\alpha_j}{2})\langle \mathbf{r}_i, \mathbf{s}_j\rangle = -\sin(\frac{\alpha_i}{2})\delta_{ij}\,.$$

We have that:

- $\mathbf{r}_i = \mathbf{s}_i$ if $i \leq d$.
- $\mathbf{r}_1 \wedge \ldots \wedge \mathbf{r}_n \wedge \mathbf{s}_{d+1} \wedge \ldots \wedge \mathbf{s}_n = (-1)^{n-d} \sin(\alpha_{d+1}) \cdot \ldots \cdot \sin(\alpha_n) \mathbf{r}_1 \wedge \ldots \wedge \mathbf{r}_n \wedge \mathbf{r}_{n+d+1} \wedge \ldots \wedge \mathbf{r}_{2n} \neq 0$.

Therefore, $\pi(\mathbf{U}, \kappa) \cap \pi(\mathbf{V}, \tau)$ contains the space spanned by $\mathbf{r}_1, \ldots, \mathbf{r}_d$ and its dimension is smaller $2n - (2n - d) = d$.

We conclude that $\pi(\mathbf{U}, \kappa) \cap \pi(\mathbf{V}, \tau)$ is the space spanned by $\mathbf{r}_1, \ldots, \mathbf{r}_d$. □

According to Souriau's formula (cf. pages 126–128 of [28]), 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 $\widetilde{U(n)}$. The Maslov index of this pair of elements is defined by:

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

where

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

It is essential that $\alpha_1, \ldots, \alpha_n \in [0, 2\pi)$, or there would be several possible values of m. It is clear that m is an integer when d is even and an half-integer when d is odd.

10.1. Maslov index of a homoclinic orbit

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 $\widetilde{U(n)}$ be such that:

$$\begin{cases} \text{span } \mathbf{U}^+(x) = \text{span}\{\widehat{\mathbf{u}}_x(x), \mathbf{a}_1^+(x), \ldots, \mathbf{a}_{n-1}^+(x)\} \\ \text{span } \mathbf{U}^-(x) = \text{span}\{\widehat{\mathbf{u}}_x(x), \mathbf{a}_1^-(x), \ldots, \mathbf{a}_{n-1}^-(x)\} \\ \kappa^+, \kappa^- \text{ are continuous} \end{cases}.$$

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

$$\begin{aligned} I_{hom}(\widehat{\mathbf{u}}) \;&=\; m\big((\mathbf{U}^-(x), \kappa^-(x)), (\mathbf{U}^+(x), \kappa^+(x))\big) \\ &\quad - \lim_{y \to +\infty} m\big((\mathbf{U}^-(-y), \kappa^-(-y)), (\mathbf{U}^+(y), \kappa^+(y))\big)\,, \end{aligned} \tag{10.4}$$

and the definition is independent of x.

In the previous definition, it is important that κ^+, κ^- are continuous. As a consequence, these functions are unique up to a shift by a multiple of 4π. If we shift κ^+ (resp. κ^-) by $4k\pi$, then the left of the minus sign is shifted by $2k$ (resp. $-2k$) and the right of the minus is shifted by $-2k$ (resp. $2k$). This guarantees that $\mathbf{I}_{hom}$ is independent of the choice of κ^+ and κ^-.

11. 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 11.1 *Suppose $\widehat{\mathbf{u}}$ is a transversely constructed homoclinic orbit with Maslov index $I_{\rm hom}(\widehat{\mathbf{u}})$. Then*

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

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 11.2 *Let $(\mathbf{U},\kappa),(\mathbf{V},\tau) \in \widetilde{U(n)}$ 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 n-forms. Then*

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

Proof First, let $\mathbf{R},\mathbf{S},\mathbf{D},\mathbf{r}_i,\mathbf{s}_i$ be matrices as defined in propositions 10.1 and 10.2.

We have:

$$\mathbf{O}_d(\mathbf{U}\mathbf{R}^\wedge,\mathbf{V}\mathbf{S}^\wedge) = \operatorname{sign}\det\begin{bmatrix}\mathbf{\Omega}(\mathbf{r}_{d+1},\mathbf{s}_{d+1}) & \dots & \mathbf{\Omega}(\mathbf{r}_{d+1},\mathbf{s}_n)\\ \vdots & & \vdots \\ \mathbf{\Omega}(\mathbf{r}_n,\mathbf{s}_{d+1}) & \dots & \mathbf{\Omega}(\mathbf{r}_n,\mathbf{s}_n)\end{bmatrix}$$
$$= \operatorname{sign}\det\begin{bmatrix}-\sin(\frac{\alpha_{d+1}}{2}) & & 0\\ & \ddots & \\ 0 & & -\sin(\frac{\alpha_n}{2})\end{bmatrix} = (-1)^{n-d} = (-1)^{n+d}$$

As a consequence we have that:

$$\begin{aligned}\mathbf{O}_d(\mathbf{U}^\wedge,\mathbf{V}^\wedge) &= \operatorname{sign}(\det \mathbf{R}\mathbf{S})\mathbf{O}_d(\mathbf{U}\mathbf{R}^\wedge,\mathbf{V}\mathbf{S}^\wedge)\\ &= \operatorname{sign}\det((\mathbf{U}_1+\mathrm{i}\mathbf{U}_2)(\mathbf{V}_1+\mathrm{i}\mathbf{V}_2)^{-1}\overline{\mathbf{D}})(-1)^{d+n}\\ &= e^{-\frac{1}{2}\mathrm{i}(\kappa-\tau)}e^{-\frac{1}{2}\mathrm{i}(\alpha_1+\cdots+\alpha_n)}(-1)^{n+d} = e^{\mathrm{i}\pi(m-\frac{1}{2}d)}(-1)^{d+n} = (-1)^{m+\frac{1}{2}d+n}\end{aligned}$$

□

The Maslov index, in the Souriau representation, for a homoclinic orbit is defined in (10.4). Use Lemma 11.2 above to conclude the proof of Theorem 11.1,

$$\begin{aligned}(-1)^{I_{hom}(\widehat{\mathbf{u}})+\frac{1}{2}} &= \mathbf{O}_1\big((\widehat{\mathbf{u}}_x\wedge\mathbf{a}_1^-(x)\wedge\ldots\wedge\mathbf{a}_{n-1}^-(x)),(\widehat{\mathbf{u}}_x\wedge\mathbf{a}_1^+(x)\wedge\ldots\wedge\mathbf{a}_{n-1}^+(x))\big)\\ &\quad\times\lim_{y\to+\infty}\mathbf{O}_0\big((\widehat{\mathbf{u}}_x\wedge\mathbf{a}_2^-(-y)\wedge\ldots\wedge\mathbf{a}_n^-(-y)),(\widehat{\mathbf{u}}_x(y)\wedge\mathbf{a}_2^+(y)\wedge\ldots\wedge\mathbf{a}_n^+(y))\big)\\ &\quad\times(-1)^{\frac{1}{2}-(n+\frac{1}{2})+n}\end{aligned} \tag{11.1}$$

But the right-hand side is just the sign of $\mathcal{T}(\widehat{\mathbf{u}})$. The right term of the product is 1, because of the normalisation. Hence, we conclude that

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

□

12. Example: coupled system of reaction-diffusion PDEs

Now consider the system of reaction-diffusion equations (1.3) with $n=3$ and $\mathbf{D}=\mathbf{I}$

$$\begin{aligned}\frac{\partial v_1}{\partial t} &= \frac{\partial^2 v_1}{\partial x^2}-4v_1+6v_1^2-c_1(v_1-v_2)+c_3(v_3-v_1)\\ \frac{\partial v_2}{\partial t} &= \frac{\partial^2 v_2}{\partial x^2}-4v_2+6v_2^2+c_1(v_1-v_2)-c_2(v_2-v_3)\\ \frac{\partial v_3}{\partial t} &= \frac{\partial^2 v_3}{\partial x^2}-4v_3+6v_3^2+c_2(v_2-v_3)-c_3(v_3-v_1)\,,\end{aligned} \tag{12.1}$$

where $\mathbf{c}=(c_1,c_2,c_3)$ is a non-zero vector-valued coupling parameter. This example generalizes the study of coupled reaction-diffusion equations with $n=2$ in [12] and §7. It can be formulated as in (1.1) by taking $\mathbf{J}$ in standard form (2.2) with $n=3$,

$$\mathbf{u}:=(\mathbf{v},\mathbf{p}):=(\mathbf{v},\mathbf{v}_x)\quad\text{and}\quad\mathbf{M}=\begin{bmatrix}\mathbf{I}&\mathbf{0}\\\mathbf{0}&\mathbf{0}\end{bmatrix}.$$

Then (12.1) can be written in the form (1.1) with

$$H(\mathbf{u})=\tfrac{1}{2}(u_4^2+u_5^2+u_6^2)-2(u_1^2+u_2^2+u_3^2)+2(u_1^3+u_2^3+u_3^2)+V(\mathbf{u})\,,$$

with

$$V(\mathbf{u})=-\tfrac{1}{2}c_1(u_1-u_2)^2-\tfrac{1}{2}c_2(u_2-u_3)^2-\tfrac{1}{2}c_3(u_3-u_1)^2\,.$$

This system has the exact steady solitary-wave solution

$$v_1=v_2=v_3:=\widehat{u}(x)=\operatorname{sech}^2(x)\,.$$

Linearizing (12.1) about the basic state $\widehat{u}$ and taking perturbations of the form

$$\mathrm{e}^{\lambda t}(v_1(x,\lambda),v_2(x,\lambda),v_3(x,\lambda))\,,$$

leads to the coupled ODE eigenvalue problem

$$\mathbf{v}_{xx} = a(x,\lambda)\mathbf{v} - \mathbf{C}\mathbf{v}\,, \tag{12.2}$$

where

$$a(x,\lambda) = \lambda + 4 + \mathrm{Trace}(\mathbf{C}) - 12\mathrm{sech}^2(x)\,,$$

and

$$\mathbf{C} = \begin{pmatrix} c_2 & c_1 & c_3 \\ c_1 & c_3 & c_2 \\ c_3 & c_2 & c_1 \end{pmatrix}\,. \tag{12.3}$$

The eigenvalue problem (12.2) can be solved explicitly. First diagonalize the symmetric matrix $\mathbf{C}$. Denote its real eigenvalues by σ_1, σ_2 and σ_3. They satisfy

$$0 = \det[\sigma\mathbf{I} - \mathbf{C}] = (\sigma - \mathrm{Trace}(\mathbf{C}))(\sigma^2 - \gamma^2)\,,$$

with

$$\gamma = \frac{1}{\sqrt{2}}\left[(c_1-c_2)^2 + (c_2-c_3)^3 + (c_3-c_1)^2\right]^{1/2}\,.$$

Hence the three eigenvalues of $\mathbf{C}$ are

$$\sigma_1 = \mathrm{Trace}(\mathbf{C})\,,\quad \sigma_2 = -\gamma\,,\quad \sigma_3 = +\gamma\,.$$

In practice it may be of interest to choose $\mathbf{c}$ so that the trivial solution of (12.1) is temporally stable. The condition for temporal stability is

$$4 + \mathrm{Trace}(\mathbf{C}) \pm \gamma > 0\,. \tag{12.4}$$

Let $\mathbf{T}$ be the 3×3 matrix whose columns are the eigenvectors of $\mathbf{C}$. Hence

$$\mathbf{T}^{-1}\mathbf{C}\mathbf{T} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}\,.$$

Use this transformation to diagonalize the eigenvalue problem (12.2). Let $\mathbf{v} = \mathbf{T}\widetilde{\mathbf{v}}$, then $\widetilde{\mathbf{v}}$ satisfies

$$\begin{pmatrix} \widetilde{v}_1 \\ \widetilde{v}_2 \\ \widetilde{v}_3 \end{pmatrix}_{xx} = \begin{bmatrix} a(x,\lambda)-\sigma_1 & 0 & 0 \\ 0 & a(x,\lambda)-\sigma_2 & 0 \\ 0 & 0 & a(x,\lambda)-\sigma_3 \end{bmatrix} \begin{pmatrix} \widetilde{v}_1 \\ \widetilde{v}_2 \\ \widetilde{v}_3 \end{pmatrix}\,. \tag{12.5}$$

12.1. Lazutkin-Treschev invariant calculation

Set $\lambda = 0$. The three decoupled ODEs in (12.5) represent the three stable and unstable subspaces. Since $\sigma_1 = \mathrm{Trace}(\mathbf{C})$, the first equation for $\widetilde{v}_1$,

$$(\widetilde{v}_1)_{xx} + 12\mathrm{sech}^2(x)\widetilde{v}_1 = 4\widetilde{v}_1$$

is the building block for $\widehat{u}_x$. The other two equations,

$$\begin{aligned}(\widetilde{v}_2)_{xx} + 12\text{sech}^2(x)\widetilde{v}_2 &= \kappa_1\,\widetilde{v}_2\\ (\widetilde{v}_3)_{xx} + 12\text{sech}^2(x)\widetilde{v}_3 &= \kappa_2\,\widetilde{v}_3\,,\end{aligned} \tag{12.6}$$

are the building blocks for $\mathbf{a}_1^\pm$ and $\mathbf{a}_2^\pm$. The coefficients κ_1 and κ_2 are

$$\kappa_1 = 4 + \text{Trace}(\mathbf{C}) + \gamma \quad \text{and} \quad \kappa_2 = 4 + \text{Trace}(\mathbf{C}) - \gamma\,. \tag{12.7}$$

Using the result in Appendix B of [12], the two equations (12.6) can be explicitly solved. Their solutions are

$$\begin{aligned}\widetilde{v}_2^\pm &= \mathrm{e}^{\mp\sqrt{\kappa_1}x}\big(\mp b_0^{(1)} + b_1^{(1)}\tanh(x) \mp b_2^{(1)}\tanh^2(x) + \tanh^3(x)\big)\\ \widetilde{v}_3^\pm &= \mathrm{e}^{\mp\sqrt{\kappa_2}x}\big(\mp b_0^{(2)} + b_1^{(2)}\tanh(x) \mp b_2^{(2)}\tanh^2(x) + \tanh^3(x)\big)\,,\end{aligned} \tag{12.8}$$

with

$$b_0^{(j)} = \frac{1}{15}\sqrt{\kappa_j}(4-\kappa_j)\,, \quad b_1^{(j)} = \frac{1}{5}(2\kappa_j - 3) \quad b_2^{(j)} = -\sqrt{\kappa_j}\,.$$

Denote the columns of $\mathbf{T}$ (eigenvectors of $\mathbf{C}$) by $\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_3$ and suppose the columns are orthonormalized. Then the basis vectors for the stable and unstable subspaces are

$$\mathbf{a}_1^\pm = \begin{pmatrix}\widetilde{v}_1^\pm\mathbf{t}_2\\ (\widetilde{v}_2^\pm)_x\mathbf{t}_2\end{pmatrix} \quad \text{and} \quad \mathbf{a}_2^\pm = \begin{pmatrix}\widetilde{v}_3^\pm\mathbf{t}_3\\ (\widetilde{v}_3^\pm)_x\mathbf{t}_3\end{pmatrix}\,.$$

Use these vectors to compute the Lazutkin-Treschev invariant

$$\mathcal{T}(\widehat{\mathbf{u}}) = \det\begin{bmatrix}\boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_1^+) & \boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_2^+)\\ \boldsymbol{\Omega}(\mathbf{a}_2^-,\mathbf{a}_1^+) & \boldsymbol{\Omega}(\mathbf{a}_2^-,\mathbf{a}_2^+)\end{bmatrix}\,. \tag{12.9}$$

Since $\mathbf{t}_2$ and $\mathbf{t}_3$ are orthogonal it follows that

$$\boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_2^+) = \boldsymbol{\Omega}(\mathbf{a}_2^-,\mathbf{a}_1^+) = 0\,,$$

and so the matrix in (12.9) is diagonal and

$$\mathcal{T}(\widehat{\mathbf{u}}) = \boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_1^+)\,\boldsymbol{\Omega}(\mathbf{a}_2^-,\mathbf{a}_2^+)\,. \tag{12.10}$$

But

$$\begin{aligned}\boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_1^+) &= \big(\widetilde{v}_2^-(\widetilde{v}_2)_x^+ - \widetilde{v}_2^+(\widetilde{v}_2)_x^-\big)\\ &= 2b_0^{(1)}\big(\sqrt{\kappa_1}b_0^{(1)} + b_1^{(1)}\big)\\ &= \tfrac{2}{225}\sqrt{\kappa_1}(4-\kappa_1)\big(\kappa_1(4-\kappa_1) + 3(2\kappa_1 - 3)\big)\,,\end{aligned}$$

which simplifies to

$$\boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_1^+) = \frac{2}{225}\sqrt{\kappa_1}(\kappa_1 - 1)(\kappa_1 - 4)(\kappa_1 - 9)\,. \tag{12.11}$$

Similarly

$$\boldsymbol{\Omega}(\mathbf{a}_2^-,\mathbf{a}_2^+) = \frac{2}{225}\sqrt{\kappa_2}(\kappa_2-1)(\kappa_2-4)(\kappa_2-9)\,, \tag{12.12}$$

where κ_1 and κ_2 are functions of $\mathbf{c}$ given in (12.7). Hence there are two independent ways that the Lazutkin-Treschev invariant can vanish giving loss of transversality: either $\boldsymbol{\Omega}(\mathbf{a}_1^-,\mathbf{a}_1^+)=0$ or $\boldsymbol{\Omega}(\mathbf{a}_2^-,\mathbf{a}_2^+)=0$.

To determine whether the vectors $\mathbf{u}_x, \mathbf{a}_1^\pm, \mathbf{a}_2^\pm$ are normalized as in (6.6), let us compute :

$$\begin{aligned}
&\lim_{x\to\infty} e^{2x(2+\sqrt{\kappa_1}+\sqrt{\kappa_2})}\mathbf{u}_x(-x)\wedge\mathbf{a}_1^-(-x)\wedge\mathbf{a}_2^-(-x)\wedge\mathbf{u}_x(x)\wedge\mathbf{a}_1^+(x)\wedge\mathbf{a}_2^+(x)\\
&= \lim_{x\to\infty} e^{2x(2+\sqrt{\kappa_1}+\sqrt{\kappa_2})}\boldsymbol{\Omega}(\mathbf{u}_x(-x),\mathbf{u}_x(x))\boldsymbol{\Omega}(\mathbf{a}_1^-(-x),\mathbf{a}_1^+(x))\boldsymbol{\Omega}(\mathbf{a}_2^-(-x),\mathbf{a}_2^+(x))\mathsf{vol}\\
&= \lim_{x\to\infty} e^{2x(2+\sqrt{\kappa_1}+\sqrt{\kappa_2})}(\tilde v_1(-x)(\tilde v_1)_x(x)-\tilde v_1(x)(\tilde v_1)_x(-x))(\tilde v_2^-(-x)(\tilde v_2)_x^+(x)-\tilde v_2^+(x)(\tilde v_2)_x^-(-x))\\
&\qquad\times(\tilde v_2^-(-x)(\tilde v_3)_x^+(x)-\tilde v_3^+(x)(\tilde v_3)_x^-(-x))\mathsf{vol}\\
&= \lim_{x\to\infty} e^{2x(2+\sqrt{\kappa_1}+\sqrt{\kappa_2})}(-2\tilde v_1(\tilde v_1)_x)(-2\tilde v_2^+(\tilde v_2)_x^+)(-2\tilde v_3^+(\tilde v_3)_x^+)\mathsf{vol}\\
&= -\lim_{x\to\infty} e^{2x(2+\sqrt{\kappa_1}+\sqrt{\kappa_2})}\big((\tilde v_1)^2\big)_x\big((\tilde v_2^+)^2\big)_x\big((\tilde v_3^+)^2\big)_x\,\mathsf{vol}.
\end{aligned}$$

The third equality was obtained by noticing that $\tilde v_1(-x)=-\tilde v_1(x)$, $\tilde v_2^-(-x)=-\tilde v_2^+(x)$, $\tilde v_3^-(-x)=-\tilde v_3^+(x)$. When x is close to $+\infty$, $(\tilde v_1^+)^2,(\tilde v_2^+)^2,(\tilde v_3^+)^2$ are decreasing. Hence, the normalising factor is positive. As a consequence, the Lazutkin-Treschev invariant has the appropriate normalization.

12.1.1. The case where $\mathbf{c}=c(1,1,1)$ In the special case $\mathbf{c}=c(1,1,1)$ the Maslov index was computed in [13]. In this case $\gamma=0$ and $\mathrm{Trace}(\mathbf{C})=3c$, and so $\kappa_1=\kappa_2:=\kappa=4+3c$ and the Lazutkin-Treschev invariant reduces to

$$\mathcal{T}(\widehat{\mathbf{u}}) = \left(\frac{2}{225}\right)^2\kappa(\kappa-1)^2(\kappa-4)^2(\kappa-9)^2\,. \tag{12.13}$$

With the condition $\kappa>0$ (which follows from (12.4)), it is non-negative. Substituting for κ in terms of c in (12.13),

$$\mathcal{T}(\widehat{\mathbf{u}}) = \left[\frac{18}{225}\sqrt{4+3c}\,(c+1)c(3c-5)\right]^2\,, \tag{12.14}$$

giving that $\mathcal{T}(\widehat{\mathbf{u}})\geq 0$. Therefore, according to the formula

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

the Maslov index must be even, and jumps by an even number when $\mathcal{T}(\widehat{\mathbf{u}})=0$. Hence we can expect a jump of two in the Maslov index when $c=-1$, $c=0$ and $c=\frac{5}{3}$. In this case, the latter observation can be checked by explicitly computing the Maslov index.

12.2. Calculation of the Maslov index

Using the results in [13] a formula for λ in the point spectrum can be computed. The eigenvalue problem (12.2) can be re-formulated in terms of the Evans function

$$\mathbf{w}_x = \mathbf{A}(x,\lambda)\mathbf{w}\,,$$

by taking $\mathbf{w} = (\mathbf{v}, \mathbf{v}_x)$ and

$$\mathbf{A}(x,\lambda) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ f(x,\lambda)+c_1+c_3 & -c_1 & -c_3 & 0 & 0 & 0 \\ -c_1 & f(x,\lambda)+c_1+c_2 & -c_2 & 0 & 0 & 0 \\ -c_3 & -c_2 & f(x,\lambda)+c_2+c_3 & 0 & 0 & 0 \end{bmatrix}, \quad (12.15)$$

with $f(x,\lambda) = \lambda + 4 - 12\,\mathrm{sech}^2(x)$. The matrix in (12.15) is Hamiltonian: $\mathbf{JA}$ is symmetric.

There are exactly nine eigenvalues in the point spectrum of this eigenvalue problem,

$$\left.\begin{array}{rcl} \lambda_j &=& -c_1 - c_2 - c_3 + \sigma_j - 3 \\ \lambda_{j+3} &=& -c_1 - c_2 - c_3 + \sigma_j \\ \lambda_{j+6} &=& -c_1 - c_2 - c_3 + \sigma_j + 5 \end{array}\right\} \quad j = 1,2,3\,.$$

When $c_1 = c_2 = c_3 := c$ then $\gamma = 0$ and $\sigma_1 = 3c$, and $\sigma_2 = \sigma_3 = 0$. Hence the nine eigenvalues reduce to

$$\big\{\lambda \;:\; \lambda = (-3-3c, -3c, -3c+5, -3, 0, 5)\big\}\,,$$

with the first three having multiplicity two. Hence the number of positive eigenvalues without multiplicity is 4, 3, 2, or 1 depending on whether $c < -1$, $-1 < c < 0$, $0 < c < 5/3$ or $c > 5/3$ respectively. According to Lemma 6 of [12], the Maslov index at λ counts the eigenvalues with multiplicities greater than λ, so the Maslov index can be explicitly written down:

c	$c < -1$	$-1 < c < 0$	$0 < c < \frac{5}{3}$	$c > \frac{5}{3}$
Maslov	8	6	4	2

As expected from the positive sign of the Lazutkin-Treschev invariant, the number $\mathsf{Maslov} = I_{hom}(\mathbf{u}) + \frac{1}{2}$ is even.

12.3. Generalization to $N-$coupled reaction-diffusion equations

This model can be generalized to $N-$coupled reaction diffusion equations. Let $\mathbf{v} = (v_1, \ldots, v_N)$ and define

$$V(\mathbf{v}) = -\tfrac{1}{2}\sum_{j=1}^{N-1} c_j(v_j - v_{j+1})^2 - \tfrac{1}{2}c_N(v_N - v_1)^2 \,.$$

Then the following system is a gradient reaction-diffusion system

$$\frac{\partial v_j}{\partial t} = \frac{\partial^2 v_j}{\partial x^2} - 4u_j + 6v_j^2 + \frac{\partial V}{\partial v_j}\,, \quad j = 1, \ldots, N\,, \tag{12.16}$$

which generalizes (12.1) to $N-$coupled equations. The steady part of this equation is a Hamiltonian system on a phase space of dimension $2N$. Taking $v_j(x) = \mathrm{sech}^2 x$ as the basic state and linearizing about it, the spectral problem can be explicitly solved in terms of the eigenvalues of the matrix $\mathbf{C}$. It appears that in this case the Lazutkin-Treschev invariant can be constructed so that it is the determinant of a diagonal matrix, e.g. $\mathcal{T}(\widehat{\mathbf{u}}) = \boldsymbol{\Omega}(\mathbf{a}_1^-, \mathbf{a}_1^+) \cdots \boldsymbol{\Omega}(\mathbf{a}_{N-1}^-, \mathbf{a}_{N-1}^+)$.

13. Concluding remarks

Our interest in this paper is in connecting transversality, the Lazutkin-Treschev invariant, the Maslov index and the Evans function. However, one of the main interests in geometric invariants of homoclinic orbits is to relate them to stability, when the homoclinic orbit represents a solitary wave. For example, the Maslov index has been related to stability of solitary waves for the nonlinear Schrödinger equation [23], gradient reaction-diffusion equations [3], Fitzhugh-Nagumo type systems [15], and some Hamiltonian PDEs [11, 12, 13]. An additional new direction that is now possible is a new proof of the sufficient condition for instability in [4, 5] in the case where the Hamiltonian PDE is multi-symplectic – the main difference is that the matrix $\mathbf{M}$ in (1.1) is skew-symmetric, and so $D'(0) = 0$ and the second derivative needs to be calculated. Combining the results in this paper with the proof in [4, 5] suggests that $D''(0)$ will be proportional to a product of the Treschev-Lazutkin invariant and the derivative of the momentum with respect to the speed of the solitary wave.

Acknowledgements

This work has been partly supported by ANR project BoND, ANR-13-BS01-0009-01.

References

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

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

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

[4] Bridges, T J and Derks, G 1999 Unstable eigenvalues and the linearization about solitary waves and fronts with symmetry, *Proc. R. Soc. Lond. A* **455** 2427–2469.

[5] Bridges, T J & Derks, G 2001 The symplectic Evans matrix, and the instability of solitary waves and fronts, *Arch. Rat. Mech. Anal.* **156** 1–87.

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

[7] Cappell, S E, Lee, R and Miller, E Y 1994 On the Maslov index, *Comm. Pure Appl. Math.* **47** 121–186.

[8] Chardard, F 2007 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.

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

[10] Chardard, F and Bridges, T J 2014 Transversality of homoclinic orbits, the Maslov index, and the symplectic Evans function: details of the '$n = 2$ case, *Technical Report.*

[11] Chardard, F, Dias, F and Bridges, T J 2009 On the Maslov index of multi-pulse homoclinic orbits, *Proc. R. Soc. Lond. A* **465** 2897–2910.

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

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

[14] Chen, C-N and Hu, X 2007 Maslov index for homoclinic orbits of Hamiltonian systems, *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **24** 589–603.

[15] Chen, C-N & Hu, X 2014 Stability analysis for standing pulse solutions to Fitzhugh-Nagumo equations, *Calc. Var.* **49** 827–845.

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

[17] Floer, A 1988 The unregularized gradient flow of the symplectic action, *Comm. Pure Appl. Math.* **41** 775–813.

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

[19] Gelfreich, V G and Lazutkin, V F 2001 Splitting of separatrices: perturbation theory and exponential smallness, *Russ. Math. Surv.* **56** 499–558.

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

[21] de Gosson, M 2006 *Symplectic Geometry and Quantum Mechanics*, Operator Theory: Advances and Applications **166**, Birkhaäuser Verlag: Basel.

[22] de Gosson, M 2009 On the usefulness of an index due to Leray for studying the intersections of Lagrangian and symplectic paths, *J. Math. Pures Appl.* **91** 598–613.

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

[24] Lang, S 2002 *Introduction to differentiable manifolds*, 2nd Edition, Springer: New York.

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

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

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

[28] Souriau, J-M 1976 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.

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