Journal of Differential Equations (in press)

Degenerate relative equilibria, curvature of the momentum map, and homoclinic bifurcation

Thomas J. Bridges

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

Abstract

A fundamental class of solutions of symmetric Hamiltonian systems is relative equilibria. In this paper the nonlinear problem near a degenerate relative equilibrium is considered. The degeneneracy creates a saddle-center and attendant homoclinic bifurcation in the reduced system transverse to the group orbit. The surprising result is that the curvature of the pullback of the momentum map to the Lie algebra determines the normal form for the homoclinic bifurcation. There is also an induced directional geometric phase in the homoclinic bifurcation. The backbone of the analysis is the use of singularity theory for smooth mappings between manifolds applied to the pullback of the momentum map. The theory is constructive and generalities are given for symmetric Hamiltonian systems on a vector space of dimension $(2n + 2)$ with an $n-$dimensional abelian symmetry group. Examples for $n = 1, 2, 3$ are presented to illustrate application of the theory.

Table of Contents

1 Introduction

A fundamental class of solutions of Hamiltonian systems with symmetry is relative equilibria. A relative equilibrium (RE) is a solution which travels along an orbit of the symmetry group at constant speed. They are pervasive in applications such as celestial mechanics, molecular dynamics, rigid-body dynamics, and fluid mechanics. An introduction to the theory of RE can be found in Chapter 4 of Marsden [16].

Consider a Hamiltonian system $\mathbf{z}_t = X_H(\mathbf{z})$ on a symplectic manifold (M, Ω) with Hamiltonian function $H : M \to \mathbb{R}$. Let G be an $n-$dimensional abelian Lie group acting symplectically on M with Lie algebra $\mathfrak{g}$. A RE is a solution of the form $\mathbf{z}(t) = \exp(t\xi)\varphi$ for some $\xi \in \mathfrak{g}$. The point $\mathbf{z}(0) = \varphi \in M$ is a critical point of the augmented Hamiltonian

$$H_\xi(\mathbf{z}) := H(\mathbf{z}) - \langle \mathbf{J}(\mathbf{z}) - \mu, \xi \rangle , \tag{1.1}$$

where $\mathbf{J} : M \to \mathfrak{g}^*$ is the momentum map, $\mu \in \mathfrak{g}^*$, and $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ is the pairing between $\mathfrak{g}$ and the dual of the Lie algebra $\mathfrak{g}^*$.

A RE is said to be non-degenerate when the second variation of H_ξ at a critical point is a non-degenerate quadratic form on the subspace consisting of vectors that are tangent to $\mathbf{J}^{-1}(\mu)$ and transverse to the group orbit (cf. Smale [33] and Chapter 5 of Marsden [16]; see also Patrick & Roberts [29] for a generalization of non-degeneracy).

There are three ways that an RE can become degenerate and they are clarified in §2 and §9. The first two are by failure of the so-called G-Morse [1] hypothesis: the dimension of the kernel of the second variation of H_ξ is greater than the dimension of the group. There are two ways that the G-Morse hypothesis can fail and for the purposes of this paper they will be called degeneracy of type II and degeneracy of type III. Type II degeneracy is a failure of the G-Morse hypothesis that arises naturally without external parameters, and type III degeneracy is failure of the G-Morse hypothesis due to external parameters in the Hamiltonian (both of these degeneracies are discussed in §9). The third form of degeneracy (which for the purposes of this paper will be called degeneracy of type I) is when the derivative of the pullback of the momentum map by φ is not surjective. Other characterizations of this degeneracy are given in §2. Degeneracy of type I is the main topic of this paper.

There is a fourth form of degeneracy that arises when $\mu \in \mathfrak{g}^*$ is a singular value of the momentum map (cf. Ortega & Ratiu [26]). However, this degeneracy is distinct from types I, II and III and is not considered here. Indeed, a blanket hypothesis throughout the paper is that $\mu \in \mathfrak{g}^*$ is a regular value of the momentum map.

A critical point φ of H_ξ in (1.1) can be viewed as a mapping from $\mathfrak{g}^{\rm RE}$ to M, where $\mathfrak{g}^{\rm RE}$ is the subset of $\mathfrak{g}$ for which critical points of H_ξ exist. Define the pullback $\mathbf{P} : \mathfrak{g}^{\rm RE} \to \mathfrak{g}^*$ of the momentum map by $\mathbf{P} = \mathbf{J} \circ \varphi$. (There is an orbit of critical points so the image of φ should be viewed as lying in a slice.) Then, taking coordinates $\mathbf{c} = (c_1, \ldots, c_n)$

[1] G-Morse is also called "equivariant Morse" and is a special case of Morse-Bott [25].

for $\mathfrak{g}$, a RE is non-degenerate if and only if it satisfies the G-Morse hypothesis and

$$\det(\mathrm{D}\mathbf{P}(\mathbf{c})) \neq 0\,, \quad \mathbf{c} \in \mathfrak{g}^{\mathrm{RE}}\,. \tag{1.2}$$

Satisfying the G-Morse hypothesis eliminates degeneracy of types II and III. Therefore, with the G-Morse hypothesis, a RE is degenerate if and only if $\det(\mathrm{D}\mathbf{P}(\mathbf{c})) = 0$.

Since the main topic of the paper is the implication of type I degeneracy, the G-Morse hypothesis is assumed throughout the paper. However, some discussion of the failure of the G-Morse hypothesis is given in §9 to illustrate how it is complementary to type I degeneracy.

The nonlinear consequences of degeneracy of RE do not appear to have been studied before. Degenerate RE have been widely observed, particularly in the $N-$body problem (cf. Palmore [27], Meyer [20]), and in fluid mechanics (cf. Bridges [4], Bridges & Donaldson [6]). Palmore [27] characterizes the subset of the space of masses in the $N-$body problem where degeneracy occurs, and shows that such degeneracies are plentiful. The "degenerate relative equilibria" discovered by Palmore [27] are however type III degeneracies (see discussion in §9).

The first implication of a type I degeneracy is that additional zero eigenvalues are generated in the linearization about degenerate RE. The connection between degenerate RE and zero eigenvalues has been observed in the literature before (e.g. Meyer [20]). Here a new proof and generalization of this result is given using symplectic Jordan chain theory. Effectively, a degenerate RE generates a saddle-center bifurcation of eigenvalues in the linearization transverse to the group orbit. Given a saddle-center in the linearization, it is well-known in Hamiltonian bifurcation theory that the nonlinear system nearby has a homoclinic bifurcation [3,21,23]. The idea is to combine the geometry of degenerate RE with this homoclinic bifurcation.

The geometry of the mapping $\mathbf{P} : \mathfrak{g}^{\mathrm{RE}} \to \mathfrak{g}^*$ is the backbone of the analysis. The theory is local, so $\mathbf{P}$ can be interpreted as a mapping from an open subset of $\mathfrak{g}^{\mathrm{RE}}$ into $\mathfrak{g}^*$. Degeneracy of the form $\det(\mathrm{D}\mathbf{P}(\mathbf{c})) = 0$ of a smooth mapping $\mathbf{P}$ between manifolds is a well-studied problem in singularity theory (cf. Chapter 2 of Arnold, Gusein-Zade & Varchenko [2], Chapter VI of Golubitsky & Guillemin [10]). The subsets of $\mathfrak{g}^{\mathrm{RE}}$ where the condition (1.2) fails are called the Thom-Boardman singularities of the mapping and are denoted in singularity theory by

$$\Sigma^k(\mathbf{P}) = \{\, \mathbf{c} \in \mathfrak{g} \ :\ \mathrm{D}\mathbf{P}(\mathbf{c}) \text{ has rank } n-k \,\}\,.$$

In this paper, attention will be restricted to the case $k = 1$. There can however be singular subsets in the image of $\Sigma^1(\mathbf{P})$ in $\mathfrak{g}^*$. The simplest such set is denoted by

$$\Sigma^{11}(\mathbf{P}) = \Sigma^1\left(\mathbf{P}\Big|_{\Sigma^1(\mathbf{P})}\right)\,.$$

That is, the set where the differential of $\mathbf{P}\Big|_{\Sigma^1(\mathbf{P})}$ has a one-dimensional kernel. Geometrically this is the set where the kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$ lies in the tangent space of $\Sigma^1(\mathbf{P})$. This classification can be continued until the dimension is exhausted [2,10]. In this paper, attention will be restricted to a study of the implications for $\mathbf{P}(\mathbf{c})$ of the hypersurface $\Sigma^1(\mathbf{P})$ and its subset $\Sigma^{11}(\mathbf{P})$.

The main result of this paper is the connection between the geometry of the manifold $\Sigma^1(\mathbf{P})$, its image in $\mathfrak{g}^*$, and the homoclinic bifurcation that occurs in the dynamics transverse to the group orbit in phase space. In particular suppose there exists a family of RE with $\mathbf{c} \in \Sigma^1(\mathbf{P}) \subset \mathfrak{g}^{\rm RE}$ but $\mathbf{c} \notin \Sigma^{11}(\mathbf{P})$. Then, in the reduced system transverse to the group orbit the nonlinearity which generates the homoclinic orbit is determined by the curvature of the graph of $\mathbf{P}$ in $\mathfrak{g} \times \mathfrak{g}^*$, and the homoclinic bifurcation transverse to the group induces a directional geometric phase along the group.

Precise hypotheses are stated in §2. These hypotheses are not the most general under which the phenomena occurs. Indeed, the hypotheses are chosen to highlight the simplest possible setting. Generalities are discussed at the end of the introduction.

The curvature of the pullback of the momentum map arises as follows. When $\mathbf{c} \in \Sigma^1(\mathbf{P})$ the kernel and cokernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$ are one dimensional. Define $\mathfrak{h} = \mathrm{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c}))$ and decompose the vector spaces $\mathfrak{g}$ and $\mathfrak{g}^*$ as follows

$$\mathfrak{g} \cong T_c\,\mathfrak{g} = \mathfrak{h} \oplus T_c\Sigma^1(\mathbf{P}) \quad \text{and} \quad \mathfrak{g}^* \cong T_{P(c)}\mathfrak{g}^* = T_{P(c)}\mathbf{P}(\Sigma^1(\mathbf{P})) \oplus \mathfrak{h}^*\,,$$

where $\mathbf{P}(\Sigma^1(\mathbf{P}))$ is the image of $\Sigma^1(\mathbf{P})$ in $\mathfrak{g}^*$, which is locally a hypersurface in $\mathfrak{g}^*$. Introduce the mapping on $\mathfrak{h} \times \mathfrak{h}^*$,

$$t \mapsto \mathscr{K}(\mathbf{c},t) := \langle \mathbf{P}(\mathbf{c}+t\boldsymbol{\eta}), \boldsymbol{\eta}\rangle\,, \quad \mathbf{c} \in \Sigma^1(\mathbf{P})\,, \quad \boldsymbol{\eta} \in \mathfrak{h}\,. \tag{1.3}$$

There are several identifications in this construction, and they are unravelled in §3. With the definition (1.3), $\mathscr{K}_t(\mathbf{c},0) = 0$ when $\mathbf{c} \in \Sigma^1(\mathbf{P})$. It is the curvature of the graph $(t, \mathscr{K}(\mathbf{c},t)) \in \mathfrak{h} \times \mathfrak{h}^*$ that appears in the normal form for the homoclinic bifurcation. The curvature of a graph in the plane at any t takes the standard form

$$\frac{\mathscr{K}_{tt}}{(1+\mathscr{K}_t^2)^{3/2}}\,.$$

However at points with $\mathscr{K}_t = 0$, the denominator reduces to unity, making the second derivative itself invariant under coordinate change. This observation is a special case of the *intrinsic second derivative* [2] of the mapping $\mathbf{P}$. The function $\mathscr{K}(\mathbf{c},t)$ is quadratic in t for t small when $\mathbf{c} \in \Sigma^1(\mathbf{P})$,

$$\mathscr{K}(\mathbf{c},t) = \mathscr{K}(\mathbf{c},0) + \tfrac{1}{2}t^2 \langle \mathrm{D}^2\mathbf{P}(\mathbf{c})(\boldsymbol{\eta},\boldsymbol{\eta}),\boldsymbol{\eta}\rangle + \cdots\,. \tag{1.4}$$

The quadratic term is precisely the intrinsic second derivative of $\mathbf{P}$ (cf. Porteous [30,31], Arnold et al. [2], page 149 of Golubitsky & Guillemin [10]). It is an interesting fact that $\mathbf{c} \in \Sigma^1(\mathbf{P}) \setminus \Sigma^{11}(\mathbf{P})$ if and only if the second intrinsic derivative is non-vanishing (a proof in the present context is given in §3).

[2] I am grateful to James Montaldi for pointing out the connection with the intrinsic second derivative.

The above curvature in $\mathfrak{h} \times \mathfrak{h}^*$ is a purely geometric property of a mapping between two manifolds with a $\Sigma^1(\mathbf{P})$ singularity. However, $\mathbf{P}$ can be related to the dynamics of X_H through the momentum map and φ, since $\mathbf{P} = \mathbf{J} \circ \varphi$. It is via this path that the curvature of $\mathbf{P}$ in the plane $\mathfrak{h} \times \mathfrak{h}^*$ shows up in the normal form theory.

The singularity in the mapping $\mathbf{P}$ manifests itself in the linearization of X_H by a symplectic Jordan chain. Using the symplectic Jordan chain, the linearization is transformed to normal form (a variant of Williamson normal form). Then standard normal form theory for vectorfields can be used to determine the nonlinear normal form to leading order. The normal form is a skew-product, with one part tangent to the group orbit, and the other transverse to the group orbit. The splitting between tangent and normal directions here is elementary because the group is abelian.

When the group is non-abelian the splitting of the Hamiltonian vectorfield into a component tangent to the group orbit and a component transverse to the group orbit requires the Guillemin-Sternberg-Marle theory and its generalizations (cf. Roberts, Wulff & Lamb [32] and references therein). It is reasonable to conjecture that the theory of this paper carries over to the non-abelian case by predicting the form of the transverse vectorfield near a degenerate RE in that case, but this generalization is not considered herein.

Take M to be a $(2n+2)$ dimensional vector space. (This is the lowest dimension in which the phenomena occurs; extension to higher dimension is discussed at the end of the introduction.) Apply normal form theory to X_H perturbed about a RE with $\mathbf{c} \in \Sigma^1(\mathbf{P})$. The leading-order normal form for the flow transverse to the group orbit is

$$\begin{aligned} -\tfrac{dv}{dt} &= I_1 - \tfrac{1}{2}\kappa u^2 + \cdots\,, \\ \tfrac{du}{dt} &= s_1 v + \cdots\,. \end{aligned} \tag{1.5}$$

where I_1 is an unfolding parameter which is a measure of the distance from the hypersurface $\mathbf{P}(\Sigma^1(\mathbf{P}))$ in the direction $\mathfrak{h}^*$, $s_1 = \pm 1$ is a symplectic sign, and κ is a real parameter. This normal form is the leading order normal form for a Hamiltonian system in the plane with a saddle-center bifurcation in the linearization (cf. Arnold et al. [3], Meyer & Hall [21], Broer et al. [9]). A classical formula for κ can be obtained in terms of second derivatives of the Hamiltonian vectorfield, and this formula is given in §6. Remarkably, the coefficient κ can also be precisely related to the curvature of $\mathbf{P}$ in the plane $\mathfrak{h} \times \mathfrak{h}^*$. It is proved in §7 that

$$\kappa = a_1^3 \frac{d^2}{dt^2}\bigg|_{t=0} \mathscr{K}(\mathbf{c}, t)\,, \tag{1.6}$$

where a_1 is a positive scale factor. When $n = 1$ the coefficient κ simplifies to

$$\kappa = a_1^3 P''(c)\,. \tag{1.7}$$

This role of the curvature in the one-dimensional case was first observed by Bridges & Donaldson [6] for the case of degenerate periodic orbits. Indeed, periodic orbits of Hamiltonian systems can be interpreted as a class of relative equilibria on the loop space

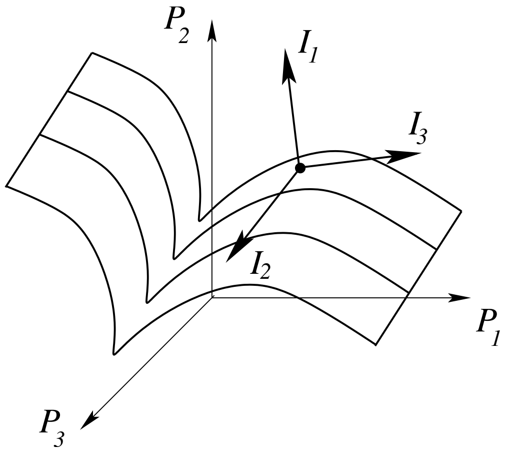

Fig. 1. Schematic of the image in $\mathfrak{g}^*$ of the hypersurface $\Sigma^1(\mathbf{P})$.

(cf. Weinstein [34], Ibort & Ontalba [12]), and so the formula (1.7) is consistent with the present theory.

The leading order normal form tangent to the group is

$$\begin{aligned} -\tfrac{dI_j}{dt} &= 0\,, \qquad\qquad j = 1, \dots, n \\ \tfrac{d\phi_1}{dt} &= u + \cdots \\ \tfrac{d\phi_j}{dt} &= s_j I_j + \cdots\,. \qquad j = 2, \dots, n\,. \end{aligned} \tag{1.8}$$

The coordinates $(I_2, \dots, I_n)$ are local coordinates for the tangent space to $\mathbf{P}(\Sigma^1(\mathbf{P}))$ in $\mathfrak{g}^*$. A schematic of the surface $\mathbf{P}(\Sigma^1(\mathbf{P}))$ for the case $n = 3$ is shown in Figure 1.

The leading order flow tangent to the group has two interesting properties. The direction of the dynamic drift along the group is determined by the symplectic signs s_j, $j = 2, \dots, n$, and these signs can be interpreted as the signs of the nonzero eigenvalues of $D\mathbf{P}(\mathbf{c})$. The second interesting feature of the flow tangent to the group is the induced holonomy. The planar system (1.5) has a homoclinic orbit. Coupling this homoclinic orbit to the normal form tangent to the group (1.8) leads to a dynamic drift along the group as well as a geometric phase shift. The dynamic drift is a familiar feature of perturbed relative equilibria (cf. Patrick [28]). The geometric phase is encoded in the equation $\frac{d\phi_1}{dt} = u + \cdots$, when u is substituted from the transverse normal form. This geometric phase can also be interpreted as a "reconstruction phase" (cf. Marsden, Montgomery & Ratiu [17]), since it appears when the reduced system (1.5) is lifted up to the full phase space. However, standard theory of reconstruction does not apply since the orbit in the reduced system is not closed. The induced geometric phase is discussed further in §8.1. A schematic of the phase shift is shown in Figure 2 for the case when the group is one-dimensional and $G = \mathbb{R}$.

A byproduct of the theory is an observation about persistence of RE: degenerate RE create barriers in $\mathfrak{g}$ or $\mathfrak{g}^*$ to persistence. In the case when $D\mathbf{P}(\mathbf{c})$ is degenerate with one-dimensional kernel, the image of $\Sigma^1(\mathbf{P})$ in $\mathfrak{g}^*$ locally divides momentum space into two regions, and locally RE exist on only one side of the surface. To see this consider the line $\mathbf{c} + t\boldsymbol{\eta}$ for $-\varepsilon < t < \varepsilon$ for some small ε with $\mathbf{c} \in \Sigma^1(\mathbf{P})$ and $\boldsymbol{\eta} \in \mathfrak{h}$. This line

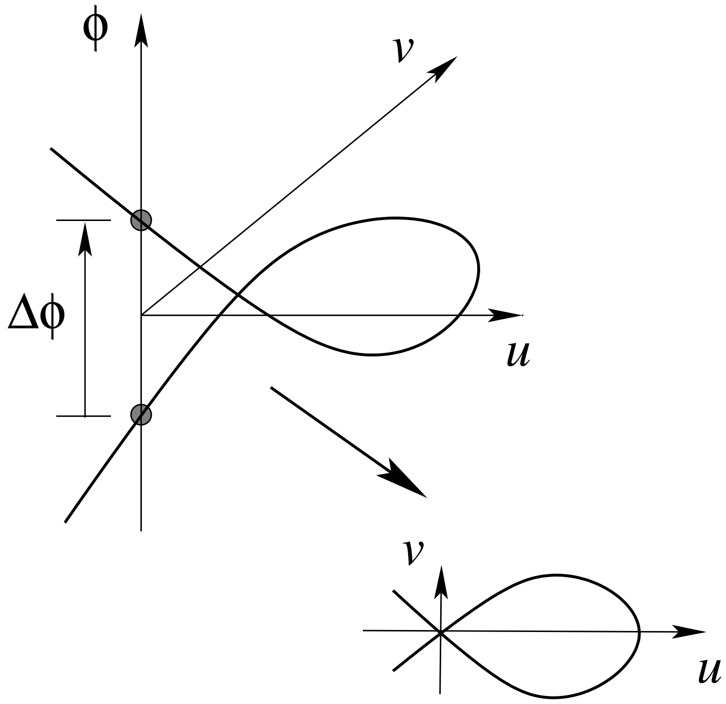

Fig. 2. Schematic of the phase shift of the orbit which is homoclinic to the RE when the orbit of the RE is a line and I_1 is fixed. The projection onto the reduced space (u,v) is also shown.

is transverse to the hypersurface $\Sigma^1(\mathbf{P})$ as long as $\mathbf{c} \notin \Sigma^{11}(\mathbf{P})$ (this is proved in §3). However, the image of this line $\mathbf{P}(\mathbf{c}+t\boldsymbol{\eta}) \in \mathfrak{g}^*$ for $-\varepsilon < t < \varepsilon$ is quadratic in t in the direction $\mathfrak{h}^*$, and hence locally the graph of a parabola. This barrier property is evident in examples. Which side of the hypersurface the RE exist is determined by the sign of κ in (1.6) and the symplectic sign s_1 (a classification is given in Figure 3). On the other hand when the G-Morse hypothesis fails, the barrier to persistence is in $\mathfrak{g}$: an example of this phenomenon is given in §9. These observations are complementary to existing results on persistence of RE (cf. Montaldi [24], Wulff [36] and references therein).

Only the leading order terms in the nonlinear normal form are considered. Global aspects of the normal form (that is, transformation to all orders) are not considered, nor is convergence of the normal forms or persistence of the homoclinic orbits in the original system. However, on these latter two points there are grounds for optimism. With the hypothesis of group dimension n and phase space dimension $2n+2$, the systems are in principle integrable and normal forms in this case are known to be extraordinarily robust [38].

The dimension of the phase space is restricted to $2n+2$ because it is the lowest dimension in which the phenomena arises. Extending the dimension brings in the usual technicalities. When the dimension of the group has dimension n and the phase space dimension is greater than $2n+2$, but the complementary dimensions are hyperbolic – even infinite dimensional – then one can use symplectic center manifold reduction (cf. Mielke [22]). With symplectic center manifold reduction, the hyperbolic directions are eliminated and one reduces to studying an ODE on $\mathbb{R}^{2n+2}$. One is again in the setting of this paper.

If the complementary dimension has an elliptic component (additional pure-imaginary eigenvalues) then formally the local normal form theory goes through but then one has well known problems with persistence of the homoclinic orbits. In this case one can expect – from related theory without symmetry [9,14] – that the homoclinic orbit will have nontrivial but exponentially-small tails.

The bifurcation associated with a degenerate RE is codimension one and therefore it should be widely observable in physical systems. The author's motivation for this theory was applications in water waves: this bifurcation arises in the analysis of wave breaking [4], is the starting point for a new branch of steady dark solitary waves in shallow water [7], and gives new results on the bifurcation of solitary waves at the interface between two fluids [8]. Some elementary examples are given in §10, §11 and §12 to illustrate application of the theory.

2 Degenerate relative equilibria

Let (M,Ω) be a symplectic manifold with the following hypothesis:

$$\text{The manifold } M \text{ is a } (2n+2)-\text{dimensional vector space.} \qquad \textbf{(H1)}$$

Let $[\cdot,\cdot] : T_zM \times T_zM \to \mathbb{R}$ be the pairing at $z \in M$. (T_zM can be identified with M but it will be useful to retain the distinction.) The symplectic form is taken to be in canonical form

$$\Omega(\mathbf{v},\mathbf{w}) = [\mathbb{J}\mathbf{v},\mathbf{w}]\,, \quad \text{for any } \mathbf{v},\mathbf{w} \in T_zM\,, \quad \text{with} \quad \mathbb{J} = \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}\,. \tag{2.1}$$

The natural setting for the analysis is Hamiltonian systems with symmetry [18]. This setting is summarized by the following hypothesis:

$$(M,\Omega,\mathrm{G},\mathbf{J},H) \text{ is a symplectic } G-\text{system.} \qquad \textbf{(H2)}$$

This hypothesis is shorthand for the following facts (cf. page 43 of [16]). G is a Lie group acting symplectically on M with Lie algebra $\mathfrak{g}$. Let $\mathfrak{g}^*$ be the dual of the Lie algebra. There exists a momentum map $\mathbf{J} : M \to \mathfrak{g}^*$ associated with G and this momentum map is Ad^*-equivariant. The function $H : M \to \mathbb{R}$ is a given smooth G$-$invariant function.

The Lie group G is restricted to be abelian:

$$\mathrm{G} = \mathbb{R}^k \times \mathbb{T}^{n-k} \text{ for some } k \text{ with } 0 \le k \le n\,. \qquad \textbf{(H3)}$$

The subgroup $\mathbb{R}^k$ is a group of affine translations in M, and $\mathbb{T}^{n-k} = S^1 \times \cdots \times S^1$ is a toral group acting on M and commuting with the translation subgroup.

The action of G on M is denoted by $\Phi_g(\mathbf{z})$, for $g \in \mathrm{G}$ and $\mathbf{z} \in M$. For any $\xi \in \mathfrak{g}$ the corresponding infinitesimal generator of the action is the vectorfield ξ_M on M defined by

$$\xi_M(\mathbf{z}) = \frac{d}{dt}\bigg|_{t=0} \Phi_{\exp(t\xi)}(\mathbf{z})\,. \tag{2.2}$$

The momentum map $\mathbf{J}: M \to \mathfrak{g}^*$ is defined by

$$\xi_M(\mathbf{z}) \,\lrcorner\, \Omega = \mathrm{d}\left(\langle \mathbf{J}(\mathbf{z}), \xi\rangle\right) \quad \text{for each} \quad \xi \in \mathfrak{g}\,. \tag{2.3}$$

Let $\{\xi_1, \ldots, \xi_n\}$ be a basis for $\mathfrak{g}$. Then the components of the momentum map are given by $J_j(\mathbf{z}) = \langle \mathbf{J}(\mathbf{z}), \xi_j\rangle$ for $j = 1, \ldots, n$, and they satisfy

$$\xi^j_M(\mathbf{z}) \,\lrcorner\, \Omega = \mathrm{d}J_j(\mathbf{z})\,, \quad j = 1, \ldots, n\,. \tag{2.4}$$

The starting point for the analysis of dynamics is the Hamiltonian system

$$\dot{\mathbf{z}} = X_H(\mathbf{z})\,, \quad \mathbf{z} \in M\,, \quad \text{with} \quad X_H \,\lrcorner\, \Omega = \mathrm{d}H\,. \tag{2.5}$$

A *relative equilibrium* of (2.5) associated with the group G is a solution of the form

$$\mathbf{z}(t) = \Phi_{\exp(t\xi)}(\varphi) \quad \text{for some} \quad \xi \in \mathfrak{g}\,,$$

where $\varphi : \mathfrak{g}^{\mathrm{RE}} \to M$ is a (group orbit of) critical point(s) of the augmented Hamiltonian (1.1). The set $\mathfrak{g}^{\mathrm{RE}}$ is the subset of $\mathfrak{g}$ for which critical points of the augmented Hamiltonian exist. In the examples in §10 and §11 $\mathfrak{g}^{\mathrm{RE}} = \mathfrak{g}$, but in the example in §12 $\mathfrak{g}^{\mathrm{RE}}$ is a proper subset of $\mathfrak{g}$. It will be sufficient for this paper that $\mathfrak{g}^{\mathrm{RE}}$ is non-empty and open.

A value $\mu \in \mathfrak{g}^*$ is regular if $\mathrm{d}\mathbf{J}(\mathbf{z})$ is surjective when $\mathbf{z} \in \mathbf{J}^{-1}(\mu)$, and it will be assumed throughout that

$$\mu \in \mathfrak{g}^* \text{ is a regular value, and } \mathfrak{g}^{\mathrm{RE}} \text{ is non-empty and open.} \tag{\textbf{H4}}$$

The pullback of the momentum map by φ induces a mapping

$$\mathbf{P} : \mathfrak{g}^{\mathrm{RE}} \to \mathfrak{g}^*\,, \quad \text{defined by} \quad \mathbf{P} = \mathbf{J} \circ \varphi\,. \tag{2.6}$$

The connection between the G-Morse hypothesis and type I degeneracy is established in the following.

Lemma 2.1. *Suppose the second variation of H_ξ evaluated at φ satisfies the G-Morse hypothesis and* Image($\mathbf{J} \circ \varphi$) *consists of regular values of the momentum map. Then the RE is non-degenerate if and only if the differential* D$\mathbf{P}$ *is surjective.*

To prove the Lemma, the terms involved need to be defined. It will be useful to introduce coordinates on $\mathfrak{g}$, although the results are independent of this choice. Take any basis for $\mathfrak{g}$ with coordinates $\mathbf{c} = (c_1, \ldots, c_n)$; that is, any $\xi \in \mathfrak{g}$ has the form $\xi = c_1\xi_1 + \cdots + c_n\xi_n$. In coordinates a critical point of the augmented Hamiltonian is denoted by $\varphi(\mathbf{c})$, the mapping $\mathbf{P}$ is denoted by $(P_1(\mathbf{c}), \ldots, P_n(\mathbf{c}))$, and the condition of type I non-degeneracy takes the form (1.2).

The G-Morse hypothesis assures that the kernel of the second variation of the augmented Hamiltonian equals the tangent space of G at the RE. The tangent space of the $\mathrm{G}-$orbit of a point $\mathbf{z}\in M$ is

$$T_z(\mathrm{G}\cdot\mathbf{z}) = \{\xi_M(\mathbf{z}) \;:\; \xi\in\mathfrak{g}\}$$

(cf. §9.3 of [18]). The second variation of the augmented Hamiltonian at φ is defined by

$$[\mathbf{L}(\mathbf{c})\mathbf{v},\mathbf{w}] := \frac{\partial^2}{\partial t_1\partial t_2} H_\xi(\varphi(\mathbf{c})+t_1\mathbf{v}+t_2\mathbf{w})\Big|_{t_1=t_2=0}\,, \quad \text{for any } \mathbf{v},\mathbf{w}\in T_\varphi M\,. \tag{2.7}$$

It follows from the invariance of the augmented Hamiltonian that $T_\varphi(\mathrm{G}\cdot\varphi)\subset\mathrm{Ker}(\mathbf{L}(\mathbf{c}))$. The G-Morse hypothesis assures equality:

$$\mathrm{Ker}(\mathbf{L}(\mathbf{c})) = T_\varphi(\mathrm{G}\cdot\varphi(\mathbf{c})) \tag{\textbf{H5}}$$

Proof (Lemma 2.1). A critical point of the augmented Hamiltonian is non-degenerate if the quadratic form $[\mathbf{L}(\mathbf{c})\mathbf{v},\mathbf{w}]$ is non-degenerate when $\mathbf{v},\mathbf{w}$ are restricted to be transverse to the group orbit and in the tangent space of $\mathbf{J}^{-1}(\mu)$. With the G-Morse hypothesis, this requirement is equivalent to

$$\det\begin{bmatrix} \mathbf{L}(\mathbf{c}) & \mathsf{Ker} & \mathrm{d}\mathbf{J}\circ\varphi \\ \mathsf{Ker}^T & \mathbf{0} & \mathbf{0} \\ (\mathrm{d}\mathbf{J}\circ\varphi)^T & \mathbf{0} & \mathbf{0}\end{bmatrix} \neq 0\,,$$

where the columns of Ker span the kernel of $\mathbf{L}(\mathbf{c})$. To avoid extraneous constants, Ker is normalized so that $\mathsf{Ker}^T\mathsf{Ker}=\mathbf{I}$. The element $\mathrm{d}\mathbf{J}\circ\varphi$ is $\mathrm{col}(\mathrm{d}J_1(\mathbf{z}),\cdots,\mathrm{d}J_n(\mathbf{z}))$ evaluated at $\mathbf{z}=\varphi$, where $\mathrm{col}(\mathbf{a},\ldots,\mathbf{b})$ is the matrix with columns $\mathbf{a},\ldots,\mathbf{b}$.

The equation

$$\mathbf{L}(\mathbf{c})\,\mathbf{W} = \mathrm{d}\mathbf{J}\circ\varphi \quad\text{and}\quad \mathsf{Ker}^T\mathbf{W} = \mathbf{0}\,,$$

is uniquely solvable for

$$\mathbf{W} = \mathrm{col}\left(\frac{\partial\varphi}{\partial c_1},\ldots,\frac{\partial\varphi}{\partial c_n}\right) + \mathsf{Ker}\,,$$

where the addition of Ker is adjusted to satisfy $\mathsf{Ker}^T\mathbf{W}=0$. This follows by differentiating the critical point equation $\mathrm{d}H\circ\varphi - \mathrm{d}\langle\mathbf{J}\circ\varphi,\xi\rangle = 0$ with respect to the coordinates $c_1,\cdots,c_n$ on $\mathfrak{g}$, and noting that $\mathrm{d}\mathbf{J}\circ\varphi$ is in the range of $\mathbf{L}(\mathbf{c})$.

Now, $\mathrm{d}\mathbf{J}\circ\varphi$ is of rank n due to the regular-value hypothesis (**H4**), and $(\mathrm{d}\mathbf{J}\circ\varphi)^T\mathbf{W} =$

$\mathrm{D}\mathbf{P}(\mathbf{c})$. Hence Proposition A.1 in Appendix A can be applied to conclude that

$$\det \begin{bmatrix} \mathbf{L}(\mathbf{c}) & \mathsf{Ker}\ \mathrm{d}\mathbf{J}\circ\varphi \\ \mathsf{Ker}^T & \mathbf{0} & \mathbf{0} \\ \mathrm{d}\mathbf{J}\circ\varphi^T & \mathbf{0} & \mathbf{0} \end{bmatrix} = (-1)^{n+1}\,\Pi\det(\mathrm{D}\mathbf{P}(\mathbf{c}))\,,$$

where Π is the product of the two non-zero eigenvalues of $\mathbf{L}(\mathbf{c})$. The G-Morse hypothesis (**H5**) assures that $\Pi \neq 0$. Hence non-degeneracy is equivalent to $\det(\mathrm{D}\mathbf{P}(\mathbf{c})) \neq 0$. ■

Corollary. *Suppose the G-Morse hypothesis* (**H5**) *is satisfied and* μ *is a regular value* (**H4**). *A RE is degenerate if and only if* $\det(\mathrm{D}\mathbf{P}(\mathbf{c})) = 0$.

In the remainder of this section, some additional properties of RE are established.

Proposition 2.2. $\mathrm{D}\mathbf{P}(\mathbf{c})$ *is a symmetric linear operator.*

Proof. Consider the pullback by $\varphi(\mathbf{c})$ of the augmented Hamiltonian to $\mathfrak{g}$,

$$\mathscr{H}(\mathbf{c}) = H_\xi \circ \varphi(\mathbf{c})\,, \quad \mathbf{c} \in \mathfrak{g}^{\mathrm{RE}}\,. \tag{2.8}$$

Then

$$\frac{\partial \mathscr{H}}{\partial c_i} = (\mathrm{d}H\circ\varphi - \langle \mathrm{d}\mathbf{J}\circ\varphi, \xi\rangle)\,\frac{\partial\varphi}{\partial c_i} - \langle \mathbf{J}\circ\varphi, \xi_i\rangle\,.$$

But the first term vanishes identically since $\varphi(\mathbf{c})$ is a critical point of the augmented Hamiltonian and the second term is $-P_i(\mathbf{c})$. Therefore

$$\frac{\partial^2 \mathscr{H}}{\partial c_i \partial c_j} = -\frac{\partial P_i}{\partial c_j}\,.$$

Symmetry of the Jacobian $\mathrm{D}\mathbf{P}(\mathbf{c})$ now follows from smoothness of the family of RE and symmetry of the second partials of $\mathscr{H}(\mathbf{c})$. ■

Since the group G is abelian, the momentum map is G$-$invariant: $\mathbf{J}\circ\Phi_g = \mathbf{J}$ for all $g \in \mathrm{G}$ and any $\mathbf{z} \in M$. There is an infinitesimal version of this property.

Proposition 2.3. *Let* ξ_i *and* ξ_j *be any two elements in the Lie algebra* $\mathfrak{g}$. *Then*

$$\Omega(\xi_M^i(\mathbf{z}), \xi_M^j(\mathbf{z})) = 0\,. \tag{2.9}$$

Proof. Invariance of the momentum map can be expressed in the form

$$J_i(\Phi_g(\mathbf{z})) = J_i(\mathbf{z}) \quad \text{for all} \quad J_i(\mathbf{z}) := \langle \mathbf{J}(\mathbf{z}), \xi_i\rangle\,,\ i = 1,\ldots,n\,,\ \text{and any } \mathbf{z} \in M\,.$$

Take $g = \exp(t\xi_j)$ and differentiate

$$0 = \frac{d}{dt}\bigg|_{t=0} J_i\left(\Phi_{\exp(t\xi_j)}(\mathbf{z})\right) = [\mathrm{d}J_i(\mathbf{z}), \xi_M^j(\mathbf{z})] = [\xi_M^i(\mathbf{z}) \lrcorner\, \Omega, \xi_M^j(\mathbf{z})] = \Omega(\xi_M^i, \xi_M^j)\,.$$

■

3 Geometry of the hypersurface $\Sigma^1(\mathbf{P})$

In this section the geometric properties of the nonlinear mapping $\mathbf{P} : \mathfrak{g}^{\mathrm{RE}} \to \mathfrak{g}^*$ are studied with the following hypothesis:

$$\mathbf{c} \in \Sigma^1(\mathbf{P})\,. \tag{\textbf{H6}}$$

This hypotheses ensures that the rank of $\mathrm{D}\mathbf{P}(\mathbf{c})$ is exactly $n-1$. Here and henceforth, $\mathbf{c} \in \Sigma^1(\mathbf{P})$ should be interpreted at $\mathbf{c} \in \Sigma^1(\mathbf{P}) \cap \mathscr{C}^{\mathrm{RE}}$ where $\mathscr{C}^{\mathrm{RE}}$ is an open subset of $\mathfrak{g}^{\mathrm{RE}}$.

Standard ideas from singularity theory of smooth nonlinear mappings between two manifolds are used [2,10]. The main issue is establishing conditions for when the kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$ is transverse to $T_c\Sigma^1(\mathbf{P})$ which is also related to when $\mathbf{c}$ is not in $\Sigma^{11}(\mathbf{P})$, where

$$\Sigma^{11}(\mathbf{P}) := \Sigma^1\left(\mathbf{P}\Big|_{\Sigma^1(\mathbf{P})}\right)\,. \tag{3.1}$$

Since $\mathfrak{g}$ and $\mathfrak{g}^*$ are vector spaces $T_c\mathfrak{g} \cong \mathfrak{g}$ and $T_{P(c)}\mathfrak{g}^* \cong \mathfrak{g}^*$. However, it will be useful to first maintain the distinction in the constructions and then to simplify via identification afterwards. Introduce a pairing on $T_{P(c)}\mathfrak{g}^*$,

$$\langle\!\langle \cdot, \cdot \rangle\!\rangle \; : T^*_{P(c)}\mathfrak{g}^* \times T_{P(c)}\mathfrak{g}^* \to \mathbb{R}\,,$$

and use the pairing $\langle \cdot, \cdot \rangle$ on $T_c\,\mathfrak{g}$. The adjoint of $\mathrm{D}\mathbf{P}(\mathbf{c})$ is then defined by

$$\langle\!\langle \mathbf{b}, \mathrm{D}\mathbf{P}(\mathbf{c})\mathbf{a} \rangle\!\rangle = \langle \mathrm{D}\mathbf{P}(\mathbf{c})^*\mathbf{b}, \mathbf{a} \rangle \quad \text{for any} \quad \mathbf{b} \in T^*_{P(c)}\mathfrak{g}^*\,, \quad \mathbf{a} \in T^*_c\mathfrak{g}\,,$$

that is $\mathrm{D}\mathbf{P}(\mathbf{c})^* : T^*_{P(c)}\mathfrak{g}^* \to T^*_c\mathfrak{g}$. The kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$ is a subspace of $T_c\,\mathfrak{g}$ and the kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})^*$ is a subspace of $T^*_{P(c)}\mathfrak{g}^*$. With the hypothesis (**H6**) these subspaces are each one dimensional. Denote these subspaces by

$$\mathfrak{h} = \mathrm{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c})) = \mathrm{span}\{\boldsymbol{\eta}\} \quad \text{and} \quad \mathrm{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c})^*) = \mathrm{span}\{\boldsymbol{\eta}^*\}\,.$$

With these preliminaries, it is clear that the function $\mathscr{K}(\mathbf{c}, t)$ in (1.3) should be defined by

$$\widehat{\mathscr{K}}(\mathbf{c}, t) := \langle\!\langle \boldsymbol{\eta}^*, \mathbf{P}(\mathbf{c} + t\boldsymbol{\eta}) \rangle\!\rangle\,, \quad \mathbf{c} \in \Sigma^1(\mathbf{P})\,.$$

The equivalence between $\widehat{\mathscr{K}}(\mathbf{c}, t)$ and $\mathscr{K}(\mathbf{c}, t)$ follows by noting that $T^*_{P(c)}\mathfrak{g}^* \cong T_c\,\mathfrak{g} \cong \mathfrak{g}$, $\mathrm{D}\mathbf{P}(\mathbf{c})$ is symmetric (Proposition 2.2) and so $\mathrm{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c})^*) \cong \mathfrak{h}$, and then transferring to the pairing on $\mathfrak{g}$. Henceforth, $\mathscr{K}(\mathbf{c}, t)$ will be used with the above identifications understood.

Transversality of $\mathfrak{h}$ and $T_c\Sigma^1(\mathbf{P})$, local smoothness of the hypersurface $\Sigma^1(\mathbf{P})$ and membership of $\mathbf{c}$ in $\Sigma^1(\mathbf{P}) \setminus \Sigma^{11}(\mathbf{P})$ are all related to nontriviality of the intrinsic second derivative

$$\frac{d^2}{dt^2}\bigg|_{t=0} \mathscr{K}(\mathbf{c},t) \neq 0\,, \quad \mathbf{c} \in \Sigma^1(\mathbf{P})\,. \tag{\textbf{H7}}$$

Lemma 3.1. *Suppose* $\mathbf{c} \in \Sigma^1(\mathbf{P})$*, and assume* $\mathbf{P} : \mathfrak{g} \to \mathfrak{g}^*$ *satisfies hypotheses* (**H6**) *and* (**H7**)*. Then the kernel of* $\mathrm{D}\mathbf{P}(\mathbf{c})$ *is transverse to the tangent space of* $\Sigma^1(\mathbf{P})$ *and*

(1) $\Sigma^1(\mathbf{P})$ *is a locally smooth submanifold of* $\mathfrak{g}$*;*
(2) $\mathfrak{g} \cong T_c\,\mathfrak{g} = \mathfrak{h} \oplus T_c\Sigma^1(\mathbf{P})$*;*
(3) $\mathbf{c} \notin \Sigma^{11}(\mathbf{P})$*.*

Proof. Let $f(\mathbf{c}) := \det(\mathrm{D}\mathbf{P}(\mathbf{c}))$. Then $\Sigma^1(\mathbf{P}) = \{\, \mathbf{c} \in \mathfrak{g} \;:\; \mathbf{c} \in f^{-1}(0)\,\}$ and this hypersurface defines a smooth submanifold of $\mathfrak{g}$ in the neighbourhood of any point where $\mathrm{d}f$ is nontrivial. Take $\boldsymbol{\eta} \in \mathfrak{h}$ and consider

$$\langle \mathrm{d}f, \boldsymbol{\eta}\rangle = \frac{d}{dt}\bigg|_{t=0} \det\left(\mathrm{D}\mathbf{P}(\mathbf{c}+t\boldsymbol{\eta})\right) = \mathrm{Tr}\left(\mathrm{D}\mathbf{P}(\mathbf{c})^{\#}\frac{d}{dt}\bigg|_{t=0}\mathrm{D}\mathbf{P}(\mathbf{c}+t\boldsymbol{\eta})\right)\,.$$

where $\mathrm{D}\mathbf{P}(\mathbf{c})^{\#}$ is the adjugate of $\mathrm{D}\mathbf{P}(\mathbf{c})$. Since $\mathrm{D}\mathbf{P}(\mathbf{c})$ is a symmetric matrix with a simple zero eigenvalue, its adjugate has the following explicit expression (see formula (A-2) in Appendix A),

$$\mathrm{D}\mathbf{P}(\mathbf{c})^{\#} = \Pi\,\frac{\boldsymbol{\eta}\boldsymbol{\eta}^T}{\boldsymbol{\eta}^T\boldsymbol{\eta}}\,,$$

where Π is the product of the $n-1$ nonzero eigenvalues of $\mathrm{D}\mathbf{P}(\mathbf{c})$, and so

$$\begin{aligned}
\langle \mathrm{d}f, \boldsymbol{\eta}\rangle &= \mathrm{Tr}\left(\mathrm{D}\mathbf{P}(\mathbf{c})^{\#}\tfrac{d}{dt}\Big|_{t=0}\mathrm{D}\mathbf{P}(\mathbf{c}+t\boldsymbol{\eta})\right)\\
&= \tfrac{\Pi}{\boldsymbol{\eta}^T\boldsymbol{\eta}}\mathrm{Tr}\left(\boldsymbol{\eta}\boldsymbol{\eta}^T \textstyle\sum_{k=1}^n \tfrac{\partial}{\partial c_k}\mathrm{D}\mathbf{P}(\mathbf{c})\eta_k\right)\\
&= \tfrac{\Pi}{\boldsymbol{\eta}^T\boldsymbol{\eta}}\boldsymbol{\eta}^T\left(\textstyle\sum_{k=1}^n \tfrac{\partial}{\partial c_k}\mathrm{D}\mathbf{P}(\mathbf{c})\eta_k\right)\boldsymbol{\eta}\\
&= \tfrac{\Pi}{\boldsymbol{\eta}^T\boldsymbol{\eta}}\textstyle\sum_{i,j,k}\tfrac{\partial}{\partial c_k}\left(\tfrac{\partial P_i}{\partial c_j}\right)\eta_i\eta_j\eta_k\\
&= \tfrac{\Pi}{\boldsymbol{\eta}^T\boldsymbol{\eta}}\langle \mathrm{D}^2\mathbf{P}(\mathbf{c})(\boldsymbol{\eta},\boldsymbol{\eta}),\boldsymbol{\eta}\rangle\,,
\end{aligned}$$

using $\mathrm{D}\mathbf{P}(\mathbf{c})_{i,j} = \frac{\partial P_i}{\partial c_j}$. Differentiating $\mathscr{K}(\mathbf{c},t)$ in (1.3),

$$\frac{d^2}{dt^2}\bigg|_{t=0}\mathscr{K}(\mathbf{c},t) = \langle \mathrm{D}^2\mathbf{P}(\mathbf{c})(\boldsymbol{\eta},\boldsymbol{\eta}),\boldsymbol{\eta}\rangle\,,$$

and this expression is non-vanishing by hypothesis (**H7**). Combining this expression with the above calculation shows that

$$\langle \mathrm{d}f, \boldsymbol{\eta}\rangle = \frac{\Pi}{\boldsymbol{\eta}^T\boldsymbol{\eta}}\frac{d^2}{dt^2}\bigg|_{t=0}\mathscr{K}(\mathbf{c},t)\,.$$

But Π is non-zero since $\mathrm{D}\mathbf{P}(\mathbf{c})$ has rank $n-1$ by hypothesis (**H6**). It follows that $\mathrm{d}f$ is non-vanishing and so $\Sigma^1(\mathbf{P})$ is a smooth hypersurface in the neighbourhood of any point $\mathbf{c}$ satisfying (**H7**). This proves point 1.

The tangent space $T_c\Sigma^1(\mathbf{P})$ is the tangent space to $f^{-1}(0)$, and every element in the tangent space of $f^{-1}(0)$ is in the kernel of $\mathrm{d}f$. But by point 1, $\boldsymbol{\eta}$ is not in the kernel of $\mathrm{d}f$ and so the kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$ is transverse to $T_c\Sigma^1(\mathbf{P})$. Standard results from singularity theory establish that the set $\Sigma^{11}(\mathbf{P})$ is distinguished by the fact that the kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$ is tangent to $\Sigma^1(\mathbf{P})$ [2,10]. Point 3 then follows from the definition of $\Sigma^{11}(\mathbf{P})$ in (3.1). ∎

The Lemma combined with standard results from linear operator theory prove the following.

Corollary. *With the above hypotheses, there exists a one-dimensional subspace $\mathfrak{h}^*$ such that*

$$\mathfrak{g}^* \cong T_{P(c)}\mathfrak{g}^* = T_{P(c)}\mathbf{P}(\Sigma^1(\mathbf{P})) \oplus \mathfrak{h}^* .$$

Remark. If a Euclidean metric is introduced in $\mathfrak{g}^*$, then $\mathfrak{h}^*$ can be identified with $\mathrm{span}\{\boldsymbol{\eta}\}$ and $\boldsymbol{\eta}$ defines a normal vector at each point on the hypersurface $\mathbf{P}(\Sigma^1(\mathbf{P}))$.

4 Degenerate RE and symplectic Jordan chain theory

In this section the algebraic implications of degenerate RE are considered. It follows from the symmetry properties of RE that the linearization has a zero eigenvalue of geometric multiplicity (at least) n and algebraic multiplicity (at least) $2n$.

Proposition 4.1. *There exists $2n$ vectors $\mathbf{v}_j$ $j=1,\ldots,2n$ such that*

$$\mathbf{L}(\mathbf{c})\mathbf{v}_j = 0\,, \quad \mathbf{L}(\mathbf{c})\mathbf{v}_{n+j} = \mathbf{v}_j \lrcorner\, \Omega\,, \quad j=1,\ldots,n\,. \tag{4.1}$$

Proof. There is a group orbit of critical points of the augmented Hamiltonian,

$$\mathrm{d}H(\Phi_g(\varphi(\mathbf{c}))) - \mathrm{d}\langle \mathbf{J}(\Phi_g(\varphi(\mathbf{c}))), \xi\rangle = 0\,, \quad \text{for all} \quad g\in \mathrm{G}\,. \tag{4.2}$$

Take $g=\exp(t_1\xi_1+\cdots+t_n\xi_n)$ where $\{\xi_1,\ldots,\xi_n\}$ is a basis for $\mathfrak{g}$. Then derivatives of (4.2) with respect to each t_j and c_j lead to

$$\begin{aligned}(\mathrm{D}^2H(\Phi_g(\varphi(\mathbf{c}))) - \mathrm{D}^2\langle \mathbf{J}(\Phi_g(\varphi(\mathbf{c}))), \xi\rangle)\, \xi^j_M(\Phi_g(\varphi(\mathbf{c}))) &= 0\\ (\mathrm{D}^2H(\Phi_g(\varphi(\mathbf{c}))) - \mathrm{D}^2\langle \mathbf{J}(\Phi_g(\varphi(\mathbf{c}))), \xi\rangle)\, \Phi_g\left(\tfrac{\partial\varphi}{\partial c_j}\right) &= \mathrm{d}J_j(\Phi_g(\varphi))\,.\end{aligned} \tag{4.3}$$

Evaluating at the identity, $t_1 = \cdots = t_n = 0$,

$$\begin{aligned} \mathbf{L}(\mathbf{c})\, \xi_M^j(\varphi(\mathbf{c})) &= 0 \\ \mathbf{L}(\mathbf{c}) \left(\tfrac{\partial \varphi}{\partial c_j}\right) &= \mathrm{d}J_j(\varphi(\mathbf{c})) = \xi_M^j(\varphi(\mathbf{c})) \,\lrcorner\, \Omega\,. \end{aligned}$$

The result now follows by taking

$$\mathbf{v}_j = \xi_M^j(\varphi(\mathbf{c})) \quad \text{and} \quad \mathbf{v}_{n+j} = \frac{\partial \varphi(\mathbf{c})}{\partial c_j}\,, \quad j = 1, \ldots, n\,.$$

■

With the hypothesis that the geometric multiplicity is exactly n (the G-Morse hypothesis), the Jordan chain terminates at $2n$ precisely when the RE is non-degenerate. The connection between degeneracy of the mapping $\mathbf{P}$ and symplectic Jordan chain theory is now established.

Proposition 4.2. *When* $\mathbf{c} \in \Sigma^1(\mathbf{P})$

$$\sum_{j=1}^n \eta_j \mathbf{v}_{n+j} \,\lrcorner\, \Omega \in \ \text{Range}(\mathbf{L}(\mathbf{c}))\,, \quad \text{where} \quad \boldsymbol{\eta} \in \text{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c}))\,,$$

and the Jordan chain continues to length $2n+2$*: there exists vectors* $\mathbf{v}_{2n+1}$ *and* $\mathbf{v}_{2n+2}$*,*

$$\mathbf{L}(\mathbf{c})\mathbf{v}_{2n+1} = \sum_{j=1}^n \eta_j \mathbf{v}_{n+j} \,\lrcorner\, \Omega \quad \text{and} \quad \mathbf{L}(\mathbf{c})\mathbf{v}_{2n+2} = \mathbf{v}_{2n+1} \,\lrcorner\, \Omega\,. \tag{4.4}$$

Proof. The first equation in (4.4) is solvable if and only if the right-hand side is in the range of $\mathbf{L}(\mathbf{c})$; that is, if

$$\sum_{j=1}^n \eta_j \Omega(\mathbf{v}_i, \mathbf{v}_{n+j}) = 0\,, \quad \text{for each} \quad i = 1, \ldots, n\,. \tag{4.5}$$

The hypothesis $\mathrm{D}\mathbf{P}(\mathbf{c})\boldsymbol{\eta} = 0$ implies, for each $i = 1, \ldots, n$, that

$$\begin{aligned} 0 = \textstyle\sum_{j=1}^n \eta_j \frac{\partial P_i}{\partial c_j} &= \textstyle\sum_{j=1}^n \eta_j \left[\mathrm{d}\langle \mathbf{J}(\varphi), \xi_i\rangle, \frac{\partial \varphi}{\partial c_j}\right] \\ &= \textstyle\sum_{j=1}^n \eta_j \left[\boldsymbol{\xi}_M^i(\varphi) \,\lrcorner\, \Omega, \frac{\partial \varphi}{\partial c_j}\right] \\ &= \textstyle\sum_{j=1}^n \eta_j \left[\mathbf{v}_i \,\lrcorner\, \Omega, \mathbf{v}_{j+n}\right] \\ &= \textstyle\sum_{j=1}^n \eta_j \Omega(\mathbf{v}_i, \mathbf{v}_{j+n})\,, \end{aligned}$$

confirming (4.5). The algebraic multiplicity of zero is always even for a symplectic linearization. Hence the algebraic multiplicity is $2n+2$ and it follows that the second equation of (4.4) is solvable. ■

The condition for termination of the Jordan chain is that

$$\mathbf{L}(\mathbf{c})\mathbf{v}_{2n+3} = \mathbf{v}_{2n+2} \lrcorner\, \Omega\,,$$

is *not solvable*: that is, there exists a constant $n-$vector $\boldsymbol{\beta}$ such that

$$\sum_{j=1}^{n} \beta_j \Omega(\mathbf{v}_j, \mathbf{v}_{2n+2}) \neq 0\,.$$

In the present case, the restriction of the dimension of M to $(2n+2)$ makes this condition irrelevant. However, a form of this condition arises in the normal form theory with $\boldsymbol{\beta} = \boldsymbol{\eta}$.

Proposition 4.3. *For* $i = 1, \ldots, n$ *and* $j = 1, \ldots, n$*,*

$$\Omega(\mathbf{v}_i, \mathbf{v}_j) = 0\,, \quad \Omega(\mathbf{v}_i, \mathbf{v}_{n+j}) = \frac{\partial P_i}{\partial c_j} \quad \text{and} \quad \Omega(\mathbf{v}_i, \mathbf{v}_{2n+1}) = 0\,.$$

Proof. $\Omega(\mathbf{v}_i, \mathbf{v}_j) = \Omega(\xi^i_M(\varphi), \xi^j_M(\varphi))$ which vanishes due to invariance of the momentum map (Proposition 2.3). For the second equation,

$$\frac{\partial P_i}{\partial c_j} = \left[\mathrm{d}J_i(\varphi), \frac{\partial \varphi}{\partial c_j} \right] = \left[\xi^i_M(\varphi) \lrcorner\, \Omega, \mathbf{v}_{n+j} \right] = \Omega(\mathbf{v}_i, \mathbf{v}_{n+j})\,.$$

For the third equation

$$\Omega(\mathbf{v}_i, \mathbf{v}_{2n+1}) = -\left[\mathbf{v}_{2n+1} \lrcorner\, \Omega, \mathbf{v}_i \right] = -\left[\mathbf{L}(\mathbf{c})\mathbf{v}_{2n+2}, \mathbf{v}_i \right] = -\left[\mathbf{L}(\mathbf{c})\mathbf{v}_i, \mathbf{v}_{2n+2} \right] = 0\,,$$

using Proposition 4.2, the fact that each $\mathbf{v}_i \in \mathrm{Ker}(\mathbf{L}(\mathbf{c}))$ and the symmetry of $\mathbf{L}(\mathbf{c})$. ■

Corollary. *Suppose* $\mathbf{c} \in \Sigma^1(\mathbf{P})$ *and let* $\mathbf{w}_1 = \eta_1 \mathbf{v}_1 + \cdots + \eta_n \mathbf{v}_n$ *where* $\boldsymbol{\eta} \in \mathrm{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c}))$*. Then*

$$\Omega(\mathbf{v}_{2n+2}, \mathbf{w}_1) \neq 0\,. \tag{4.6}$$

Proof. The proof proceeds by showing that

$$\mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{2n+1}\} \subset \mathrm{Ker}\left(\mathbf{w}_1 \lrcorner\, \Omega \right)\,. \tag{4.7}$$

Then using the fact that the dimension of $\mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{2n+2}\} = 2n+2$ and nontriviality of $\mathbf{w}_1$ (which follows from linear independence of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and nontriviality of $\boldsymbol{\eta}$) the result is proved. The fact that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is in the kernel of $\mathbf{w}_1 \lrcorner\, \Omega$ follows from invariance of the momentum map (2.9),

$$[\mathbf{w}_1 \lrcorner\, \Omega, \mathbf{v}_k] = \sum_{j=1}^{n} \eta_j [\mathbf{v}_j \lrcorner\, \Omega, \mathbf{v}_k] = \sum_{j=1}^{n} \Omega(\mathbf{v}_j, \mathbf{v}_k) = 0\,, \quad \text{for } k = 1, \ldots, n\,,$$

using the first part of Proposition 4.3. For $k = n+1, \ldots, 2n$,

$$[\mathbf{w}_1 \lrcorner\, \Omega, \mathbf{v}_k] = \sum_{j=1}^{n} \eta_j [\mathbf{v}_j \lrcorner\, \Omega, \mathbf{v}_k] = \sum_{j=1}^{n} \eta_j \Omega(\mathbf{v}_j, \mathbf{v}_k) = \sum_{j=1}^{n} \eta_j \frac{\partial P_j}{\partial c_k} = 0\,,$$

using the second part of Proposition 4.3 and the fact that $\boldsymbol{\eta} \in \mathrm{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c}))$. For $k = 2n+1$ apply the third part of Proposition 4.3

$$[\mathbf{w}_1 \lrcorner\, \Omega, \mathbf{v}_{2n+1}] = \sum_{j=1}^{n} \eta_j [\mathbf{v}_j \lrcorner\, \Omega, \mathbf{v}_{2n+1}] = \sum_{j=1}^{n} \eta_j \Omega(\mathbf{v}_j, \mathbf{v}_{2n+1}) = 0\,.$$

This completes the proof. ■

5 Normal form for the linearization about degenerate RE

The linearization of the Hamiltonian system (2.5) about a degenerate RE takes the form

$$\mathbb{J}\mathbf{z}_t = \mathbf{L}(\mathbf{c})\,\mathbf{z}\,,$$

where $\mathbf{L}(\mathbf{c})$ has a Jordan chain of length $(2n+2)$ defined by (4.1) and (4.4). The theory for the transformation of the pair $(\mathbb{J}, \mathbf{L}(\mathbf{c}))$ to symplectic Jordan normal form is classical (WILLIAMSON [35]; see also DELLNITZ & MELBOURNE [19] for the equivariant case), and the following Lemma is a special case of this theory. However, there are a few interesting observations to be made, and in order to develop the nonlinear normal form near a degenerate RE some precise properties of the linear normal form transformations need to be established.

Lemma 5.1. *Suppose that* $(\mathbb{J}, \mathbf{L}(\mathbf{c}))$ *has the Jordan structure (4.1)-(4.4). Then there exists a transformation* $\mathbf{F}$ *such that*

$$\mathbf{F}^T \mathbf{L}(\mathbf{c}) \mathbf{F} = \mathbf{L}^{\mathrm{ref}} \quad \text{and} \quad \mathbf{F}^T\, \mathbb{J}\, \mathbf{F} = \mathbb{J}$$

where

$$\mathbf{L}^{\mathrm{ref}} = \left[\begin{array}{cccccccccc}
0 & \cdot & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & & & \vdots & \vdots & & & & \vdots \\
\vdots & & & \vdots & 0 & & & & \vdots \\
0 & \cdots & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\
0 & \cdots & 0 & 1 & 0 & 0 & \cdots & \cdots & 0 \\
\vdots & & & 0 & 0 & s_2 & & & \vdots \\
\vdots & & & \vdots & \vdots & & \ddots & & \vdots \\
\vdots & & & \vdots & \vdots & & & s_n & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 & s_1
\end{array}\right]\,. \tag{5.1}$$

$s_1, \ldots, s_n$ are symplectic invariants and take values ± 1, with the values determined by

$$s_1 = \operatorname{sign} \Omega(\mathbf{v}_{2n+2}, \mathbf{w}_1)\,, \tag{5.2}$$

and s_j for $j = 2, \ldots, n$ are the signs of the non-zero eigenvalues of $\mathrm{D}\mathbf{P}(\mathbf{c})$ *(with any chosen ordering of the non-zero eigenvalues).*

An explicit proof of this result is given in Appendix B. The sign s_1 is familiar in symplectic Jordan theory [35], it is the sign of an appropriate projection onto the top of the Jordan chain $\mathbf{v}_{2n+2}$. The other $(n-1)$ signs, which appear in the Williamson theory in a purely algebraic way, can be given a geometric interpretation in the present context. The eigenvalues of $\mathrm{D}\mathbf{P}(\mathbf{c})$ are related to the curvature of the pullback of the augmented Hamiltonian $\mathscr{H} : \mathfrak{g} \to \mathbb{R}$ defined in (2.8), since

$$\mathrm{D}^2\mathscr{H}(\mathbf{c}) = -\mathrm{D}\mathbf{P}(\mathbf{c})\,.$$

The signs of the eigenvalues of $\mathrm{D}\mathbf{P}(\mathbf{c})$ are therefore equal to minus the signs of the curvatures of the graph of $\mathscr{H}(\mathbf{c})$. The signs $s_2, \ldots, s_n$ show up in the dynamics by determining the direction of dynamic drift along the group associated with perturbations about the family of degenerate RE.

A sketch of the properties of the new symplectic basis are recorded here. The details of their construction are given in Appendix B. Introduce new symplectic coordinates

$$\widehat{\mathbf{z}} = (\phi_1, \ldots, \phi_n, u, I_1, \ldots, I_n, v)\,,$$

defined by $\mathbf{z}(t) = \mathbf{F}\widehat{\mathbf{z}}(t)$. Or, using the definition of $\mathbf{F}$ from Appendix B

$$\mathbf{F} = \mathrm{col}\left(\widehat{\mathbf{w}}_1, \cdots, \widehat{\mathbf{w}}_{n+1}, -s_1\widehat{\mathbf{w}}_{2n+2}, s_2\widehat{\mathbf{w}}_{n+2}, \cdots, s_n\widehat{\mathbf{w}}_{2n}, s_1\widehat{\mathbf{w}}_{2n+1}\right)\,,$$

the new coordinates are defined by

$$\begin{aligned}\mathbf{z}(t) &= \phi_1(t)\widehat{\mathbf{w}}_1 + \cdots + \phi_n(t)\widehat{\mathbf{w}}_n + u(t)\widehat{\mathbf{w}}_{n+1} \\ &\qquad -s_1 I_1(t)\widehat{\mathbf{w}}_{2n+2} + s_2 I_2(t)\widehat{\mathbf{w}}_{n+2} + \cdots + s_n I_n(t)\widehat{\mathbf{w}}_{2n} + s_1 v(t)\widehat{\mathbf{w}}_{2n+1}\,.\end{aligned} \tag{5.3}$$

The Jordan chain associated with the new symplectic basis is

$$\begin{aligned}&\mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_j = 0\,, \quad \mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{n+j} = \widehat{\mathbf{w}}_j \,\lrcorner\, \Omega\,, \quad j = 1, \ldots, n\,, \\ &\mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{2n+1} = \widehat{\mathbf{w}}_{n+1} \,\lrcorner\, \Omega \quad \text{and} \quad \mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{2n+2} = \widehat{\mathbf{w}}_{2n+1} \,\lrcorner\, \Omega\,.\end{aligned} \tag{5.4}$$

The following property of the new symplectic basis will be needed in the analysis of the nonlinear normal form.

Proposition 5.2.

$$\Omega(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+1}) = -\Omega(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+2}) = s_1\,.$$

Proof. For the first equality,

$$
\begin{aligned}
\Omega(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+1}) &= -[\widehat{\mathbf{w}}_{2n+1} \lrcorner\, \Omega, \widehat{\mathbf{w}}_{n+1}] \\
&= -[\mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{2n+2}, \widehat{\mathbf{w}}_{n+1}] \quad \text{(using Jordan chain (5.4))} \\
&= -[\mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+2}] \quad \text{(symmetry of } \mathbf{L}(\mathbf{c})\text{)} \\
&= -[\widehat{\mathbf{w}}_1 \lrcorner\, \Omega, \widehat{\mathbf{w}}_{2n+2}] \quad \text{(using Jordan chain (5.4))} \\
&= -\Omega(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+2})
\end{aligned}
$$

For the second equality, use the definition of s_1 in (5.2) and the explicit form for $\widehat{\mathbf{w}}_{2n+2}$ given in equation (B-7) in the Appendix,

$$
\begin{aligned}
-\Omega(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+2}) &= -\Omega(a_1\mathbf{w}_1, a_1\mathbf{w}_{2n+2} + \textstyle\sum_{k=1}^n m_k \mathbf{w}_{n+k}) \\
&= -a_1^2\Omega(\mathbf{w}_1, \mathbf{w}_{2n+2}) - a_1 \textstyle\sum_{k=1}^n m_k \Omega(\mathbf{w}_1, \mathbf{w}_{n+k}) \\
&= s_1 - a_1 \textstyle\sum_{k=1}^n m_k \Omega(\mathbf{w}_1, \mathbf{w}_{n+k})\,,
\end{aligned}
$$

using the definition of a_1 in (B-4). The second term vanishes because $\mathbf{w}_{n+k}$ is a linear combination of $\mathbf{v}_{n+1}, \ldots, \mathbf{v}_{2n}$ and

$$
\Omega(\mathbf{w}_1, \mathbf{v}_{n+k}) = \sum_{\ell=1}^n \eta_\ell\, \Omega(\mathbf{v}_\ell, \mathbf{v}_{n+k}) = \sum_{\ell=1}^n \eta_\ell \frac{\partial P_\ell}{\partial c_k} = 0\,,
$$

for each $k = 1, \ldots, n$ using Proposition 4.3 and the fact that $\boldsymbol{\eta} \in \mathrm{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c}))$. ■

6 Nonlinear normal form near a degenerate RE

Perturb the Hamiltonian system (2.5) about the family of RE. Taking advantage of the vector space structure of M,

$$
\mathbf{z}(t) = \Phi_{g(t)}(\varphi(\mathbf{c}) + V(\mathbf{c}, t))\,, \quad g(t) = \exp(t\xi)\,.
$$

Substuting into (2.5)

$$
\Phi_{g(t)} V_t + \xi_M(\Phi_{g(t)}(\varphi + V)) = X_H(\Phi_{g(t)}(\varphi + V))\,.
$$

Take the interior product with Ω,

$$
\Phi_{g(t)} V_t \lrcorner\, \Omega = \mathrm{d}H(\Phi_{g(t)}(\varphi + V)) - \mathrm{d}\langle \mathbf{J}(\Phi_{g(t)}(\varphi + V)), \xi\rangle\,.
$$

Use the invariance of H and the momentum map, expand the right-hand side in a Taylor series about φ and use the relation $V_t \lrcorner\, \Omega = \mathbb{J}V_t$ to obtain the leading order system for the perturbation about the family of RE

$$
\mathbb{J}V_t = \mathbf{L}(\mathbf{c})V + \tfrac{1}{2}\mathrm{D}^3 H(V, V) + \cdots\,, \tag{6.1}
$$

where the third derivative $\mathrm{D}^3 H$ is evaluated at the point $\varphi(\mathbf{c})$ and defined by

$$\mathrm{D}^3 H(\mathbf{a}, \mathbf{b}) := \left.\frac{\partial^2}{\partial t_1 \partial t_2}\right|_{t_1=t_2=0} \mathrm{d}H(\varphi(\mathbf{c}) + t_1 \mathbf{a} + t_2 \mathbf{b})\,.$$

To leading order use the linear transformation (5.3) This results in the linear Hamiltonian system in the new coordinates

$$\begin{aligned}
-\tfrac{dI_j}{dt} &= 0\,, && j = 1, \dots, n \\
-\tfrac{dv}{dt} &= I_1\,, \\
\tfrac{d\phi_1}{dt} &= u \\
\tfrac{d\phi_j}{dt} &= s_j I_j\,, && j = 2, \dots, n \\
\tfrac{du}{dt} &= s_1 v\,.
\end{aligned} \tag{6.2}$$

The transverse system

$$\begin{aligned}
-v_t &= I_1 \\
u_t &= s_1 v\,,
\end{aligned}$$

is in standard form for a saddle-center bifurcation in the linearization, and the nonlinear normal form to leading order takes the well-known form

$$\begin{aligned}
-v_t &= I_1 - \tfrac{1}{2}\kappa u^2 + \cdots \\
u_t &= s_1 v + \cdots\,,
\end{aligned} \tag{6.3}$$

(cf. pages 263-264 of Arnold, Koslov & Neistadt [3], page 188 of Meyer & Hall [21]). The main result of this paper is summarized in the following.

Theorem 6.1. *Consider a Hamiltonian system on a symplectic manifold (M, Ω) with a family of relative equilibria satisfying the hypotheses* (**H1**) *to* (**H7**). *Take $\mathbf{c} \in \Sigma^1(\mathbf{P}) \cap \mathscr{C}^{\mathrm{RE}}$ fixed. Then there exists a symplectic transformation with new coordinates*

$$(\phi_1, \dots, \phi_n, u, I_1, \dots, I_n, v)\,,$$

such that the leading order terms in the Hamiltonian vectorfield take the form

$$X_H = u\frac{\partial}{\partial \phi_1} + s_2 I_2 \frac{\partial}{\partial \phi_2} + \cdots + s_n I_n \frac{\partial}{\partial \phi_n} + s_1 v \frac{\partial}{\partial u} - (I_1 - \tfrac{1}{2}\kappa u^2)\frac{\partial}{\partial v} + \cdots\,, \tag{6.4}$$

with I_1 the component of the momentum map at $\mathbf{P}(\mathbf{c})$ in the $\mathfrak{h}^*$ direction, κ given by the formula (1.6), and the symplectic signs s_j defined in Lemma 5.1.

The flow tangent to the group follows from the linear normal form theory in Lemma 5.1, and the flow transverse to the group follows from (6.3). The property of I_1 is established in Proposition 6.3 below. The proof of the formula (1.6) will be split into two steps. First classical normal form theory is used to determine a formula for κ in terms of derivatives of the Hamiltonian functional (Lemma 6.2). Then the relation with the curvature (1.6) is established in Lemma 7.3.

Lemma 6.2.

$$\kappa = -[\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{n+1}] - 3[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_1), \widehat{\mathbf{w}}_{2n+2}] + 3[\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+1}), \widehat{\mathbf{w}}_1]\,.$$

Proof. To determine a formula for the coefficient κ in the nonlinear normal form, the strategy proposed in IOOSS & ADELMEYER [13] for computing normal form coefficients is used. The idea is to introduce a polynomial approximation in the new coordinates, substitute into the perturbed ODE and equate terms proportional to like powers to zero.

Expand V in a Taylor series in $(\phi_1,\ldots,\phi_n,u,I_1,\ldots,I_n,v)$ to the order desired and substitute into the governing equations (6.1). The key step is then to replace derivatives of the coordinates $(\dot\phi_1,\ldots,,\dot v)$ by their normal form expressions in (1.5) and (1.8). This strategy results is a large system of linear equations at each order, and solvability leads to conditions for the existence of the coefficients.

Let

$$V(t) = V_1(t) + V_2(t) + \cdots\,, \tag{6.5}$$

where each V_j is a homogeneous power series of degree j in the coordinates $(\phi_1,\ldots,v)$ with V_1 given in (5.3) and

$$V_2(t) = \phi_1(t)^2\Upsilon_1 + \phi_1(t)u(t)\Upsilon_2 + \phi_1(t)I_1(t)\Upsilon_3 + \phi_1(t)v(t)\Upsilon_4 + u(t)^2\Upsilon_5 + u(t)v(t)\Upsilon_6 + \cdots$$

where $\Upsilon_1,\ldots,\Upsilon_6$ are $t-$independent vectors that are determined as part of the calculation. There are $\frac{1}{2}n(n+1)$ terms in V_2, but only the six terms listed above are needed to compute κ. It is remarkable that the number of terms in V_2 needed to determine κ is independent of n!

Substituting the expressions for V_1 and V_2 into (6.1) results in the following coupled system of linear equations

$$\begin{aligned}
\phi_1^2 &\;:\; 0 &&= \mathbf{L}(\mathbf{c})\Upsilon_1 + \tfrac{1}{2}\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_1)\\
\phi_1 u &\;:\; 2\mathbb{J}\Upsilon_1 &&= \mathbf{L}(\mathbf{c})\Upsilon_2 + \mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{n+1})\\
\phi_1 I_1 &\;:\; -\mathbb{J}\Upsilon_4 &&= \mathbf{L}(\mathbf{c})\Upsilon_3 - s_1\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+2})\\
\phi_1 v &\;:\; s_1\mathbb{J}\Upsilon_2 &&= \mathbf{L}(\mathbf{c})\Upsilon_4 + s_1\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+1})\\
u^2 &\;:\; \tfrac{1}{2}s_1\kappa\mathbb{J}\widehat{\mathbf{w}}_{2n+1} + \mathbb{J}\Upsilon_2 &&= \mathbf{L}(\mathbf{c})\Upsilon_5 + \tfrac{1}{2}\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{n+1})\\
uv &\;:\; \mathbb{J}\Upsilon_4 + 2s_1\mathbb{J}\Upsilon_5 &&= \mathbf{L}(\mathbf{c})\Upsilon_6 + s_1\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+1})\,.
\end{aligned} \tag{6.6}$$

Add the third and sixth equations

$$\mathbf{L}(\mathbf{c})(\Upsilon_3 + \Upsilon_6) = 2s_1 \mathbb{J}\Upsilon_5 + s_1 \mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+2}) - s_1 \mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+1})\,.$$

This equation is solvable for $(\Upsilon_3 + \Upsilon_6)$ if and only if the right hand side is in the range of $\mathbf{L}(\mathbf{c})$

$$s_1[2\mathbb{J}\Upsilon_5 + \mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+2}) - \mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+1}), \widehat{\mathbf{w}}_j] = 0\,, \quad j = 1, \ldots, n\,.$$

The coefficient κ drops out of all these equations except the case $j = 1$: only the first equation is needed to determine κ,

$$2[\mathbb{J}\Upsilon_5, \widehat{\mathbf{w}}_1] = -[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+2}), \widehat{\mathbf{w}}_1] + [\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+1}), \widehat{\mathbf{w}}_1]\,. \tag{6.7}$$

The left-hand side can be recast as follows

$$\begin{aligned} 2[\mathbb{J}\Upsilon_5, \widehat{\mathbf{w}}_1] &= -2[\mathbb{J}\widehat{\mathbf{w}}_1, \Upsilon_5] \quad \text{(skew-symmetry of } \mathbb{J}) \\ &= -2[\mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{n+1}, \Upsilon_5] \quad \text{(since } \mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{n+1} = \mathbb{J}\widehat{\mathbf{w}}_1) \\ &= -2[\mathbf{L}(\mathbf{c})\Upsilon_5, \widehat{\mathbf{w}}_{n+1}] \quad \text{(symmetry of } \mathbf{L}(\mathbf{c})) \\ &= -2[\tfrac{1}{2}s_1\kappa\mathbb{J}\widehat{\mathbf{w}}_{2n+1} + \mathbb{J}\Upsilon_2 - \tfrac{1}{2}\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{n+1}]\,, \end{aligned}$$

using, in the last equality, the Υ_5 equation in (6.6). Substitute this expression into (6.7), noting from Proposition 5.2 that $[\mathbb{J}\widehat{\mathbf{w}}_{2n+1}, \widehat{\mathbf{w}}_{n+1}] = -s_1$, to obtain

$$\begin{aligned} \kappa = {} & 2[\mathbb{J}\Upsilon_2, \widehat{\mathbf{w}}_{n+1}] - [\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{n+1}] \\ & -[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+2}), \widehat{\mathbf{w}}_1] + [\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+1}), \widehat{\mathbf{w}}_1]\,. \end{aligned} \tag{6.8}$$

To eliminate the Υ_2 term, use the second equation of (6.6),

$$\begin{aligned} 2[\mathbb{J}\Upsilon_2, \widehat{\mathbf{w}}_{n+1}] &= -2[\mathbb{J}\widehat{\mathbf{w}}_{n+1}, \Upsilon_2] \quad \text{(skew-symmetry of } \mathbb{J}) \\ &= -2[\mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{2n+1}, \Upsilon_2] \quad \text{(since } \mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{2n+1} = \mathbb{J}\widehat{\mathbf{w}}_{n+1}) \\ &= -2[\mathbf{L}(\mathbf{c})\Upsilon_2, \widehat{\mathbf{w}}_{2n+1}] \quad \text{(symmetry of } \mathbf{L}(\mathbf{c})) \\ &= -2[2\mathbb{J}\Upsilon_1 - \mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{2n+1}] \quad \text{(using } \Upsilon_2 \text{ equation in (6.6))} \end{aligned}$$

Substitute into (6.8), using symmetry of the third derivative of H,

$$\begin{aligned} \kappa = {} & -4[\mathbb{J}\Upsilon_1, \widehat{\mathbf{w}}_{2n+1}] - [\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{n+1}] \\ & -[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{2n+2}), \widehat{\mathbf{w}}_1] + 3[\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{2n+1}), \widehat{\mathbf{w}}_1]\,. \end{aligned} \tag{6.9}$$

It remains to eliminate the Υ_1 term, following similar lines as above

$$\begin{aligned}
4[\mathbb{J}\Upsilon_1, \widehat{\mathbf{w}}_{2n+1}] &= -4[\mathbb{J}\widehat{\mathbf{w}}_{2n+1}, \Upsilon_1] \\
&= -4[\mathbf{L}(\mathbf{c})\widehat{\mathbf{w}}_{2n+2}, \Upsilon_1] \\
&= -4[\mathbf{L}(\mathbf{c})\Upsilon_1, \widehat{\mathbf{w}}_{2n+2}] \\
&= -4[-\tfrac{1}{2}\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_1), \widehat{\mathbf{w}}_{2n+2}] \\
&= 2[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_1), \widehat{\mathbf{w}}_{2n+2}]\,.
\end{aligned}$$

Substituting this expression into (6.9) completes the proof. ■

At this stage, the formula for κ is what one would expect for a coefficient of a quadratic term in the normal form. It involves evaluating the quadratic nonlinearity (or third derivative of the Hamiltonian) on eigenvectors or generalized eigenvectors.

The property of I_1 in Theorem 6.1 is established in the following.

Proposition 6.3. *Consider the perturbation,* $\mathbf{z}(t) = \Phi_{g(t)}(\varphi(\mathbf{c}) + V)$ *with* $g(t) = \exp(t\xi)$*, of a RE with* $\mathbf{c} \in \Sigma^1(\mathbf{P})$*. For* $\|V\|$ *small,* I_1 *is the component of the perturbation of the momentum map in the direction* $\mathfrak{h}^*$*.*

Proof. With the usual identifications, $\langle \mathbf{J}(\mathbf{z}(t)), \boldsymbol{\eta}\rangle$ is the component of the momentum map in the direction $\mathfrak{h}^* \subset \mathfrak{g}^*$. Consider $\langle \mathbf{J}(\mathbf{z}(t)), \boldsymbol{\eta}\rangle$ with $\|V\|$ small,

$$\begin{aligned}
\langle \mathbf{J}(\mathbf{z}(t)), \boldsymbol{\eta}\rangle &= \langle \mathbf{J}(\varphi(\mathbf{c}) + V), \boldsymbol{\eta}\rangle \quad \text{(invariance of momentum map)} \\
&= \langle \mathbf{J}(\varphi(\mathbf{c}) + V_1 + V_2 + \cdots), \boldsymbol{\eta}\rangle \quad \text{(using (6.5))} \\
&= \langle \mathbf{J}(\varphi), \boldsymbol{\eta}\rangle + [\mathrm{d}\langle \mathbf{J}(\varphi(\mathbf{c})), \boldsymbol{\eta}\rangle, V_1] + \mathscr{O}(\|V\|^2) \\
&= \mathscr{K}(\mathbf{c}, 0) + \textstyle\sum_{j=1}^n \eta_j [\xi_M^j(\varphi) \lrcorner\, \Omega, V_1] + \mathscr{O}(\|V\|^2) \quad \text{(definition of } \mathscr{K} \text{ and (2.4))} \\
&= \mathscr{K}(\mathbf{c}, 0) + [\mathbf{w}_1 \lrcorner\, \Omega, V_1] + \mathscr{O}(\|V\|^2) \quad \text{(definition of } \mathbf{w}_1) \\
&= \mathscr{K}(\mathbf{c}, 0) + [\mathbf{w}_1 \lrcorner\, \Omega, -s_1 I_1 \widehat{\mathbf{w}}_{2n+2}] + \mathscr{O}(\|V\|^2) \quad \text{(using (4.7) and (5.3))} \\
&= \mathscr{K}(\mathbf{c}, 0) - s_1 \tfrac{I_1}{a_1}[\widehat{\mathbf{w}}_1 \lrcorner\, \Omega, \widehat{\mathbf{w}}_{2n+2}] + \mathscr{O}(\|V\|^2) \quad \text{(definition of } \widehat{\mathbf{w}}_1) \\
&= \mathscr{K}(\mathbf{c}, 0) + \tfrac{I_1}{a_1} + \mathscr{O}(\|V\|^2) \quad \text{(using Proposition 5.2 and } s_1^2 = +1).
\end{aligned}$$

Hence, to leading order $I_1 = a_1(\langle \mathbf{J}(\mathbf{z}(t)), \boldsymbol{\eta}\rangle - \mathscr{K}(\mathbf{c}, 0))$ where $\mathscr{K}(\mathbf{c}, 0) = \boldsymbol{\eta} \lrcorner\, \mathbf{P}(\mathbf{c})$. ■

7 The role of curvature of the momentum map

In this section the relationship between the formula for κ in Lemma 6.2 and $\mathscr{K}(\mathbf{c},t)$ is established. Now,

$$\mathscr{K}(\mathbf{c},t) = \langle \mathbf{P}(\mathbf{c}+t\boldsymbol{\eta}),\boldsymbol{\eta}\rangle = \langle \mathbf{J}(\varphi(\mathbf{c}+t\boldsymbol{\eta})),\boldsymbol{\eta}\rangle\,,$$

and so

$$\frac{d}{dt}\mathscr{K}(\mathbf{c},t) = \left[\mathrm{d}\langle \mathbf{J}(\varphi(\mathbf{c}+t\boldsymbol{\eta})),\boldsymbol{\eta}\rangle, \frac{d}{dt}\varphi(\mathbf{c}+t\boldsymbol{\eta})\right] = \left[\Xi_M(\varphi(\mathbf{c}+t\boldsymbol{\eta}))\,\lrcorner\,\Omega, \frac{d}{dt}\varphi(\mathbf{c}+t\boldsymbol{\eta})\right]\,,$$

where

$$\Xi_M(\cdot) := \frac{d}{dt}\bigg|_{t=0}\Phi_{\exp(t\Xi)}(\cdot) \quad \text{with} \quad \Xi = \eta_1\xi_1+\cdots+\eta_n\xi_2\,. \tag{7.10}$$

Differentiate again and set $t=0$

$$\begin{aligned}\tfrac{d^2}{dt^2}\Big|_{t=0}\mathscr{K}(\mathbf{c},t) = {} & \left[\Xi_M\left(\tfrac{d}{dt}\Big|_{t=0}\varphi(\mathbf{c}+t\boldsymbol{\eta})\right)\lrcorner\,\Omega, \tfrac{d}{dt}\Big|_{t=0}\varphi(\mathbf{c}+t\boldsymbol{\eta})\right] \\ & + \left[\Xi_M(\varphi(\mathbf{c}))\,\lrcorner\,\Omega, \tfrac{d^2}{dt^2}\Big|_{t=0}\varphi(\mathbf{c}+t\boldsymbol{\eta})\right]\,,\end{aligned}$$

But

$$\frac{d}{dt}\bigg|_{t=0}\varphi(\mathbf{c}+t\boldsymbol{\eta}) = \sum_{j=1}^n \eta_j\frac{\partial\varphi}{\partial c_j} = \sum_{j=1}^n \eta_j\mathbf{v}_{n+j} = \mathbf{w}_{n+1}\,,$$

and

$$\Xi_M(\varphi(\mathbf{c})) = \sum_{j=1}^n \eta_j\xi_M^j(\varphi(\mathbf{c})) = \sum_{j=1}^n \eta_j\mathbf{v}_j = \mathbf{w}_1\,.$$

The completes the first step in determining a formula for $\tfrac{d^2}{dt^2}\Big|_{t=0}\mathscr{K}(\mathbf{c},t)$.

Proposition 7.1.

$$\frac{d^2}{dt^2}\bigg|_{t=0}\mathscr{K}(\mathbf{c},t) = [A_2\,\lrcorner\,\Omega, \mathbf{w}_{n+1}] + [\mathbf{w}_1\,\lrcorner\,\Omega, A_3]\,,$$

where

$$A_2 = \Xi_M\left(\frac{d}{dt}\bigg|_{t=0}\varphi(\mathbf{c}+t\boldsymbol{\eta})\right) \quad \text{and} \quad A_3 = \frac{d^2}{dt^2}\bigg|_{t=0}\varphi(\mathbf{c}+t\boldsymbol{\eta})\,. \tag{7.11}$$

To make further progress, equations for A_2 and A_3 are required.

Proposition 7.2.

$$\begin{aligned}\mathbf{L}(\mathbf{c})\,A_2 &= -\tfrac{1}{a_1^2}\mathrm{D}^3H(\widehat{\mathbf{w}}_1,\widehat{\mathbf{w}}_{n+1}) + A_1\,\lrcorner\,\Omega \\ \mathbf{L}(\mathbf{c})\,A_3 &= -\tfrac{1}{a_1^2}\mathrm{D}^3H(\widehat{\mathbf{w}}_{n+1},\widehat{\mathbf{w}}_{n+1}) + 2A_2\,\lrcorner\,\Omega\,,\end{aligned}$$

where

$$\mathbf{L}(\mathbf{c})\, A_1 = -\frac{1}{a_1^2}\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_1)\,, \quad \text{with} \quad A_1 = \frac{d^2}{dt^2}\bigg|_{t=0} \Phi_{\exp(t\Xi)}(\varphi(\mathbf{c}))\,,$$

and $\widehat{\mathbf{w}}_1$ *and* $\widehat{\mathbf{w}}_{n+1}$ *are defined in (B-7) in Appendix B.*

Proof. Consider (4.2) with $\mathbf{c} \mapsto \mathbf{c} + t\boldsymbol{\eta}$ and the group action replaced by the action in (7.10),

$$\mathrm{d}H(\Phi_{g(s)}(\varphi(\mathbf{c}+t\boldsymbol{\eta}))) - \mathrm{d}\langle \mathbf{J}(\Phi_{g(s)}(\varphi(\mathbf{c}+t\boldsymbol{\eta}))), \widehat{\xi}(t)\rangle = 0\,, \quad g(s) = \exp(s\Xi)\,, \qquad (7.12)$$

and where

$$\widehat{\xi}(t) := (c_1 + t\eta_1)\xi_1 + \cdots + (c_n + t\eta_n)\xi_n\,.$$

Differentiate with respect to s and t and set $s = t = 0$,

$$\mathbf{L}(\mathbf{c})\, A_2 + \mathrm{D}^3 H\left(\frac{d}{dt}\bigg|_{t=0} \varphi(\mathbf{c}+t\boldsymbol{\eta}), \Xi_M(\varphi(\mathbf{c}))\right) - \Xi_M\left(\frac{d}{ds}\Phi_{g(s)}(\varphi)\bigg|_{s=0}\right) \lrcorner\, \Omega = 0\,,$$

using the fact that $D^3\langle \mathbf{J}(\cdot), \xi\rangle = 0$. Substituting in the definitions for A_1, $\mathbf{w}_1$ and $\mathbf{w}_{n+1}$, this is

$$\mathbf{L}(\mathbf{c})\, A_2 = -\mathrm{D}^3 H(\mathbf{w}_{n+1}, \mathbf{w}_1) + A_1 \lrcorner\, \Omega\,.$$

The resulting equation for A_2 follows by using the definitions in (B-7). The verification of the equations for A_3 and A_1 follows the same argument. ■

There is now enough information to complete the analysis of $\frac{d^2}{dt^2}\Big|_{t=0} \mathscr{K}(\mathbf{c}, t)$.

Lemma 7.3. *Let* $\mathscr{K}(\mathbf{c}, t)$ *be the function defined in (1.3) with* $\mathbf{c} \in \Sigma^1(\mathbf{P})$ *and* $\boldsymbol{\eta} \in \mathfrak{h}$. *Then*

$$\begin{aligned} a_1^3 \tfrac{d^2}{dt^2}\Big|_{t=0} \mathscr{K}(\mathbf{c}, t) &= 3\,[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{2n+1}] - 3\,[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_1), \widehat{\mathbf{w}}_{2n+2}] \\ &\quad -[\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{n+1}] \end{aligned}$$

Proof. Evaluate the terms in the formula in Proposition 7.2,

$$\begin{aligned} [\mathbf{w}_1 \lrcorner\, \Omega, A_3] &= [\mathbf{L}(\mathbf{c})\mathbf{w}_{n+1}, A_3] \\ &= [\mathbf{L}(\mathbf{c})A_3, \mathbf{w}_{n+1}] \\ &= \left[-\tfrac{1}{a_1^2}\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{n+1}) + 2A_2 \lrcorner\, \Omega, \mathbf{w}_{n+1}\right] \\ &= -\tfrac{1}{a_1^3}[\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{n+1}] + 2[A_2 \lrcorner\, \Omega, \mathbf{w}_{n+1}]\,, \end{aligned}$$

and so

$$\frac{d^2}{dt^2}\bigg|_{t=0} \mathscr{K}(\mathbf{c}, t) = 3[A_2 \lrcorner\, \Omega, \mathbf{w}_{n+1}] - \frac{1}{a_1^3}[\mathrm{D}^3 H(\widehat{\mathbf{w}}_{n+1}, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{n+1}]\,.$$

Now consider the first term

$$
\begin{aligned}
[A_2 \lrcorner\, \Omega, \mathbf{w}_{n+1}] &= -[\mathbf{w}_{n+1} \lrcorner\, \Omega, A_2] \\
&= -[\mathbf{L}(\mathbf{c})\mathbf{w}_{2n+1}, A_2] \\
&= -[\mathbf{L}(\mathbf{c})A_2, \mathbf{w}_{2n+1}] \\
&= -\left[-\tfrac{1}{a_1^2}\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{n+1}) + A_1 \lrcorner\, \Omega, \mathbf{w}_{2n+1}\right] \\
&= \tfrac{1}{a_1^3}\left[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_{n+1}), \widehat{\mathbf{w}}_{2n+1}\right] - [A_1 \lrcorner\, \Omega, \mathbf{w}_{2n+1}]\,.
\end{aligned}
$$

but

$$
\begin{aligned}
-[A_1 \lrcorner\, \Omega, \mathbf{w}_{2n+1}] &= [\mathbf{w}_{2n+1} \lrcorner\, \Omega, A_1] \\
&= [\mathbf{L}(\mathbf{c})\mathbf{w}_{2n+2}, A_1] \\
&= [\mathbf{L}(\mathbf{c})A_1, \mathbf{w}_{2n+2}] \\
&= -\tfrac{1}{a_1^2}[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_1), \mathbf{w}_{2n+2}] \\
&= -\tfrac{1}{a_1^3}[\mathrm{D}^3 H(\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_1), \widehat{\mathbf{w}}_{2n+2}]\,.
\end{aligned}
$$

Combining the above expressions proves the Lemma. ■

Combining Lemma 7.3 with Lemma 6.2 proves the formula (1.6).

8 The bifurcating homoclinic manifold and its geometric phase

The leading order normal form is integrable, and when the higher-order terms are neglected, an explicit solution can be obtained. The flow along the group satisfies

$$
\dot{I}_1 = \cdots = \dot{I}_n = 0\,, \quad \dot{\phi}_1 = u\,, \quad \dot{\phi}_j = s_j I_j\,, \quad j = 2, \ldots n\,.
$$

$I_1, \ldots, I_n$ are constant and

$$
\phi_1(t) = \int u(t)\,\mathrm{d}t + \phi_1^0\,, \quad \phi_j = s_j I_j\, t + \phi_j^0\,, \quad j = 2, \ldots, n\,, \tag{8.1}
$$

with the contribution from the homoclinic orbit $u(t)$ determined from the reduced system

$$
-v_t = I_1 - \tfrac{1}{2}\kappa u^2 \quad \text{and} \quad u_t = s_1 v\,.
$$

The reduced system can be explicitly solved: $v = s_1 u_t$ and

$$
u(t) = \nu - 3\nu\,\mathrm{sech}^2(\gamma t)\,,
$$

with

$$
\gamma = \frac{1}{2}\sqrt{s_1 \kappa \nu} \quad \text{and} \quad \nu = \begin{cases} -\sqrt{\frac{2I_1}{\kappa}} & \text{if } s_1\kappa < 0 \\ +\sqrt{\frac{2I_1}{\kappa}} & \text{if } s_1\kappa > 0. \end{cases}
$$

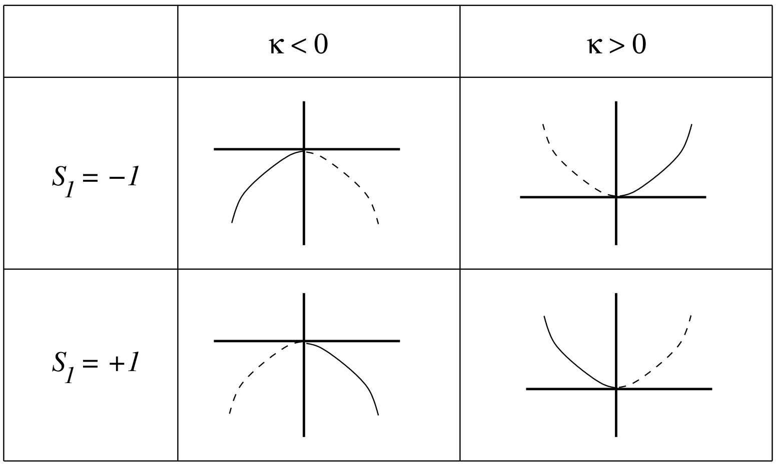

Fig. 3. Curves of RE of the normal form plotted in $\mathfrak{h} \times \mathfrak{h}^*$ (in coordinates, the $(\nu, I_1)-$plane) as a function of s_1 and κ. The dashed (solid) lines identify branches of hyperbolic (elliptic) RE. The horizontal axis in each plot is the line $I_1 = 0$.

Clearly existence requires

$$\kappa I_1 > 0 \quad \text{and} \quad s_1 \kappa \nu > 0\,. \tag{8.2}$$

The line $I_1 = 0$ defines the local tangent space of the image of $\Sigma^1(\mathbf{P})$ in $\mathfrak{g}^*$. Given the sign of κ, the first condition in (8.2) indicates whether RE persist for $I_1 > 0$ or $I_1 < 0$. Given s_1 and the sign of κ, the second inequality determines which branch of RE is hyperbolic. There are four cases and Figure 3 shows them as a function of s_1 and κ.

8.1 The induced geometric phase

The phase to leading order can now be determined by substituting the expression for $u(t)$ into the ϕ_1 equation in (8.1)

$$\phi_1(t) = \nu t - \frac{3\nu}{\gamma}\tanh(\gamma t) + \phi_1^0\,,$$

with $\phi_2(t), \ldots, \phi_n(t)$ retaining their form in (8.1). The geometric part of the phase shift is

$$\Delta\phi_1 = \Big[\phi_1(t) - \nu t\Big]_{-\infty}^{+\infty} = -\frac{6}{\gamma}\nu = -12\,\frac{s_1}{\kappa}\,(2\kappa I_1)^{1/4}\,. \tag{8.3}$$

The geometric phase has a direction in the group. This direction is clearly not invariant under coordinate change. In normal form coordinates it is the ϕ_1 direction. In the original coordinates, the direction is determined by unwrapping the normal form transformations.

The *dynamic phase* to leading order in the normal form also has a direction, $\mathbf{v}$, since to leading order it is

$$\phi^{\rm dyn}(t) = \phi^0 + \mathbf{v}t\,, \quad \text{with} \quad \mathbf{v} = (\nu, s_2 I_2, \ldots, s_n I_n)\,.$$

The terms “geometric phase” and “reconstruction phase” are used here informally.

After normal form transformations the geometric and dynamic phases can be explicitly determined, to leading order, so formalization of such phases is not pursued here.

In order to give a general theory for the fully nonlinear problem in a neighborhood of the degenerate RE, there are two issues that need to be addressed. First, existing theory of reconstruction (e.g. Chapter 5 of [17]) would have to be modified to take into account that the orbit in the reduced space is a homoclinic orbit – not a closed orbit. The second issue is the more familiar issue of introducing an appropriate connection that would enable precise distinction between the horizontal and vertical subspaces in the definition of geometric phase.

9 Intermezzo: failure of the G-Morse hypothesis

This section is a slight digression from the main theme of the paper. The purpose is threefold: to give a mechanism for failure of the G-Morse hypothesis (the type II degeneracy) showing that it is complementary to the type I degeneracy of $\mathrm{D}\mathbf{P}(\mathbf{c})$; secondly to give a simple example illustrating the mechanism; thirdly to show that the degeneracies that have been studied in the $N-$body problem correspond to failure of the G-Morse hypothesis but by the type III degeneracy.

Consider the type III degeneracy first. When the Hamiltonian function depends on additional parameters, the matrix representation of the second variation of the augmented Hamiltonian is a parameter-dependent matrix. Generically, in such a multi-parameter family of matrices, additional zero eigenvalues can arise, resulting in failure of the G-Morse hypothesis. This degeneracy is called type III in this paper. Determination of the codimension and a precise characterization of the above observation can be obtained using the theory of versal deformation of matrices (cf. Arnold [1]). It is this type III degeneracy that is called "degenerate relative equilibria" in Palmore [27], in the context of the $N-$body problem. In [27], the value of the momentum map is fixed (see definition of S_m on page 423 of [27]), and it is the masses $m_1,\ldots,m_n$ which are varied. Hence the second variation is a matrix dependent on the mass parameters and Palmore determines subsets of the mass parameter space where the RE are degenerate.

The other way that the G-Morse hypothesis can fail – called type II degeneracy here – without any additional parameters in the Hamiltonian, that is natural in applications, is when the mapping $(\mathbf{c},\mathbf{P}(\mathbf{c}))\in\mathfrak{g}\times\mathfrak{g}^*$ fails to be a graph. To see this re-parameterize the RE as $(\mathbf{c}(\mu),\mu)\in\mathfrak{g}\times\mathfrak{g}^*$; that is, take the values of the momentum map as the parameters.

Lemma 9.1. *Suppose the family of RE is parameterized by* $\mu\in\mathfrak{g}^*$ *and* $\det(\mathrm{D}\mathbf{c}(\boldsymbol{\mu}))=0$. *Then*

$$\dim\mathrm{Ker}(\mathrm{D}^2H_\xi(\varphi))>\dim T_\varphi(\mathrm{G}\cdot\varphi)\,.$$

Proof. It is already clear that $T_\varphi(\mathrm{G}\cdot\varphi)\subset\mathrm{Ker}(\mathbf{L}(\mathbf{c}))$. Hence the statement is proved if there exists at least one additional independent eigenvector in $\mathrm{Ker}(\mathbf{L}(\mathbf{c}))$. With the hypothesis $\det(\mathrm{D}\mathbf{c}(\boldsymbol{\mu}))=0$ there exists a vector $\boldsymbol{\gamma}\in T_\mu\mathfrak{g}^*$ satisfying

$$\mathrm{D}\mathbf{c}(\boldsymbol{\mu})\boldsymbol{\gamma}=0\,.$$

Consider the equation satisfied by φ but considered as a function of μ,

$$\mathrm{d}H(\varphi)-\mathrm{d}\langle\mathbf{J}(\varphi),\xi(\mu)\rangle=0\,. \tag{9.1}$$

Take a basis for $\mathfrak{g}$ and a dual basis for $\mathfrak{g}^*$; in terms of this basis $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)$. Differentiate the equation (9.1) with respect to μ_j,

$$\mathbf{L}(\boldsymbol{\mu})\frac{\partial\varphi}{\partial\mu_j}=\sum_{k=1}^n\frac{\partial c_k}{\partial\mu_j}\,\mathrm{d}J_k(\varphi)\,.$$

where $\mathbf{L}(\boldsymbol{\mu})$ is the second variation of the augmented Hamiltonian parameterized by $\boldsymbol{\mu}$.

Consider

$$\mathbf{V}=\sum_{j=1}^n\gamma_j\frac{\partial\varphi}{\partial\mu_j}\,,\qquad \boldsymbol{\gamma}\in\mathrm{Ker}(\mathrm{D}\mathbf{c}(\boldsymbol{\mu}))\,.$$

Then

$$\begin{aligned}\mathbf{L}(\boldsymbol{\mu})\mathbf{V}&=\textstyle\sum_{j=1}^n\gamma_j\mathbf{L}(\boldsymbol{\mu})\frac{\partial\varphi}{\partial\mu_j}\\&=\textstyle\sum_{j=1}^n\gamma_j\sum_{k=1}^n\frac{\partial c_k}{\partial\mu_j}\mathrm{d}J_k(\varphi)\\&=\textstyle\sum_{k=1}^n\left(\sum_{j=1}^n\gamma_j\frac{\partial c_k}{\partial\mu_j}\right)\mathrm{d}J_k(\varphi)\\&=0\,.\end{aligned}$$

This proves that $\mathbf{V}$ is in the kernel of $\mathbf{L}(\boldsymbol{\mu})$. It remains only to show that the $(n+1)$ vectors $\mathbf{v}_1,\ldots,\mathbf{v}_n,\mathbf{V}$ are linearly independent. This set is linearly independent if and only if there exists real parameters $\alpha_1,\ldots,\alpha_{n+1}$ such that

$$\alpha_1\mathbf{v}_1+\cdots+\alpha_n\mathbf{v}_n+\alpha_{n+1}\mathbf{V}=0\quad\Rightarrow\quad\alpha_1=\cdots=\alpha_{n+1}=0\,,$$

which is equivalent to

$$\alpha_1\mathbf{v}_1\;\lrcorner\;\Omega+\cdots+\alpha_n\mathbf{v}_n\;\lrcorner\;\Omega+\alpha_{n+1}\mathbf{V}\;\lrcorner\;\Omega=0\quad\Rightarrow\quad\alpha_1=\cdots=\alpha_{n+1}=0\,. \tag{9.2}$$

Now, the set $\mathbf{v}_1,\ldots,\mathbf{v}_n$ is linearly independent, hence the result will follow if (9.2) implies $\alpha_{n+1}=0$.

Consider the following property of the momentum map

$$\mu_j=J_j(\varphi)\quad\Rightarrow\quad\delta_{jk}=\left[\mathrm{d}J_j(\varphi),\frac{\partial\varphi}{\partial\mu_k}\right]=\left[\mathbf{v}_j\;\lrcorner\;\Omega,\frac{\partial\varphi}{\partial\mu_k}\right]\,.$$

Hence

$$\gamma_j=\sum_{k=1}^n\gamma_k\delta_{jk}=\sum_{k=1}^n\left[\mathbf{v}_j\;\lrcorner\;\Omega,\gamma_k\frac{\partial\varphi}{\partial\mu_k}\right]=-\left[\mathbf{V}\;\lrcorner\;\Omega,\mathbf{v}_j\right]\,.$$

Now pairing (9.2) with each $\mathbf{v}_j$, using Proposition 2.3 (or the first part of Proposition 4.3) and noting that $\boldsymbol{\gamma}$ is not identically zero proves that $\alpha_1=\cdots=\alpha_{n+1}=0$. ■

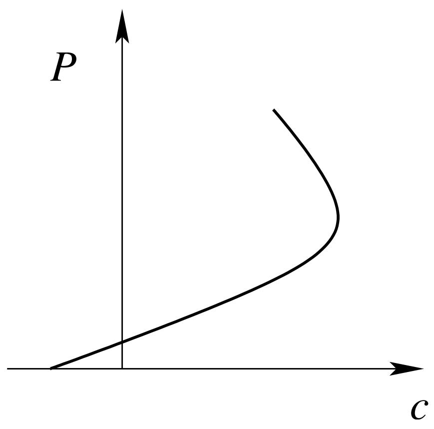

Fig. 4. A parametric representation of the mapping $(c(u), P(u))$, parameterized by $u \in \mathbb{R}^+$, in $\mathfrak{g} \times \mathfrak{g}^*$ for RE of the Hamiltonian system (9.3).

9.1 An example with G-Morse degeneracy

The following example illustrates the failure of the G-Morse hypothesis. It is a model for water waves constructed by Zufiria & Saffman [37]. It is an S^1−equivariant Hamiltonian system on $\mathbb{R}^4$ with coordinates $\mathbf{z} = (q_1, q_2, p_1, p_2)$ and standard symplectic structure. It will be easier to work with the complex coordinates

$$a_j = \frac{1}{\sqrt{2}}(q_j + \mathrm{i}p_j)\,, \quad j = 1, 2\,.$$

In these coordinates, the Hamiltonian function is

$$\begin{aligned} H(\mathbf{z}) = \tfrac{1}{277}\Bigg(&\tfrac{1}{3}a_1\overline{a_1}(-509 + \tfrac{3557}{4}a_1\overline{a_1} + \tfrac{257}{3}(a_1\overline{a_1})^2) \\ &-922a_2\overline{a_2} + (a_1^2\overline{a_2} + \overline{a_1}^2 a_2)\left(269 - \tfrac{353}{4}a_1\overline{a_1} + \tfrac{21}{4}(a_1\overline{a_1})^2\right)\Bigg)\,, \end{aligned} \tag{9.3}$$

where $\overline{(\;)}$ denotes complex conjugation. The Hamiltonian is S^1−invariant with the action of S^1 given by $(\mathrm{e}^{\mathrm{i}\theta}a_1, \mathrm{e}^{2\mathrm{i}\theta}a_2)$. The Lie algebra is $\mathbb{R}$ and the momentum map is

$$J(\mathbf{z}) = a_1\overline{a_1} + 2a_2\overline{a_2}\,.$$

It is straightforward to construct families of RE for this system (see [37]) and a parametric representation of (c, P) for one of the families is shown in Figure 4. When $c = 3$ there is a point where the slope is vertical and by Lemma 9.1 this corresponds to a point where the dimension of the kernel of the second variation of the augmented Hamiltonian is greater than one. Note that the point of degeneracy creates a local barrier in $\mathfrak{g}$ for the persistence of RE: locally, RE exist in $\mathfrak{g}$ for $c \le 3$.

One can vary the coefficients in the Hamiltonian for this example to also find points where the standard degeneracy $P'(c) = 0$ occurs as well as examples where both types of degeneracy occur in the same family of RE.

10 Creating solitary waves via a degenerate RE

When the group is one-dimensional and the manifold is a vector space of dimension four the theory simplifies dramatically and most details can be worked out explicitly. It is however an instructive example, highlighting in the simplest possible context some features of the theory, and at the same time its use in applications can be illustrated.

Consider the following system of PDEs, an example of a Boussinesq system which model water waves in shallow water [7]

$$\begin{aligned} \tfrac{\partial h}{\partial t} + u\tfrac{\partial h}{\partial x} + h\tfrac{\partial u}{\partial x} + \tfrac{1}{3}\tfrac{\partial^3 u}{\partial x^3} &= 0 \\ \tfrac{\partial u}{\partial t} + \tilde{g}\tfrac{\partial h}{\partial x} + u\tfrac{\partial u}{\partial x} &= 0 \end{aligned} \tag{10.1}$$

The functions $h(x,t)$ and $u(x,t)$ represent the depth and velocity of the fluid layer, and $\tilde{g}$ is a given strictly positive constant representing gravity. Scale coordinates so that $\tilde{g}$ can be taken to be unity. Then steady solutions of this system satisfy

$$\frac{d}{dx}\left(hu + \tfrac{1}{3}u_{xx}\right) = 0 \quad \text{and} \quad h + \tfrac{1}{2}u^2 = R\,,$$

for some strictly positive constant R. Integrating the first equation and substituting for h from the second equation leads to a planar system for $u(x)$ which can be analyzed completely. This planar system has a homoclinic orbit which represents a solitary wave solution of the Boussinesq system, which in turn models the famous Korteweg-DeVries solitary wave solution of the water wave problem [7].

These solitary waves (homoclinic orbits) can be related to degenerate RE as follows. Introduce new coordinates $\mathbf{z} = (q_1, q_2, p_1, p_2)$ with $u = q_2$ and

$$q_2 = \frac{dq_1}{dx}\,, \quad p_2 = -\tfrac{1}{3}\frac{dq_2}{dx}\,, \quad p_1 = R\,q_2 - \frac{1}{2}q_2^3 - \frac{dp_2}{dx}\,.$$

The above steady system can be represented by the Hamiltonian system

$$\mathbf{z}_x = X_H(\mathbf{z})\,, \quad \mathbf{z} \in M = \mathbb{R}^4\,, \tag{10.2}$$

with $\Omega = \mathrm{d}q_1 \wedge \mathrm{d}p_1 + \mathrm{d}q_2 \wedge \mathrm{d}p_2\,,$ $H(\mathbf{z}, R) = p_1 q_2 - \tfrac{3}{2}p_2^2 - \tfrac{1}{2}Rq_2^2 + \tfrac{1}{8}q_2^4$ and so

$$X_H(\mathbf{z}) = q_2\frac{\partial}{\partial q_1} - 3p_2\frac{\partial}{\partial q_2} + (Rq_2 - p_1 - \tfrac{1}{2}q_2^3)\frac{\partial}{\partial p_2}\,.$$

This Hamiltonian system has a one-parameter affine translation symmetry $\mathrm{G} = \mathbb{R}$ with generator $\xi_M(\mathbf{z}) = \frac{\partial}{\partial q_1}$, and momentum map

$$J(\mathbf{z}) = \langle \mathbf{J}(\mathbf{z}), \xi \rangle = p_1\,.$$

The family of RE associated with this group is $\mathbf{z}(x) = \Phi_{g(x)}(\varphi)$. The family of critical points of the augmented Hamiltonian is

$$\varphi(c) = (0, c, cR - \tfrac{1}{2}c^3, 0)\,.$$

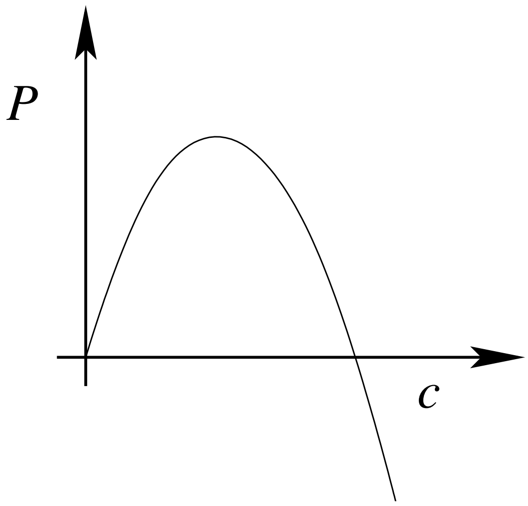

Fig. 5. Graph of the pullback of the momentum map $P : \mathfrak{g} \to \mathfrak{g}^*$ for RE of the Hamiltonian system (10.2) taking $R > 0$ fixed.

This family exists for all $c \in \mathbb{R}$ and so $\mathfrak{g}^{\rm RE} = \mathfrak{g}$. The pullback of the momentum by φ is

$$P(c) = J \circ \varphi = Rc - \tfrac{1}{2}c^3 \quad \text{and} \quad {\rm D}P(c) = R - \frac{3}{2}c^2 \,.$$

Degeneracy occurs when $c^2 = \frac{3}{2}R$ and so

$$\Sigma^1(P) = \left\{ \pm \left(\frac{3}{2}R \right)^{1/2} \right\}$$

and there is no further degeneracy as long as $R > 0$. The graph of the mapping $P(c)$ is shown in Figure 5 for $c \geq 0$. The image of $\Sigma^1(P)$ in $\mathfrak{g}^*$ consists of the two points $\pm\frac{\sqrt{6}}{8}R^{3/2}$. Figure 5 shows in a simple context how the image of $\Sigma^1(P)$ is a barrier in $\mathfrak{g}^*$ to the continuation of the family of RE. When $c > 0$ RE exist only in the region $P < \frac{\sqrt{6}}{8}R^{3/2}$. This barrier is local, since globally there exist RE with $P > \frac{\sqrt{6}}{8}R^{3/2}$ when c is negative.

The leading-order nonlinear normal form when $c \in \Sigma^1(P)$ is given by Theorem 6.1 with $n = 1$ and

$$\kappa = a_1^3 P''(c) = -3a_1^3 c \,, \quad c \in \Sigma^1(P) \,, \tag{10.3}$$

for some positive constant a_1. Hence ${\rm sign}(\kappa) = -{\rm sign}(c)$. From this normal form, the homoclinic orbit representing the solitary wave can be constructed, and the local region in parameter space where it exists identified. Further detail, particularly the sign s_1 requires computation of the eigenvectors.

10.1 Constructive aspects of the case $n = 1$

To illustrate the constructive aspects of the theory, an explicit computation of the linear normal form will be given. First reduce the formulas from Appendix B to the case $n = 1$. The starting point is the Jordan chain $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ satisfying (4.1)-(4.4) with $n = 1$. The differential ${\rm D}\mathbf{P}(\mathbf{c})$ is one-dimensional and so $\boldsymbol{\eta}$ is unity and the transformed symplectic basis is

$$\{\,\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_2, -s_1\widehat{\mathbf{w}}_4, s_1\widehat{\mathbf{w}}_3\,\}$$

with

$$\widehat{\mathbf{w}}_1 = a_1\mathbf{v}_1\,, \quad \widehat{\mathbf{w}}_2 = a_1\mathbf{v}_2\,,$$

$$\widehat{\mathbf{w}}_3 = a_1\mathbf{v}_3 + m_1\mathbf{v}_1\,, \quad \widehat{\mathbf{w}}_4 = a_1\mathbf{v}_4 + m_1\mathbf{v}_2\,,$$

with

$$a_1 = |\Omega(\mathbf{v}_1,\mathbf{v}_4)|^{-1/2}\,, \quad s_1 = \mathrm{sign}\,\Omega(\mathbf{v}_4,\mathbf{v}_1)\,,$$

and

$$m_1 = \tfrac{1}{2}s_1a_1^3\,\Omega(\mathbf{v}_3,\mathbf{v}_4)\,.$$

Apply this theory to the system (10.2) perturbed about a degenerate RE. The elements of the Jordan chain $\mathbf{v}_1,\ldots,\mathbf{v}_4$ are easily computed,

$$\mathbf{v}_1 = \frac{\partial}{\partial q_1}\,, \quad \mathbf{v}_2 = \frac{\partial}{\partial q_2}\,, \quad \mathbf{v}_3 = -\frac{1}{3}\frac{\partial}{\partial p_2}\,, \quad \mathbf{v}_4 = \frac{1}{3}\frac{\partial}{\partial p_1}\,.$$

The parameters are $m_1 = 0$,

$$s_1 = -\mathrm{sign}\,\Omega(\mathbf{v}_1,\mathbf{v}_4) = -\mathrm{sign}\left(\frac{1}{3}\right) = -1\,, \quad a_1 = |\Omega(\mathbf{v}_4,\mathbf{v}_1)|^{-1/2} = \sqrt{3}\,.$$

Hence the scaled vectors which generate a symplectic basis are

$$\widehat{\mathbf{w}}_1 = \sqrt{3}\frac{\partial}{\partial q_1}\,, \quad \widehat{\mathbf{w}}_2 = \sqrt{3}\frac{\partial}{\partial q_2}\,, \quad \widehat{\mathbf{w}}_3 = -\frac{\sqrt{3}}{3}\frac{\partial}{\partial p_2}\,, \quad \widehat{\mathbf{w}}_4 = \frac{\sqrt{3}}{3}\frac{\partial}{\partial p_1}\,.$$

Since $a_1 = \sqrt{3}$, $\kappa = -9\sqrt{3}c$. For $c > 0$ RE persist for $I_1 < 0$ which is consistent with Figure 5. The other critical parameter in (8.2) is $s_1\kappa\nu$. For $c > 0$, $s_1\kappa > 0$ and so it is the right $P'(c) < 0$ branch of RE which is hyperbolic in Figure 5.

11 Re-interpretation as a two-dimensional group

The example of §10 has an additional parameter, R. This parameter can be interpreted as an element of a larger dimensional momentum map. In this section the example in §10 is re-interpreted as a Hamiltonian system with no external parameters, but a two-component momentum map.

Introduce new coordinates $\mathbf{z} = (q_1,q_2,q_3,p_1,p_2,p_3)$ with $u = q_3$,

$$q_3 = \frac{dq_2}{dx}\,, \quad p_3 = -\tfrac{1}{3}\frac{dq_3}{dx}\,,$$

and

$$p_1 = \frac{dq_1}{dx} + \tfrac{1}{2}q_3^2\,, \quad p_2 = p_1q_3 - \frac{1}{2}q_3^3 - \frac{dp_3}{dx}\,.$$

The coordinate p_1 represents the parameter R in §10 and it satisfies the trivial equation $\frac{dp_1}{dx} = 0$, but there is additional geometry due to its presence in the momentum map.

The Hamiltonian vectorfield, representing the steady part of (10.1), is

$$X_H(\mathbf{z}) = (p_1 - \tfrac{1}{2}q_3^2)\frac{\partial}{\partial q_1} + q_3\frac{\partial}{\partial q_2} - 3p_3\frac{\partial}{\partial q_3} + (p_1q_3 - p_2 - \tfrac{1}{2}q_3^3)\frac{\partial}{\partial p_3}\,. \tag{11.1}$$

with the standard symplectic structure and Hamiltonian function

$$H(\mathbf{z}) = p_2q_3 - \tfrac{3}{2}p_3^2 + \frac{1}{2}p_1^2 - \frac{1}{2}p_1q_3^2 + \frac{1}{8}q_3^4\,.$$

The group G is a two-parameter affine translation symmetry with generators

$$\xi_M^1(\mathbf{z}) = \frac{\partial}{\partial q_1} \quad \text{and} \quad \xi_M^2(\mathbf{z}) = \frac{\partial}{\partial q_2}\,,$$

associated with a basis $\{\xi_1, \xi_2\}$ for $\mathfrak{g}$. The symmetry group $\mathbb{R}^2$ represents the fact that H is independent of q_1 and q_2. Now

$$\xi_M^1 \lrcorner\, \Omega = \mathrm{d}p_1 \quad \text{and} \quad \xi_M^2 \lrcorner\, \Omega = \mathrm{d}p_2\,,$$

and so the momentum map, using the dual basis $\{\xi_1^*, \xi_2^*\}$, is

$$\mathbf{J}(\mathbf{z}) = J_1(\mathbf{z})\xi_1^* + J_2(\mathbf{z})\xi_2^*\,, \quad \text{with} \quad J_1(\mathbf{z}) = p_1 \quad \text{and} \quad J_2(\mathbf{z}) = p_2\,.$$

The family of RE associated with this group is

$$\mathbf{z}(x) = \Phi_{\exp(x\xi)}(\varphi)\,, \quad \text{with} \quad \xi = c_1\xi_1 + c_2\xi_2\,,$$

and $\varphi(\mathbf{c}) = (0, 0, c_2, c_1 + \frac{1}{2}c_2^2, c_1, c_2, 0)$ and so

$$P_1(\mathbf{c}) = c_1 + \tfrac{1}{2}c_2^2 \quad \text{and} \quad P_2(\mathbf{c}) = c_1c_2\,, \tag{11.2}$$

with Jacobian

$$\mathrm{D}\mathbf{P}(\mathbf{c}) = \begin{pmatrix} 1 & c_2 \\ c_2 & c_1 \end{pmatrix}.$$

The family of critical points $\varphi(\mathbf{c})$ exists for all $\mathbf{c} \in \mathfrak{g}$ and so $\mathfrak{g}^{\mathrm{RE}} = \mathfrak{g}$.

The Jacobian is degenerate when $c_2^2 = c_1$ and so

$$\Sigma^1(\mathbf{P}) = \{\, \mathbf{c} \in \mathfrak{g} \;:\; c_2^2 - c_1 = 0 \,\}\,.$$

The Jacobian $\mathrm{D}\mathbf{P}(\mathbf{c})$ has rank 1 as long as $1 + c_1 \neq 0$. But $\mathbf{c} \in \Sigma^1(\mathbf{P})$ implies $c_1 \geq 0$ and so the rank can not drop to zero.

In this example the set $\Sigma^{11}(\mathbf{P})$ is not empty. The tangent space of $\Sigma^1(\mathbf{P})$ is

$$T_c\Sigma^1(\mathbf{P}) = \left\{ \operatorname{span}\begin{pmatrix} 2c_2 \\ 1 \end{pmatrix} \right\} \subset T_c\,\mathfrak{g}\,,$$

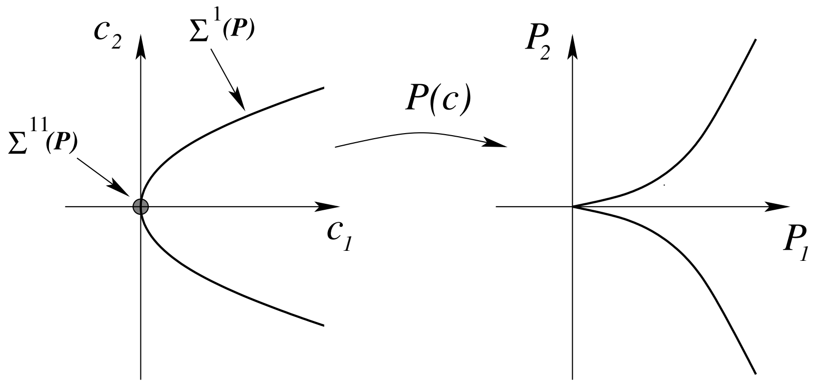

Fig. 6. The set $\Sigma^1(\mathbf{P})$ in $\mathfrak{g}$ and its image under P in $\mathfrak{g}^*$

For $\mathbf{c} \in \Sigma^1(\mathbf{P})$,

$$\mathrm{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c})) = \mathrm{span}\{\boldsymbol{\eta}\}\,, \quad \boldsymbol{\eta} = \frac{1}{\sqrt{1+c_1}} \begin{pmatrix} -c_2 \\ 1 \end{pmatrix} .$$

The geometry of the curve of degeneracy in momentum space is illustrated in Figure 6. The kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$ is in $T_c\Sigma^1(\mathbf{P})$ when $c_2 = 0$. Hence $\Sigma^{11}(\mathbf{P}) = \{\mathbf{c} = (c_1, 0)\} \cap \Sigma^1(\mathbf{P})$ which consists of just the origin in $\mathfrak{g}$. As shown in Figure 6, the image of the point in $\Sigma^{11}(\mathbf{P})$ is the cusp point in the $\mathfrak{g}^*$ plane. This is an example of the Whitney cusp (cf. Chapter 1 of [2]).

The leading order nonlinear normal form for $\mathbf{c} \in \Sigma^1(\mathbf{P}) \backslash \Sigma^{11}(\mathbf{P})$ is given by Theorem 6.1 with $n = 2$. The signs s_2 and the sign of κ can be computed without the eigenvectors. s_2 is the sign of the nonzero eigenvalue of $\mathrm{D}\mathbf{P}(\mathbf{c})$ and $\lambda_2 = \mathrm{Trace}(\mathrm{D}\mathbf{P}(\mathbf{c})) = 1 + c_1 > 0$ and so $s_2 = +1$. To determine κ,

$$\mathscr{K}(\mathbf{c}, t) = \eta_1((c_1 + t\eta_1) + \tfrac{1}{2}(c_2 + t\eta_2)^2) + \eta_2(c_1 + t\eta_1)(c_2 + t\eta_2)\,.$$

Hence

$$\left.\frac{d^2}{dt^2}\right|_{t=0} \mathscr{K}(\mathbf{c}, t) = 3\eta_1\eta_2^2$$

and so

$$\kappa = a_1^3 \left.\frac{d^2}{dt^2}\right|_{t=0} \mathscr{K}(\mathbf{c}, t) = 3\eta_1\eta_2^2 a_1^3 = -3c_2 \left(\frac{a_1}{\sqrt{1+c_1}}\right)^3 ,$$

and so $\mathrm{sign}(\kappa) = -\mathrm{sign}(c_2)$. To go further and compute an explicit expression for a_1 and to determine the symplectic sign s_1, the eigenvectors need to be computed.

Given the Jordan chain $\mathbf{v}_1, \ldots, \mathbf{v}_6$ and the kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$, the normalized Jordan chain is constructed using the theory in Appendix B with $n = 2$,

$$\begin{aligned}
\widehat{\mathbf{w}}_1 &= a_1(\eta_1 \mathbf{v}_1 + \eta_2 \mathbf{v}_2)\,, \quad \widehat{\mathbf{w}}_2 = a_2(-\eta_2 \mathbf{v}_1 + \eta_1 \mathbf{v}_2)\,, \\
\widehat{\mathbf{w}}_3 &= a_1(\eta_1 \mathbf{v}_3 + \eta_2 \mathbf{v}_4) \\
\widehat{\mathbf{w}}_4 &= a_2(-\eta_2 \mathbf{v}_3 + \eta_1 \mathbf{v}_4) + B_{12}(\eta_1 \mathbf{v}_1 + \eta_2 \mathbf{v}_2) \\
\widehat{\mathbf{w}}_5 &= a_1 \mathbf{v}_5 + m_1(\eta_1 \mathbf{v}_1 + \eta_2 \mathbf{v}_2) + m_2(-\eta_2 \mathbf{v}_1 + \eta_1 \mathbf{v}_2) \\
\widehat{\mathbf{w}}_6 &= a_1 \mathbf{v}_6 + m_1(\eta_1 \mathbf{v}_3 + \eta_2 \mathbf{v}_4) + m_2(-\eta_2 \mathbf{v}_3 + \eta_1 \mathbf{v}_4)\,,
\end{aligned}$$

with the parameters determined by

$$\begin{aligned}
a_1 &= |\Omega(\eta_1 \mathbf{v}_1 + \eta_2 \mathbf{v}_2, \mathbf{v}_6)|^{-1/2}\,, \quad s_1 = \operatorname{sign}\Omega(\mathbf{v}_6, \eta_1 \mathbf{v}_1 + \eta_2 \mathbf{v}_2)\,, \\
a_2 &= |\lambda_2|^{-1/2}\,, \quad s_2 = \operatorname{sign}(\lambda_2) \\
B_{12} &= s_1 a_1^2 a_2\, \Omega(-\eta_2 \mathbf{v}_3 + \eta_1 \mathbf{v}_4, \mathbf{v}_6)\,, \quad m_2 = -\tfrac{a_1}{\lambda_2}\Omega(-\eta_2 \mathbf{v}_1 + \eta_1 \mathbf{v}_2, \mathbf{v}_6) \\
m_1 &= \tfrac{1}{2} s_1 a_1 \left(a_1^2\, \Omega(\mathbf{v}_5, \mathbf{v}_6) - \lambda_2 m_2^2\right)\,,
\end{aligned}$$

Apply this theory to the RE of the Hamiltonian system (11.1) with $\mathbf{c} \in \Sigma^1(\mathbf{P})$. The natural eigenvectors associated with the basis $\{\xi_1, \xi_2\}$ for $\mathfrak{g}$ are

$$\begin{aligned}
\mathbf{v}_1 &= \tfrac{\partial}{\partial q_1}\,, & \mathbf{v}_2 &= \tfrac{\partial}{\partial q_2}\,, & \mathbf{v}_3 &= \tfrac{\partial}{\partial p_1} + c_2 \tfrac{\partial}{\partial p_2} \\
\mathbf{v}_4 &= \tfrac{\partial}{\partial q_3} + c_2 \tfrac{\partial}{\partial p_1} + c_1 \tfrac{\partial}{\partial p_2}\,, & \mathbf{v}_5 &= -\tfrac{1}{3}\eta_2 \tfrac{\partial}{\partial p_3}\,, & \mathbf{v}_6 &= \tfrac{1}{3}\eta_2 \tfrac{\partial}{\partial p_2}\,.
\end{aligned}$$

whence

$$\Omega(\eta_1 \mathbf{v}_1 + \eta_2 \mathbf{v}_2, \mathbf{v}_6) = \frac{1}{3}\eta_2^2\,,$$

and so

$$s_1 = -1\,, \quad a_1 = \frac{\sqrt{3}}{|\eta_2|}\,, \quad s_2 = \operatorname{sign}(\lambda_2) = +1\,.$$

The other parameters are

$$B_{12} = 0\,, \quad m_1 = \frac{\eta_1^2}{2\sqrt{3}|\eta_2|\lambda_2} \quad \text{and} \quad m_2 = -\frac{\eta_1\eta_2}{\sqrt{3}|\eta_2|\lambda_2}\,.$$

The normalized eigenvectors are therefore

$$\widehat{\mathbf{w}}_1 = \sqrt{3}\left(-c_2\tfrac{\partial}{\partial q_1} + \tfrac{\partial}{\partial q_2}\right), \quad \widehat{\mathbf{w}}_2 = -\tfrac{1}{\lambda_2}\left(\tfrac{\partial}{\partial q_1} + c_2\tfrac{\partial}{\partial q_2}\right),$$

$$\widehat{\mathbf{w}}_3 = \sqrt{3}\tfrac{\partial}{\partial q_3}, \quad \widehat{\mathbf{w}}_4 = -\tfrac{c_2}{\lambda_2}\tfrac{\partial}{\partial q_3} - \tfrac{\partial}{\partial p_1} - c_2\tfrac{\partial}{\partial p_2},$$

$$\widehat{\mathbf{w}}_5 = -\tfrac{\sqrt{3}}{6}\tfrac{c_2}{\lambda_2^2}\left((1+\lambda_2)\tfrac{\partial}{\partial q_1} + c_2\tfrac{\partial}{\partial q_2}\right) - \tfrac{1}{\sqrt{3}}\tfrac{\partial}{\partial p_3},$$

$$\widehat{\mathbf{w}}_6 = -\tfrac{\sqrt{3}}{6}\tfrac{c_2}{\lambda_2^2}\left(c_2\tfrac{\partial}{\partial q_3} + 2\lambda_2\tfrac{\partial}{\partial p_1} + 2c_2\lambda_2\tfrac{\partial}{\partial p_2}\right) + \tfrac{1}{\sqrt{3}}\tfrac{\partial}{\partial p_2},$$

The full formula for κ is now

$$\kappa = -\frac{3c_2}{(1+c_1)^{3/2}}a_1^3 = -9\sqrt{3}\,c_2\,.$$

Note that this expression for κ agrees with that found in (10.3) in §10 as it should, since the reduced system is the same, with only the dimension of the group changing. As with the one-dimensional example, the RE persist for $I_1 < 0$ when $c_2 > 0$, and the branch of RE with $\nu > 0$ is hyperbolic.

12 Creating dark solitary waves: an example with $\mathbf{G} = \mathbb{R}^2 \times S^1$

In this section a Hamiltonian system on $\mathbb{R}^8$ is considered which has a three-dimensional symmetry group consisting of a direct product of S^1 and a two-dimensional group of affine translations. This example is a simplified version of a problem that appears in the theory of water waves, where degenerate relative equilibria are associated with the concept of criticality in fluid mechanics and the creation of a new class of dark solitary waves coupled to a mean flow [7].

The following simplifed system of differential equations is a variant of the steady Benney-Roskes equation modelling water waves introduced in [5]

$$\begin{aligned} aA_{xx} + 2\mathrm{i}bA_x + \beta|A|^2A &= -2(\ell h_x + mu_x)A \\ r_1h_{xx} + r_3u_{xx} &= \ell(|A|^2)_x \\ r_3h_{xx} + r_2u_{xx} &= m(|A|^2)_x\,, \end{aligned} \tag{12.1}$$

where a, b, β, ℓ, m, r_1, r_2 and r_3 are given real parameters with $r_1r_2 - r_3^2 \neq 0$ and $\mathrm{i} = \sqrt{-1}$. The function $A(x)$ is complex valued and the functions $h(x)$ and $u(x)$ are real valued.

The first equation in (12.1) is a steady nonlinear Schrödinger equation coupled to mean flow terms and $A(x)$ is associated with spatial modulation of a periodic water wave. The equations for h and u are the shallow water equations coupled to the wave modulation, with $h(x)$ representing the mean depth and $u(x)$ representing the mean velocity. Note that the second and third equations in (12.1) can be integrated and substituted into the first equation. However a geometric approach leads to more information about solutions, and extends to more general systems such as water waves.

From the viewpoint of this paper, this system can be characterized as a Hamiltonian system on $\mathbb{R}^8$ with a 3−dimensional symmetry group. Introduce coordinates $\mathbf{z} = (q_1, \dots, q_4, p_1, \dots, p_4)$, defined by

$$A = q_1 + \mathrm{i}q_2\,, \quad h = q_3 \quad \text{and} \quad u = q_4$$

and

$$p_1 = a\tfrac{dq_1}{dx} - bq_2\,, \quad p_3 = r_1\tfrac{dq_3}{dx} + r_3\tfrac{dq_4}{dx} - \ell(q_1^2 + q_2^2)$$

$$p_2 = a\tfrac{dq_2}{dx} + bq_1\,, \quad p_4 = r_2\tfrac{dq_4}{dx} + r_3\tfrac{dq_3}{dx} - m(q_1^2 + q_2^2)\,.$$

The symplectic form is the standard one, $\Omega = \mathrm{d}q_1 \wedge \mathrm{d}p_1 + \cdots + \mathrm{d}q_4 \wedge \mathrm{d}p_4$, and the Hamiltonian function in these coordinates is

$$\begin{aligned} H = {} & \frac{1}{2a}(p_1^2 + p_2^2) + \frac{r_2}{2\delta}p_3^2 - \frac{r_3}{\delta}p_3p_4 + \frac{r_1}{2\delta}p_4^2 \\ & + \frac{b}{a}(p_1q_2 - p_2q_1) + \frac{b^2}{2a}(q_1^2 + q_2^2) + \tfrac{1}{4}\widetilde{\beta}(q_1^2 + q_2^2)^2 \\ & + \frac{1}{\delta}[(r_2\ell - r_3m)p_3 + (r_1m - r_3\ell)p_4](q_1^2 + q_2^2)\,, \end{aligned}$$

where

$$\delta = r_1r_2 - r_3^2 \quad \text{and} \quad \widetilde{\beta} = \beta - \frac{2}{\delta}\det\begin{bmatrix} r_1 & r_3 & \ell \\ r_3 & r_2 & m \\ \ell & m & 0 \end{bmatrix}\,,$$

The governing equations are now

$$\mathbf{z}_x = X_H(\mathbf{z}) \quad \mathbf{z} \in M := \mathbb{R}^8\,, \quad X_H \,\lrcorner\, \Omega = \mathrm{d}H\,.$$

An arbitrary constant can be added to q_3 and q_4 and this generates the subgroup $\mathbb{R}^2$ of affine translations. The equations (12.1) are equivariant with respect to rotations acting on A. This symmetry leads to the action of S^1 associated with (q_1, q_2) and (p_1, p_2).

The infinitesimal generators of the group acting on M are

$$\begin{aligned} \xi_M^1(\mathbf{z}) &= \tfrac{\partial}{\partial q_3} \\ \xi_M^2(\mathbf{z}) &= \tfrac{\partial}{\partial q_4} \\ \xi_M^3(\mathbf{z}) &= -q_2\tfrac{\partial}{\partial q_1} + q_1\tfrac{\partial}{\partial q_2} - p_2\tfrac{\partial}{\partial p_1} + p_1\tfrac{\partial}{\partial p_2} \end{aligned}$$

Let $\xi_1^*, \xi_2^*, \xi_3^*$ be a basis for $\mathfrak{g}^*$. The momentum map is

$$\mathbf{J}(\mathbf{z}) = J_1(\mathbf{z})\xi_1^* + J_2(\mathbf{z})\xi_2^* + J_3(\mathbf{z})\xi_3^*\,, \tag{12.2}$$

with

$$J_1(\mathbf{z}) = p_3\,, \quad J_2(\mathbf{z}) = p_4 \quad \text{and} \quad J_3(\mathbf{z}) = p_2 q_1 - p_1 q_2\,. \tag{12.3}$$

Relative equilibria associated with this group are of the form

$$\mathbf{z}(x) := \Phi_{\exp(x\xi)}(\varphi(\mathbf{c}))\,, \quad \xi = c_1\xi_1 + c_2\xi_2 + c_3\xi_3\,,$$

with $\varphi(\mathbf{c})$ a critical point of the augmented Hamiltonian

$$\mathrm{d}H(\varphi) = c_1 \mathrm{d}J_1(\varphi) + c_2 \mathrm{d}J_2(\varphi) + c_3 \mathrm{d}J_3(\varphi)\,. \tag{12.4}$$

This system consists of set of simple algebraic equations with solution

$$\varphi(\mathbf{c}) = \begin{pmatrix} \widehat{q}_1(\mathbf{c}) \\ \widehat{q}_2(\mathbf{c}) \\ 0 \\ 0 \\ -(b+ac_3)\widehat{q}_2(\mathbf{c}) \\ (b+ac_3)\widehat{q}_1(\mathbf{c}) \\ (r_1c_1 + r_3c_2) - \ell(\widehat{q}_1(\mathbf{c})^2 + \widehat{q}_2(\mathbf{c})^2) \\ (r_3c_1 + r_2c_2) - m(\widehat{q}_1(\mathbf{c})^2 + \widehat{q}_2(\mathbf{c})^2) \end{pmatrix}$$

where

$$\widehat{q}_1(\mathbf{c})^2 + \widehat{q}_2(\mathbf{c})^2 = \frac{1}{\beta}(ac_3^2 + 2bc_3 - 2\ell c_1 - 2mc_2)\,. \tag{12.5}$$

For real solutions the right-hand side must be positive and this restriction defines $\mathfrak{g}^{\mathrm{RE}}$ by

$$\mathfrak{g}^{\mathrm{RE}} = \{\, \mathbf{c} \in \mathfrak{g} \;:\; \beta(ac_3^2 + 2bc_3 - 2\ell c_1 - 2mc_2) > 0 \,\}\,. \tag{12.6}$$

This set is a semi-algebraic variety in $\mathfrak{g}$ and its boundary consists of points where momentum map is singular. Hence points $\mathbf{c} \in \mathfrak{g}^{\mathrm{RE}}$ automatically satisfy hypothesis (**H4**).

The pullback of the momentum map by $\varphi(\mathbf{c})$ has components

$$\begin{aligned} P_1(\mathbf{c}) &= \left(r_1 + \tfrac{2\ell^2}{\beta}\right) c_1 + \left(r_3 + \tfrac{2m\ell}{\beta}\right) c_2 - \tfrac{\ell}{\beta}(ac_3^2 + 2bc_3) \\ P_2(\mathbf{c}) &= \left(r_3 + \tfrac{2m\ell}{\beta}\right) c_1 + \left(r_2 + \tfrac{2m^2}{\beta}\right) c_2 - \tfrac{m}{\beta}(ac_3^2 + 2bc_3) \\ P_3(\mathbf{c}) &= \tfrac{(b+ac_3)}{\beta}\left(ac_3^2 + 2bc_3 - 2\ell c_1 - 2mc_2\right)\,. \end{aligned}$$

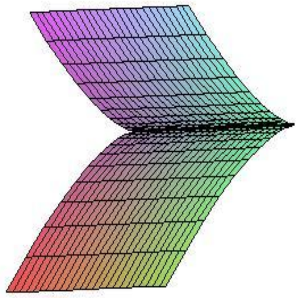

Fig. 7. The image of $\Sigma^1(\mathbf{P})$ in $\mathfrak{g}^*$ for parameter values: $a=1$, $b=-.25$, $\beta=-3$, $\ell=-1$, $m=1.25$, $r_1=-.5$, $r_2=1$ and $r_3=.25$, and $a\widetilde{\beta}=-6<0$

A straightforward calculation then leads to

$$\begin{aligned} \mathrm{D}\mathbf{P}(\mathbf{c}) &= \begin{bmatrix} \frac{\partial P_1}{\partial c_1} & \frac{\partial P_1}{\partial c_2} & \frac{\partial P_1}{\partial c_3} \\ \frac{\partial P_2}{\partial c_1} & \frac{\partial P_2}{\partial c_2} & \frac{\partial P_2}{\partial c_3} \\ \frac{\partial P_3}{\partial c_1} & \frac{\partial P_3}{\partial c_2} & \frac{\partial P_3}{\partial c_3} \end{bmatrix} \\ &= \frac{1}{\beta}\begin{bmatrix} r_1\beta+2\ell^2 & r_3\beta+2\ell m & -2\ell(b+ac_3) \\ r_3\beta+2\ell m & r_2\beta+2m^2 & -2m(b+ac_3) \\ -2\ell(b+ac_3) & -2m(b+ac_3) & 3(b+ac_3)^2-b^2-2\ell ac_1-2mac_2 \end{bmatrix} \end{aligned}$$

and so

$$\det(\mathrm{D}\mathbf{P}(\mathbf{c})) = -\frac{1}{\beta^2}\left(\delta\widetilde{\beta}(2a\ell c_1+2amc_2+b^2)-\delta(2\beta+\widetilde{\beta})(b+ac_3)^2\right).$$

For fixed values of the parameters, $\det(\mathrm{D}\mathbf{P}(\mathbf{c}))=0$ is linear in c_1 and c_2 and quadratic in c_3. Hence the set $\Sigma^1(\mathbf{P})$ is a simple parabola in $\mathfrak{g}$

$$\Sigma^1(\mathbf{P}) = \left\{\mathbf{c}\in\mathfrak{g} \;:\; 2a\ell c_1+2amc_2+b^2-\frac{(2\beta+\widetilde{\beta})}{\widetilde{\beta}}(b+ac_3)^2=0\right\}.$$

Even though $\Sigma^1(\mathbf{P})\cap\mathfrak{g}$ is always non-empty, the set $\Sigma^1(\mathbf{P})\cap\mathfrak{g}^{\mathrm{RE}}$ may be empty. Substituting the definition of $\Sigma^1(\mathbf{P})$ into $\mathfrak{g}^{\mathrm{RE}}$ shows that

$$\Sigma^1(\mathbf{P})\cap\mathfrak{g}^{\mathrm{RE}} \quad \text{is nonempty} \quad \Leftrightarrow \quad a\widetilde{\beta}<0\,.$$

The image of $\Sigma^1(\mathbf{P})$ in $\mathfrak{g}^*$ can have singularities. A typical example is shown in Figure 7. The singularities are in the image of $\mathbf{P}(\mathbf{c})$ when $\mathbf{c}\in\Sigma^{11}(\mathbf{P})$. To compute $\Sigma^{11}(\mathbf{P})$, the kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$ and the function $\mathscr{K}(\mathbf{c},t)$ need to be computed.

The kernel of $\mathrm{D}\mathbf{P}(\mathbf{c})$ can be taken to be a column of the adjugate matrix, since,

$$\boldsymbol{\eta} = \begin{pmatrix} \frac{\partial P_1}{\partial c_2}\frac{\partial P_2}{\partial c_3} - \frac{\partial P_1}{\partial c_3}\frac{\partial P_2}{\partial c_2} \\ \frac{\partial P_1}{\partial c_3}\frac{\partial P_2}{\partial c_1} - \frac{\partial P_1}{\partial c_1}\frac{\partial P_2}{\partial c_3} \\ \frac{\partial P_1}{\partial c_1}\frac{\partial P_2}{\partial c_2} - \frac{\partial P_1}{\partial c_2}\frac{\partial P_2}{\partial c_1} \end{pmatrix} \quad \Rightarrow \quad \mathrm{D}\mathbf{P}(\mathbf{c})\,\boldsymbol{\eta} = \det[\mathrm{D}\mathbf{P}(\mathbf{c})] \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{12.7}$$

Calculating,

$$\boldsymbol{\eta} = \mathsf{Constant} \begin{pmatrix} 2(b+ac_3)(\ell r_2 - m r_3) \\ 2(b+ac_3)(m r_1 - \ell r_3) \\ \delta\widetilde{\beta} \end{pmatrix}.$$

Use this expression to compute the curvature

$$\left.\frac{d^2}{dt^2}\right|_{t=0} \mathscr{K}(\mathbf{c},t) = \frac{a}{\beta}\delta^3\widetilde{\beta}^2(\widetilde{\beta}+5\beta)(b+ac_3)\,.$$

Hence, if the parameters a, δ, $\widetilde{\beta}$ and $(\widetilde{\beta}+5\beta)$ are nonzero, then $\mathbf{c}\in\Sigma^1(\mathbf{P})\setminus\Sigma^{11}(\mathbf{P})$ when $b+ac_3\neq 0$. On the other hand,

$$\begin{aligned} \Sigma^{11}(\mathbf{P}) &= \{\, \mathbf{c}\in\Sigma^1(\mathbf{P}) \ :\ \mathrm{Ker}(\mathrm{D}\mathbf{P}(\mathbf{c})) \subset T_c\Sigma^1(\mathbf{P})\,\} \\ &= \left\{ \mathbf{c}\in\mathfrak{g} \ :\ c_3 = -\tfrac{b}{a} \quad \text{and} \quad \ell c_1 + m c_2 = -\tfrac{b^2}{2a} \right\}. \end{aligned}$$

The set $\Sigma^{11}(\mathbf{P})$ is a line in $\mathfrak{g}$ and the image of this line in $\mathfrak{g}^*$ is a curve of cusp points. An example is shown in Figure 7. Note that $\Sigma^{11}(\mathbf{P})$ also corresponds with the boundary of $\mathfrak{g}^{\mathrm{RE}}$ and so consists of singular values of the momentum map.

For all $\mathbf{c}\in\Sigma^1(\mathbf{P})\setminus\Sigma^{11}(\mathbf{P})$ the theory of this paper applies to give the local existence of a branch of homoclinic orbits which for the model (12.1) correspond to dark solitary waves coupled to a mean flow. Details of solutions of this type are given in [7].

— Appendix —

A Properties of bordered matrices

In this appendix some elementary properties of bordered matrices are recorded for use in the proof of Lemma 2.1 in §2. The basic ideas can be found in Magnus & Neudecker [15], Greenberg, Maddocks & Rogers [11] and references therein.

Let $\mathbf{A}$ be an $m \times m$ symmetric matrix, and let $\mathbf{b}$ be an $m \times 1$ matrix. Then

$$\det \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & 0 \end{bmatrix} = -\mathbf{b}^T \mathbf{A}^{\#} \mathbf{b}\,, \tag{A-1}$$

where $\mathbf{A}^{\#}$ is the *adjugate* of $\mathbf{A}$ (transpose of the cofactor matrix of $\mathbf{A}$). This is Theorem 4 on page 43 of [15].

Suppose $\mathbf{A}$ has a simple zero eigenvalue with eigenvector $\mathbf{v}$, then

$$\mathbf{A}^{\#} = \Pi \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\,, \tag{A-2}$$

where Π is the product of the non-zero eigenvalues of $\mathbf{A}$. This is Theorem 3 on page 41 of [15].

Combining these two results: suppose $\mathbf{A}$ has a simple zero eigenvalue with eigenvector $\mathbf{v}$, then

$$\det \begin{bmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{v}^T & 0 \end{bmatrix} = -\Pi \, \|\mathbf{v}\|^2\,. \tag{A-3}$$

Now, suppose that the kernel of $\mathbf{A}$ has dimension n with $n < m$ and let Ker be the $m \times n$ matrix whose columns are the normalized eigenvectors, and so $\mathsf{Ker}^T\mathsf{Ker} = \mathbf{I}_n$, then a straightforward generalization of (A-3) is

$$\det \begin{bmatrix} \mathbf{A} & \mathsf{Ker} \\ \mathsf{Ker}^T & 0 \end{bmatrix} = -\Pi\,, \tag{A-4}$$

where Π is the product of the $m - n$ non-zero eigenvalues of $\mathbf{A}$.

The main result needed in the characterization of non-degeneracy of a critical point is the following.

Proposition A.1. *Suppose* $\mathbf{A}$ *is an* $m \times m$ *symmetric matrix with kernel of dimension* n *where* $n < m$. *Let* Ker *be the* $m \times n$ *matrix whose columns are the normalized eigenvectors spanning* $\mathrm{Ker}(\mathbf{A})$. *Suppose* $\mathbf{B}$ *is an* $m \times n$ *matrix of rank* n *satisfying* $\mathsf{Ker}^T\mathbf{B} = 0$. *Then*

$$\det \begin{bmatrix} \mathbf{A} & \mathsf{Ker} & \mathbf{B} \\ \mathsf{Ker}^T & \mathbf{0} & \mathbf{0} \\ \mathbf{B}^T & \mathbf{0} & \mathbf{0} \end{bmatrix} = (-1)^{n+1}\, \Pi \det(\mathbf{B}^T\mathbf{W})\,, \tag{A-5}$$

where Π *is the product of the non-zero eigenvalues of* $\mathbf{A}$ *and* $\mathbf{W}$ *is the unique* $m \times n$ *matrix satisfying*

$$\mathbf{A}\,\mathbf{W} = \mathbf{B} \quad \text{and} \quad \mathsf{Ker}^T\mathbf{W} = \mathbf{0}\,. \tag{A-6}$$

Proof. The condition $\mathsf{Ker}^T\mathbf{B} = 0$ assures that the columns of $\mathbf{B}$ are in the range of $\mathbf{A}$ and so (A-6) is solvable, and the second condition in (A-6) assures that $\mathbf{W}$ unique. Define

$$\widehat{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathsf{Ker} \\ \mathsf{Ker}^T & \mathbf{0} \end{bmatrix}, \quad \widehat{\mathbf{B}} = \begin{pmatrix} \mathbf{B} \\ \mathbf{0} \end{pmatrix} \quad \text{and} \quad \widehat{\mathbf{W}} = \begin{pmatrix} \mathbf{W} \\ \mathbf{0} \end{pmatrix}.$$

Then $\widehat{\mathbf{A}}$ is invertible by (A-4) and

$$\det \begin{bmatrix} \widehat{\mathbf{A}} & \widehat{\mathbf{B}} \\ \widehat{\mathbf{B}}^T & \mathbf{0} \end{bmatrix} = \det \left\{ \begin{pmatrix} \widehat{\mathbf{A}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \widehat{\mathbf{B}}^T & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & \widehat{\mathbf{W}} \\ \mathbf{0} & -\mathbf{B}^T\mathbf{W} \end{pmatrix} \right\} = \det(\widehat{\mathbf{A}}) \det(-\mathbf{B}^T\mathbf{W}).$$

since (A-6) is equivalent to $\widehat{\mathbf{W}} = \widehat{\mathbf{A}}^{-1}\widehat{\mathbf{B}}$. The required result now follows using (A-4).■

B Williamson normal form for the linearization about degenerate RE

In this appendix, an explicit proof of Lemma 5.1 is given. The starting point is the pair of linear operators $(\mathbf{L}, \mathbb{J})$ (suppressing the dependence on $\mathbf{c}$) with $\mathbb{J}$ in standard form (2.1) and $\mathbf{L}$ having a Jordan chain of length $2n+2$; specifically

$$\mathbf{L}\mathbf{v}_j = 0 \quad \text{and} \quad \mathbf{L}\mathbf{v}_{n+j} = \mathbb{J}\mathbf{v}_j, \quad j = 1, \ldots, n,$$

$$\mathbf{L}\mathbf{v}_{2n+1} = \sum_{j=1}^{n} \eta_j \mathbb{J}\mathbf{v}_{n+j} \quad \text{and} \quad \mathbf{L}\mathbf{v}_{2n+2} = \mathbb{J}\mathbf{v}_{2n+1},$$

for some $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)$ with

$$\sum_{j=1}^{n} \eta_j \, [\mathbb{J}\mathbf{v}_{2n+2}, \mathbf{v}_j] \neq 0.$$

The object is to construct a transformation matrix $\mathbf{F}$ such that

$$\mathbf{F}^T\mathbb{J}\mathbf{F} = \mathbb{J} \quad \text{and} \quad \mathbf{F}^T\mathbf{L}\mathbf{F} = \mathbf{L}^{\text{ref}}, \tag{B-1}$$

where $\mathbf{L}^{\text{ref}}$ is defined in (5.1), with the symplectic signs defined in Lemma 5.1.

The symplectic transformation $\mathbf{F}$ is decomposed into four parts

$$\mathbf{F} = \mathbf{F}_1\mathbf{F}_2\mathbf{F}_3\mathbf{F}_4,$$

with $\mathbf{F}_1$ being the matrix whose columns are the eigenvectors of $\mathbf{L}$

$$\mathbf{F}_1 = \text{col}(\mathbf{v}_1, \cdots, \mathbf{v}_{2n+2}).$$

This first transformation results in

$$\mathbf{L}\mathbf{F}_1 = \mathbb{J}\,\mathbf{F}_1\mathbf{N}\,, \quad \mathbf{N} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_n & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\eta} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$\mathbf{F}_1^T \mathbb{J}\mathbf{F}_1 := \mathbf{K} = \begin{bmatrix} \mathbf{0} & -\mathrm{D}\mathbf{P}(\mathbf{c}) & 0 & \boldsymbol{\sigma}_1 \\ \mathrm{D}\mathbf{P}(\mathbf{c}) & \widehat{\mathbf{K}} & -\boldsymbol{\sigma}_1 & \boldsymbol{\sigma}_2 \\ 0 & \boldsymbol{\sigma}_1^T & 0 & \alpha \\ -\boldsymbol{\sigma}_1^T & -\boldsymbol{\sigma}_2{}^T & -\alpha & 0 \end{bmatrix}, \tag{B-2}$$

where

$$\boldsymbol{\sigma}_1 = -\begin{pmatrix} \Omega(\mathbf{v}_1, \mathbf{v}_{2n+2}) \\ \vdots \\ \Omega(\mathbf{v}_n, \mathbf{v}_{2n+2}) \end{pmatrix}, \quad \boldsymbol{\sigma}_2 = -\begin{pmatrix} \Omega(\mathbf{v}_{n+1}, \mathbf{v}_{2n+2}) \\ \vdots \\ \Omega(\mathbf{v}_{2n}, \mathbf{v}_{2n+2}) \end{pmatrix}, \quad \alpha = -\Omega(\mathbf{v}_{2n+1}, \mathbf{v}_{2n+2})\,,$$

and $\widehat{\mathbf{K}}$ is the $n \times n$ skew-symmetric matrix with entries

$$\widehat{\mathbf{K}}_{ij} = -\Omega(\mathbf{v}_{n+i}, \mathbf{v}_{n+j})\,, \quad i, j = 1, \ldots, n\,.$$

The form of the entries in $\mathbf{K}$ in (B-2) follow from skew-symmetry of $\mathbf{K}$ and Proposition 4.3. The symplectic operator $\mathbf{K}$ is not in canonical form. The purpose of the three transformations $\mathbf{F}_2, \mathbf{F}_3$ and $\mathbf{F}_4$ is to transform $\mathbf{K}$ into $\mathbb{J}$ while preserving the form of the transformed Jordan chain.

The transformation $\mathbf{F}_2$ is chosen to diagonalize $\mathrm{D}\mathbf{P}(\mathbf{c})$ in $\mathbf{K}$. Let $\mathbf{T}$ be the $n \times n$ orthogonal matrix which diagonalizes $\mathrm{D}\mathbf{P}(\mathbf{c})$ with the first column taken to be $\boldsymbol{\eta}$. Then

$$\mathbf{T}^T \mathrm{D}\mathbf{P}(\mathbf{c})\mathbf{T} = \mathrm{diag}[0, \lambda_2, \ldots, \lambda_n]\,,$$

and $\mathbf{F}_2$ can be taken to be the block diagonal matrix

$$\mathbf{F}_2 = \begin{bmatrix} \mathbf{T} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} & \mathbf{0} \\ 0 & 0 & \mathbf{I}_2 \end{bmatrix}.$$

Define these rotated eigenvectors by $\mathbf{w}_1, \ldots, \mathbf{w}_{2n+2}$; in matrix notation

$$\mathbf{F}_1\mathbf{F}_2 := \mathrm{col}(\mathbf{w}_1, \ldots, \mathbf{w}_{2n+2})\,.$$

The modified Jordan chain is

$$
\begin{gathered}
\mathbf{L}\mathbf{w}_j = 0 \quad \text{and} \quad \mathbf{L}\mathbf{w}_{n+j} = \mathbb{J}\mathbf{w}_j\,, \quad j = 1, \dots, n\,, \\
\mathbf{L}\mathbf{w}_{2n+1} = \mathbb{J}\mathbf{w}_{n+1} \quad \text{and} \quad \mathbf{L}\mathbf{w}_{2n+2} = \mathbb{J}\mathbf{w}_{2n+1}\,.
\end{gathered}
\tag{B-3}
$$

The object is to introduce new vectors $\widehat{\mathbf{w}}_1, \dots, \widehat{\mathbf{w}}_{2n+2}$ with the same Jordan chain structure but resulting in a canonical symplectic form that is a permutation of $\mathbb{J}$.

The transformation matrix $\mathbf{F}_3$ is the one which does the most work. It transforms the $\{\mathbf{w}_j\}$ set to the set $\{\widehat{\mathbf{w}}_j\}$. Let

$$
\mathbf{F}_3 = \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{m} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} & \mathbf{m} \\ 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & a_1 \end{bmatrix} .
$$

where $\mathbf{A} = \text{diag}[a_1, \dots, a_n]$, $\mathbf{m}$ is an $n-$vector and $\mathbf{B}$ is a nilpotent upper triangular matrix. The following expressions are found for the entries of $\mathbf{F}_3$,

$$
a_1 = \frac{1}{|\Omega(\mathbf{w}_1, \mathbf{w}_{2n+2})|^{1/2}} \quad \text{and} \quad s_1 = -\text{sign}\left(\Omega(\mathbf{w}_1, \mathbf{w}_{2n+2})\right)\,, \tag{B-4}
$$

$$
a_j = \frac{1}{\sqrt{|\lambda_j|}} \quad \text{and} \quad s_j = \text{sign}\left(\lambda_j\right)\,, \quad j = 2, \dots, n\,. \tag{B-5}
$$

$$
m_j = \begin{cases} -\frac{1}{2} s_1 a_1 \left(\alpha a_1^2 + \sum_{k=2}^{n} \lambda_k m_k^2\right) & j = 1 \\ -\dfrac{a_1}{\lambda_j}\Omega(\mathbf{w}_j, \mathbf{w}_{2n+2}) & j \geq 2\,, \end{cases} \tag{B-6}
$$

where $\alpha = -\Omega(\mathbf{w}_{2n+1}, \mathbf{w}_{2n+2})$. The first row of $\mathbf{B}$ has entries

$$
B_{1j} = s_1 a_1^2 a_j \left(\Omega(\mathbf{w}_{n+j}, \mathbf{w}_{2n+2}) + \sum_{k=2}^{n} \frac{1}{\lambda_k}\Omega(\mathbf{w}_k, \mathbf{w}_{2n+2})\Omega(\mathbf{w}_{n+j}, \mathbf{w}_{n+k}) \right)\,, \quad j \geq 2\,.
$$

Note that B_{11} is identically zero in this definition, since $\Omega(\mathbf{w}_{n+1}, \mathbf{w}_{2n+2}) = 0$ and $\Omega(\mathbf{w}_{n+1}, \mathbf{w}_{n+k}) = 0$ for $k = 1, \dots, n$. The other rows of $\mathbf{B}$ have entries

$$
B_{ij} = \frac{a_j}{\lambda_j}\Omega(\mathbf{w}_{n+i}, \mathbf{w}_{n+j})\,, \quad i \geq 2 \quad \text{and} \quad j > i\,.
$$

With

$$
\mathbf{F}_1\mathbf{F}_2\mathbf{F}_3 := \text{col}(\widehat{\mathbf{w}}_1, \dots, \widehat{\mathbf{w}}_{2n+2})\,.
$$

the new vectors have the form

$$\begin{aligned}
\widehat{\mathbf{w}}_j &= a_j\,\mathbf{w}_j\,, \quad j=1,\ldots,n \\
\widehat{\mathbf{w}}_{n+j} &= a_j\,\mathbf{w}_{n+j} + \textstyle\sum_{k=1}^n B_{kj}\mathbf{w}_k\,, \quad j=1,\ldots,n \\
\widehat{\mathbf{w}}_{2n+1} &= a_1\mathbf{w}_{2n+1} + \textstyle\sum_{k=1}^n m_k\mathbf{w}_k \\
\widehat{\mathbf{w}}_{2n+2} &= a_1\mathbf{w}_{2n+2} + \textstyle\sum_{k=1}^n m_k\mathbf{w}_{n+k}\,.
\end{aligned} \tag{B-7}$$

These vectors satisfy the same Jordan chain equations as the $\{\mathbf{w}_j\}$ in (B-3).

The fourth matrix $\mathbf{F}_4$ is just a permutation matrix which puts the plus and minus ones in the correct slots using the symplectic signs $s_1,\ldots,s_n$. It is defined by

$$\mathbf{F}_4 = \begin{bmatrix} \mathbf{I}_{n+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^T \end{bmatrix} \quad \text{with} \quad \mathbf{S} = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 & -s_1 \\ s_2 & 0 & \cdots & & \cdots & 0 \\ 0 & s_3 & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & s_n & 0 & 0 \\ 0 & \cdots & \cdots & 0 & s_1 & 0 \end{bmatrix},$$

with s_1 defined in (5.2) and $s_j = \mathrm{sign}(\lambda_j)$ for $j=2,\ldots,n$. The complete transformation $\mathbf{F}$ satisfying (B-1) is then

$$\mathbf{F} = \mathbf{F}_1\mathbf{F}_2\mathbf{F}_3\mathbf{F}_4 := \mathrm{col}\left(\widehat{\mathbf{w}}_1,\ldots,\widehat{\mathbf{w}}_{n+1}, -s_1\widehat{\mathbf{w}}_{2n+2}, s_2\widehat{\mathbf{w}}_{n+2},\ldots,s_n\widehat{\mathbf{w}}_{2n}, s_1\widehat{\mathbf{w}}_{2n+1}\right)\,.$$

Acknowledgements

Helpful discussions with Neil Donaldson, James Montaldi and Mark Roberts are gratefully acknowledged.

References

[1] V.I. Arnold. *Matrices depending on parameters*, Russ. Math. Surveys **26** 29–43 (1970).

[2] V.I. Arnold, S.M. Gusein-Zade & A.N. Varchenko. *Singularities of Differentiable Maps, Volume I*, Birkhäuser: Boston (1985).

[3] V.I. Arnol'd, V.V. Kozlov & A.I. Neishtadt. *Mathematical aspects of classical and celestial mechanics*, Dynamical systems III, Encycl. Math. Sci. **3**, Springer: Berlin (1993).

[4] T.J. Bridges. *Superharmonic instability, homoclinic torus bifurcation and water-wave breaking*, J. Fluid Mech. **505** 153–162 (2004).

[5] T.J. Bridges. *Steady dark solitary waves emerging from wave-generated meanflow: the role of modulation equations*, Chaos **15** 037113 (2005).

[6] T.J. Bridges & N.M. Donaldson. *Degenerate periodic orbits and homoclinic torus bifurcation*, Phys. Rev. Lett. **95**(10) 104301 (2005).

[7] T.J. Bridges & N.M. Donaldson. *Secondary criticality of water waves. Part 1. Definition, bifurcation and solitary waves*, J. Fluid Mech. **565** 381–417 (2006).

[8] T.J. Bridges & N.M. Donaldson. *Reappraisal of criticality for two-layer flows and its implication for internal solitary waves*, Phys. Fluids **19** 072111 (2007).

[9] H.W. Broer, S.-N. Chow, Y.-I. Kim & G. Vegter. *The Hamiltonian double-zero eigenvalue*, Fields Inst. Comm. **4**, 1–19 (1995).

[10] M. Golubitsky & V. Guillemin. *Stable Mappings and Their Singularities*, Springer-Verlag: New York (1973).

[11] L. Greenberg, J.H. Maddocks & K.A. Rogers. *The bordered operator and the index of a constrained critical point*, Math. Nachr. **219** 109–124 (2000).

[12] A. Ibort & C. Martínez Ontalba. *Periodic orbits of Hamiltonian systems and symplectic reduction*, J. Phys. A: Math. Gen **29** 675–687 (1996).

[13] G. Iooss & M. Adelmeyer. *Topics in Bifurcation Theory and Applications*, World Scientific: Singapore (1992).

[14] G. Iooss & E. Lombardi. *Polynomial normal forms with exponentially-small remainder for analytic vectorfields*, J. Diff. Eqns. **212** 1–61 (2005).

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

[16] J.E. Marsden. *Lectures on Mechanics*, London Mathematical Society Lecture Notes **174**, Cambridge University Press (1992).

[17] J.E. Marsden, R. Montgomery & T.S. Ratiu. *Reduction, Symmetry and Phases in Mechanics*, American Mathematical Society Memoirs **436**, AMS Providence (1990).

[18] J.E. Marsden & T.S. Ratiu. *Introduction to Mechanics and Symmetry*, Springer-Verlag: New York (1994).

[19] I. Melbourne & M. Dellnitz. *Normal forms for linear Hamiltonian vector fields commuting with the action of a compact Lie group*, Math. Proc. Camb. Philos. Soc. **114** 235–268 (1993).

[20] K.R. Meyer. *Periodic solutions of the $N-$body problem*, J. Diff. Eqns. **39** 2–38 (1981).

[21] K.R. Meyer & G. Hall. *Introduction to Hamiltonian Dynamical Systems and the $N-$body problem*, Appl. Math. Sci. **90**, Springer: New York (1992).

[22] A. Mielke. *Hamiltonian and Lagrangian flows on center-manifolds with applications to elliptic variational problems*, Lect. Notes Math. **1489**, Springer-Verlag: Berlin (1991).

[23] J. Montaldi. *Symmetric Hamiltonian bifurcations*, in *Peyresq Lectures on Geometric Mechanics and Symmetry*, LMS Lecture Notes **306** 357–402, CUP: Cambridge (2005).

[24] J. Montaldi. *Persistence and stability of relative equilibria*, Nonlinearity **10** 449–466 (1997).

[25] J.P. Ortega. *Relative normal modes for nonlinear Hamiltonian systems*, Proc. Roy. Soc. Edin. A **133** 665–704 (2003).

[26] J.P. Ortega & T.S. Ratiu. *Momentum Maps and Hamiltonian Reduction*, Birkhäuser: Boston (2004).

[27] J.I. Palmore. *Measure of degenerate relative equilibria. I*, Annals of Math. **104** 421–429 (1976).

[28] G.W. Patrick. *Relative equilibria of Hamiltonian systems with symmetry: linearization, smoothness, and drift*, J. Nonlin. Sci. **5** 373–418 (1995).

[29] G.W. Patrick & R.M. Roberts. *The transversal relative equilibria of a Hamiltonian system with symmetry*, Nonlinearity **13** 2089–2105 (2000).

[30] I.R. Porteous. *Simple singularities of maps*, in Proc. Liverpool Singularities Symposium, pp. 286–307. Lecture Notes in Math. **192** Springer-Verlag: Berlin (1971).

[31] I.R. Porteous. *Geometric Differentiation*, Second Edition, Cambridge University Press (2001).

[32] R.M. Roberts, C. Wulff & J.S.W. Lamb. *Hamiltonian systems near relative equilibria*, J. Diff. Eqns **179** 562–604 (2002).

[33] S. Smale. *Topology and Mechanics. II The planar $N-$ body problem*, Invent. Math. **11**, 45–64 (1970).

[34] A. Weinstein. *Bifurcations and Hamilton's principle*, Math. Z. **159** 235–248 (1978).

[35] J. Williamson. *On an algebraic problem concerning the normal forms of linear dynamical systems*, Amer. J. Math. **58** 141–163 (1936).

[36] C. Wulff. *Persistence of relative equilibria in Hamiltonian systems with non-compact symmetry*, Nonlinearity **16** 67–91 (2003).

[37] J.A. Zufiria & P.G. Saffman. *An example of stability exchange in a Hamiltonian wave system*, Stud. Appl. Math. **74** 85–91 (1986).

[38] N.T. Zung. *Convergence versus integrability in Birkhoff normal form*, Annals of Math. **161** 141–156 (2005).