JOURNAL OF FUNCTIONAL ANALYSIS 94, 308–348 (1990)

Stability Theory of Solitary Waves in the Presence of Symmetry, II*

MANOUSSOS GRILLAKIS

Department of Mathematics, University of Maryland, College Park, Maryland 20742

JALAL SHATAH

Courant Institute, New York University, New York, New York 10012

AND

WALTER STRAUSS

Department of Mathematics, Brown University, Providence, Rhode Island 02912

Communicated by the Editors

Received March 28, 1989

Consider an abstract Hamiltonian system which is invariant under a group of operators. We continue to study the effect of the group invariance on the stability of solitary waves. Applications are given to bound states and traveling wave solutions of nonlinear wave equations. © 1990 Academic Press, Inc

1. INTRODUCTION

This paper is a continuation of [1]. We consider the abstract Hamiltonian system

$$\frac{du}{dt} = JE'(u(t)), \qquad u(t) \in X,$$

* This work was supported in part by the National Science Foundation under Grants DMS-87-22331 and DMS-86-00247, by an IBM Visiting Members Grant at the Courant Institute, and by the Sloan Foundation.

0022-1236/90 $3.00

which is invariant under the action T of a Lie group G. By a "solitary wave" or "bound state" we mean a solution $z(t)$ of (1.1) whose time evolution is given by a one-parameter subgroup of G:

$$z(t) = T(e^{t\omega})\,\varphi_\omega,$$

where ω belongs to the Lie algebra $\mathfrak{g}$ and $\varphi_\omega \in X$. If K is a subgroup of G, we say this solitary wave is *K-stable* if a solution $u(t)$ of (1.1) exists for all $t \geqslant 0$ and forever remains near the orbit $\{T(g)\,\varphi_\omega \mid g \in K\}$ in the norm of X provided its initial datum $u(0)$ is sufficiently close to φ_ω.

Central to our analysis, as in [1], are the "charge" functionals Q_σ for $\sigma \in \mathfrak{g}$, defined by $Q'_\sigma = J^{-1}T_\sigma$, where T_σ denotes the differential of T. (See Section 2 for a precise definition.) We define the linear operator $H_\omega = E''(\varphi_\omega) - Q''_\omega(\varphi_\omega)$ and the scalar $d(\omega) = E(\varphi_\omega) - Q_\omega(\varphi_\omega)$. Under appropriate assumptions (see Section 2), we have the following main results.

STABILITY THEOREM. *Given $\omega \in \mathfrak{g}$, consider the scalar function d restricted to the centralizer $\mathfrak{g}_\omega$. Assume that it is non-degenerate at ω and let $p(d'')$ be the number of positive eigenvalues of its Hessian at ω. Let $n(H)$ be the number of negative eigenvalues of H_ω. Then $p(d'') \leqslant n(H)$. If $p(d'') = n(H)$, then the solitary wave is G_ω-stable, where G_ω is the centralizer subgroup.*

INSTABILITY THEOREM. *If d is non-degenerate at ω and $n(H) - p(d'')$ is odd, then the solitary wave is G_ω-unstable.*

The linearized equation around the solitary wave (1.2) is $dv/dt = JH_\omega v$, with generator JH_ω. Our assumptions are not on this generator, but rather on the linearized Hamiltonian H_ω.

The oddness assumption in the instability theorem is necessary. For instance, take a pair of harmonic oscillators with the zero solution as a stable center, with $X = \mathbb{R}^4$ and G trivial so that $p(d'') = 0$, but $n(H) = p(H) = 2$.

In [1] we studied one-dimensional groups $G = \mathbb{R}^1$ or S^1. In that case d'' is a scalar function on $g = \mathbb{R}$, so that $p(d'')$ is either zero (if $d''(\omega) \leqslant 0$) or one (if $d''(\omega) > 0$). The non-degeneracy means that $d'' \neq 0$. The Stability Theorem asserts that we have stability if either $n(H) = 0$ or else $n(H) = 1$ with $d''(\omega) > 0$. The Instability Theorem asserts that we have instability if $n(H) = 1$ with $d''(\omega) < 0$. Thus we recover the main results of [1]. Note, however, that the present paper allows $n(H) > 1$.

In Section 2 we set up the basic framework of this paper. In Section 3 we define the key operator in the analysis, the reduced Hamiltonian, and we relate the number of its negative eigenvalues to that of the full Hamiltonian H_ω. Section 4 is devoted to the Stability Theorem, as well as to its generalization to the degenerate case. The proof follows the same main lines as in [1].

Sections 5 and 6 are devoted to the Instability Theorem. This result is much more general than in [1] because $n(H) - p(d'')$ is allowed to be odd rather than just equal to one. This necessitates a completely different proof of instability from that of [1]. In Section 5 we prove the "linearized instability" and in Section 6 we deduce the "nonlinear" instability (as defined above). Specifically, we use a topological argument to show that the generator JH_ω has a pair of real eigenvalues $\pm\lambda$ (Theorem 5.1). (A much simpler proof can be made if $n(H) - p(d'') = 1$.) We also obtain an upper bound on the number of eigenvalues of JH_ω off the imaginary axis. In Section 6 we deduce the instability with respect to G_ω, as well as with respect to the whole group G in certain cases.

The rest of the paper is devoted to several applications to nonlinear wave equations. In Section 7 we consider a coupled system $u_{tt} - u_{xx} + f(u) = 0$ with $u \in \mathbb{R}^3$. We saw in Example A of [1] that traveling waves $u = \varphi(x - ct)$ are never stable. Here we construct waves of the form $u = \exp(tS)\,\varphi(x - ct)$ for skew-symmetric matrices S and show they are stable for certain choices of S and c. The group G consists of the rotations in u and the translations in x. Thus a rotation in the dependent variables can be a stability mechanism.

In Section 8 we consider harmonic maps from a Lorentz manifold $S^1 \times \mathbb{R}$ into S^2. We show that a plane wave solution $\exp[A(k\vartheta + \omega t)]v$ is stable if $\omega^2 > k^2$ but unstable if $\omega^2 < k^2$. Our technique is to embed S^2 in $\mathbb{R}^3$, thereby approximating the problem by a system of the type of Section 7.

In a similar way, in Section 9 we consider solutions of the form $u = \exp(tS)\,\varphi(x - ct)$ of a system of coupled nonlinear Schrödinger equations. In Section 10 we resume our discussion of the optical wave guide by considering states with higher nodal structure. Some of these states are unstable because $n(H) > 1$ is odd. More elaborate examples can be constructed. A list of errata to [1] is added at the end of this paper.

Many of our results were obtained [8] as early as 1985. A number of references to related work have been given in [1]. Some more recent references are the following: In [7], Oh considers the stability theory for Hamiltonian systems on finite-dimensional manifolds. In [5] Jones uses a dynamical systems approach to instability. In [3] Grillakis proves instability if $d''(\omega) < 0$ and G is one-dimensional, no matter what $n(H)$ is. In [4] Grillakis analyzes conditions under which a pair of eigenvalues of JH can bifurcate off the imaginary axis. In [2] the theory is applied to logarithmic and other singular nonlinearities.

We thank C. Jones, Y.-G. Oh, and M. Weinstein for their interest in this work.

2. Formulation

Let X, J, and E be the same as in [1]. For the sake of completeness we repeat the assumptions here. Let X be a real Hilbert space with inner product (,) and dual space X^*. Let $I: X \to X^*$ be the natural isomorphism defined by $\langle Iu, v \rangle = (u, v)$. We shall identify X^{**} with X. Adjoints will refer to the pairing $\langle\ ,\ \rangle$ not the inner product (,).

Let $J: D(J) \to X$ be a linear operator with dense domain $D(J) \subset X^*$. We assume that J is one–one, onto, and skew-symmetric (with respect to $\langle\ ,\ \rangle$). (In [1] we did not assume it was one–one.) Let $E: X \to \mathbb{R}$ be a C^2 functional defined on all of X.

Let G be a finite-dimensional Lie group with Lie algebra $\mathfrak{g}$. We write $e^{\omega} = \exp(\omega)$ for $\omega \in \mathfrak{g}$. For any $\omega \in \mathfrak{g}$, let

$$\mathfrak{g}_{\omega} = \{\sigma \in \mathfrak{g} \mid [\sigma, \omega] = 0\}$$

and let G_{ω} be the subgroup generated by the centralizer $\mathfrak{g}_{\omega}$. Let T be a unitary representation of G on X. Thus, for all $g \in G$, $T(g)$ is a unitary operator on X, which in our notation means that $T^*(g)I = IT(g^{-1})$. We assume that E is invariant under the group action; that is,

$$E(T(g)u) = E(u) \qquad \text{for} \quad g \in G,\ u \in X. \tag{2.1}$$

Differentiation leads to the identities

$$T^*(g)\, E'(T(g)u) = E'(u) \qquad \text{and} \qquad T^*(g) E''(T(g)u)\, T(g) = E''(u). \tag{2.2}$$

The differential of T maps $\mathfrak{g}$ into the skew-adjoint operators on X. We denote it by $\omega \to T_{\omega} = dT_e\omega(e)$. Regarding $\mathfrak{g}$ as the set of left-invariant vector fields, we have $(d/dt)\, T(\exp(t\omega)) = T(\exp(t\omega))\, T_{\omega}$. Thus each T_{ω} is a skew-adjoint operator on X (with a dense domain). Differentiation of (2.1) with respect to g leads to

$$\langle E'(u), T_{\omega}u \rangle = 0 \qquad \text{for} \quad u \in D(T_{\omega}). \tag{2.3}$$

Assume that

$$T(g)J = JT^*(g^{-1}) \qquad \text{for} \quad g \in G. \tag{2.4}$$

Equivalently, $J^{-1}T(g) = T^*(g^{-1})\, J^{-1}$. Differentiating, we get $J^{-1}T_{\omega} = -T_{\omega}^* J^{-1}$; whence $(J^{-1}T_{\omega})^* = -(J^{-1})^* T_{\omega}^{**} = J^{-1}T_{\omega}$. Thus $J^{-1}T_{\omega}$ is a symmetric operator with domain $D(T_{\omega})$. Assume that

$$J^{-1}T_{\omega} \text{ extends to a bounded operator } B_{\omega}: X \to X^* \tag{2.5}$$

for every $\omega \in \mathfrak{g}$. Then B_ω is also symmetric. Define

$$Q_\omega(u) = \tfrac{1}{2}\langle B_\omega u, u\rangle \qquad \text{for} \quad u \in X, \omega \in \mathfrak{g},$$

so that $Q'_\omega(u) = B_\omega u$. Assume that

$$Q_\omega(T(g)u) = Q_\omega(u) \qquad \text{for} \quad g \in G_\omega, u \in X.$$

Remark. Probably (2.7) is a consequence of the previous assı due to the following argument. Since $T_\omega T_\sigma - T_\sigma T_\omega = T_{[\omega,\sigma]}$, $JA = 0$, where

$$A = B_\omega J B_\sigma - B_\sigma J B_\omega - B_{[\omega,\sigma]}.$$

Hence

$$\begin{aligned}\left.\frac{d}{ds}\right|_{s=0} Q_\omega(T(e^{s\sigma})u) &= \langle B_\omega u, T_\sigma u\rangle \\ &= \langle B_\omega u, J B_\sigma u\rangle \\ &= \tfrac{1}{2}\langle B_\omega u, J B_\sigma u\rangle - \tfrac{1}{2}\langle B_\sigma u, J B_\omega u\rangle \\ &= \tfrac{1}{2}\langle u, B_{[\omega,\sigma]} u\rangle + \tfrac{1}{2}\langle u, Au\rangle.\end{aligned}$$

Letting $v = J^{-1}u$, we have $\langle u, Au\rangle = \langle Jv, Au\rangle = \langle v, JAu\rangle = 0$. H

$$\left.\frac{d}{ds}\right|_{s=0} Q_\omega(T(e^{s\sigma})u) = Q_{[\omega,\sigma]}(u),$$

which imples (2.7). This argument is correct provided u belon domains of the appropriate operators. Whether these domains hav intersection is the only questionable point in the argument.

Differentiation of (2.7) leads to

$$T(g)^* Q'_\omega T(g) = Q'_\omega \qquad \text{for} \quad g \in G_\omega, \qquad \text{where} \quad Q'_\omega = B_\omega$$

whence $Q'_\omega T_\sigma = -T^*_\sigma Q'_\omega$ for $\sigma \in \mathfrak{g}_\omega$.

Assumption 1 (Local existence of solutions). For each $u_0 \in$ exists $t_0 > 0$ depending only on μ, where $\|u_0\| \leqslant \mu$, and there solution of

$$\frac{du}{dt} = JE'(u(t)), \qquad u \in \mathscr{C}(\mathscr{I}; X)$$

in the interval $\mathscr{I} = [0, t_0)$ such that $u(0) = u_0$ and $E(u(t))$ $Q_\sigma(u(t)) = Q_\sigma(u_0)$ for $t \in \mathscr{I}$ and $\sigma \in \mathfrak{g}$.

By a "solution" of (2.10) we mean a weak solution as defined in [1]. As shown there, $T(g)\,u(t)$ is a solution if $u(t)$ is one and $g \in G$.

This assumption is the same as in [1]. Note that the functionals E and Q_σ are formally conserved. Indeed, a smooth solution satisfies

$$\frac{dQ_\sigma(u)}{dt} = \left\langle Q'_\sigma(u), \frac{du}{dt}\right\rangle = -\langle JQ'_\sigma(u), E'(u)\rangle = -\langle T_\sigma u, E'(u)\rangle = 0$$

by (2.3), and similarly for $E(u)$.

DEFINITION. By a *bound state* or *solitary wave* we mean a solution of the evolution equation (2.10) of the form

$$u(t) = T(e^{t\omega})\varphi, \qquad \text{where} \quad \omega \in \mathfrak{g} \text{ and } \varphi \in X. \tag{2.11}$$

LEMMA 2.1. *$T(e^{t\omega})\varphi$ is a bound state if $\varphi \in D(T_\omega)$ and*

$$E'(\varphi) = Q'_\omega(\varphi). \tag{2.12}$$

Proof. $\operatorname{Exp}(s\omega)$ commutes with $\exp(t\omega)$. By differentiation, T_ω commutes with $T(\exp(t\omega))$. Using the notation (2.11) and $g = \exp(t\omega)$, we deduce

$$\begin{aligned}\frac{du}{dt} - JE'(u) &= T(e^{t\omega})\,T_\omega\varphi - JE'(T(e^{t\omega})\,\varphi)\\ &= T(g)\{T_\omega\varphi - T(g^{-1})\,JE'(T(g)\varphi)\}\\ &= T(g)\{T_\omega\varphi - JE'(\varphi)\}\\ &= T(g)\{T_\omega\varphi - JQ'_\omega(\varphi)\} = 0\end{aligned}$$

by (2.2), (2.4), (2.5), (2.6), and (2.12).

DEFINITION. If K is a subgroup of G, we say the bound state is *K-stable* if for all $\varepsilon > 0$ there exists $\delta > 0$ such that: whenever $u(t)$ is a solution of (2.10) with $\|u(0) - \varphi\| < \delta$, the solution $u(t)$ exists for all $t \geqslant 0$ and

$$\sup_{0 \leqslant t < \infty} \ \inf_{g \in K} \|u(t) - T(g)\varphi\| < \varepsilon. \tag{2.13}$$

Otherwise it is called *K-unstable.*

Assumption 2 (Existence of bound states). There is a non-empty set $\Omega \subset \mathfrak{g}$ and a mapping from Ω into X, denoted by $\omega \mapsto \varphi_\omega$, such that

(a) the mapping is C^1,

(b) $E'(\varphi_\omega) = Q'_\omega(\varphi_\omega)$,

(c) $\varphi_\omega \in D(T_\sigma^2)$ for all $\sigma \in \mathfrak{g}$,

(d) $\Omega \cap \mathfrak{g}_\omega$ is open in $\mathfrak{g}_\omega$ for all $\omega \in \mathfrak{g}$.

As in [1] we define

$$H_\omega = E''(\varphi_\omega) - Q''_\omega(\varphi_\omega) \colon X \to X^*.$$

LEMMA 2.2. *$H_\omega(T_\sigma \varphi_\omega) = Q'_{[\sigma,\omega]}\varphi_\omega$ for all $\sigma \in \mathfrak{g}$ and $\omega \in \Omega$.*

Proof. For all $g \in G$ we have

$$\begin{aligned} E'(T(g)\,\varphi_\omega) - Q'_\omega(T(g)\,\varphi_\omega) &= T^*(g)^{-1} E'(\varphi_\omega) - J^{-1}T_\omega T(g)\,\varphi_\omega \\ &= J^{-1}T(g)\,J \cdot J^{-1}T_\omega \varphi_\omega - J^{-1}T_\omega T(g)\,\varphi_\omega \\ &= J^{-1}[T(g), T_\omega]\,\varphi_\omega. \end{aligned}$$

Setting $g = \exp(s\sigma)$ and taking the derivative at $s = 0$, we get

$$\{E''(\varphi_\omega) - Q''_\omega\}\; T_\sigma \varphi_\omega = J^{-1} T_{[\sigma,\omega]}\varphi_\omega.$$

DEFINITION. Let $Z = \{T_\sigma \varphi_\omega \mid \sigma \in \mathfrak{g}_\omega\}$. By Lemma 2.2, Z is contained in the kernel of H_ω.

Assumption 3 (Spectral decomposition of H_ω). The space X is decomposed as a direct sum,

$$X = N + Z + P, \tag{2.14}$$

where Z is defined above, N is a finite-dimensional subspace such that

$$\langle H_\omega u, u\rangle < 0 \qquad \text{for} \quad 0 \neq u \in N, \tag{2.15}$$

and P is a closed subspace such that

$$\langle H_\omega u, u\rangle \geqslant \delta\, \|u\|^2 \qquad \text{for} \quad u \in P \tag{2.16}$$

for some constant $\delta > 0$.

A few words of explanation are in order. The direct sum is in the sense of vector spaces: each u can be written uniquely as $n + z + p$ with $n \in N$, $z \in Z$, and $p \in P$. It follows that Z equals the kernel of H_ω and that N and P are maximal subspaces with the properties (2.15) and (2.16).

It may be helpful to introduce some general terminology. Let X be a real vector space and $h\colon X \times X \to \mathbb{R}$ be a symmetric bilinear form. A subspace N (or P) is called *negative* (or *positive*) if $h(u, u) < 0$ (or >0) for all $0 \neq u \in N$ (or P). The *kernel* of h is $Z = \{u \in X \mid h(u, v) = 0$ for all $v \in X\}$. Then all maximal negative subspaces have the same dimension ($\leqslant \infty$), which we call the *negative index* $n(h)$. Similarly, there is a positive index $p(h)$. In case $h(u, u) = \langle Hu, u\rangle$, where X is a Hilbert space, we write $n(H)$ and $p(H)$ for these indices. We also write $Z(H)$ for the kernel of H (=kernel of h) and

$z(H) = \dim Z(H)$. In our case (Assumption 3), where $H = H_\omega$, we have $n(H_\omega) = \dim N < \infty$ and $p(H_\omega) = \dim P = \infty$ (if $\dim X = \infty$). The decomposition (2.14) is orthogonal if N and P are chosen to be spectral subspaces for H_ω. Thus $n(H_\omega)$ is the number of negative eigenvalues of H_ω.

We define $d: \Omega \mapsto \mathbb{R}$ by

$$d(\omega) = E(\varphi_\omega) - Q_\omega(\varphi_\omega).$$

For each $\omega \in \Omega$, we define $d''(\omega)$ to be the Hessian of this function restricted to $\Omega \cap \mathfrak{g}_\omega$. It is a symmetric bilinear form on $\mathfrak{g}_\omega$, so that we can speak of $n(d'')$, $z(d'')$, and $p(d'')$, which are the number of its eigenvalues which are negative, zero, and positive, respectively.

We shall write φ for φ_ω, H for H_ω, where there is no possibility of confusion.

3. The Reduced Hamiltonian

For fixed $\omega \in \Omega$, we write $H = H_\omega$ and define the subspace

$$X_1 = \{x \in X \mid \langle Q'_\sigma(\varphi_\omega), u \rangle = 0, \forall \sigma \in \mathfrak{g}_\omega\}$$

and let Π_1 be the orthogonal projection of X onto X_1. We define the *reduced Hamiltonian*

$$H_1 = \Pi_1^* H, \qquad \text{restricted to } X_1.$$

Let $z_0 = \dim\{\sigma \in \mathfrak{g}_\omega \mid T_\sigma \varphi_\omega = 0\}$. By (3.6) and (3.9) below, $z_0 \leqslant z(d'')$.

THEOREM 3.1. *The reduced Hamiltonian has the negative index*

$$n(H_1) = n(H) - p(d'') - (z(d'') - z_0), \tag{3.1}$$

and the null index

$$z(H_1) = z(H) + (z(d'') - z_0). \tag{3.2}$$

COROLLARY 3.2. $p(d'') \leqslant n(H)$.

Proof of Theorem 3.1. Differentiating the equation $E'(\varphi_\omega) = Q'_\omega(\varphi_\omega)$ with respect to ω in the direction $\sigma \in \mathfrak{g}_\omega$, we get

$$H_\omega(\partial_\sigma \varphi_\omega) = Q'_\sigma(\varphi_\omega), \tag{3.3}$$

where ∂_σ denotes differentiation in $\mathfrak{g}_\omega$ in the σ direction. That is, $\partial_\sigma \varphi_\omega = (d/d\varepsilon)|_{\varepsilon=0} \varphi_{\omega+\varepsilon\sigma}$. Differentiating $d(\omega) = E(\varphi_\omega) - Q_\omega(\varphi_\omega)$, we obtain

$$\partial_\sigma d(\omega) = -Q_\sigma(\varphi_\omega). \tag{3.4}$$

Differentiating once more,

$$\begin{aligned}\partial_\tau \partial_\sigma d(\omega) &= -\langle Q'_\sigma(\varphi_\omega), \partial_\tau \varphi_\omega\rangle \\ &= -\langle H_\omega(\partial_\sigma \varphi_\omega), \partial_\tau \varphi_\omega\rangle \end{aligned} \tag{3.5}$$

for any pair of directions $\tau, \sigma \in \mathfrak{g}_\omega$. From now on we write $\varphi = \varphi_\omega$ and $H = H_\omega$. Let

$Y = \{\partial_\sigma \varphi \mid \sigma \in \mathfrak{g}_\omega\}$,

$Y_1 = Y \cap X_1 = \{\partial_\tau \varphi \mid \partial_\tau \partial_\sigma d = \langle Q'_\sigma(\varphi_\omega), \partial_\tau \varphi\rangle = 0 \text{ for all } \sigma \in \mathfrak{g}_\omega\}$,

$Y_2 = Y \cap X_1^\perp$ (orthogonality with respect to (,)).

Thus $\partial_\tau \varphi \in Y_1$ if and only if $\partial_\tau \varphi \in X_1$ if and only if $\partial_\tau \partial_\sigma d = 0$ for all $\sigma \in \mathfrak{g}_\omega$. Hence

$$z(d'') = \dim Y_1. \tag{3.6}$$

We denote the restriction of any operator A to a subspace M by $A \mid M$. By (3.5) and the definition of Y_1, we have

$$p(d'') = n(H \mid Y) = N(H \mid Y_2)$$

and

$$0 = z(H \mid Y_2).$$

Let $u \in X_1 + Y = X_1 \oplus Y_2$. Write $u = u_1 + y$, where $u_1 \in X_1$ and $y \in Y_2$. Then $\langle Hy, u_1\rangle = 0$ by (3.3) and the definitions of X_1 and Y. So we have the decomposition

$$\langle Hu, u\rangle = \langle Hu_1, u_1\rangle + \langle Hy, y\rangle$$

on the direct sum $X_1 \oplus Y_2$. It follows that the non-positive dimensions of H add up, in the sense that

$$\begin{aligned} n(H \mid X_1 + Y) + z(H \mid X_1 + Y) &= n(H_1) + z(H_1) + n(H \mid Y_2) + z(H \mid Y_2) \\ &= n(H_1) + z(H_1) + p(d''). \end{aligned} \tag{3.7}$$

Consider the kernel $Z(H_1)$ of H_1. If $u \in Z(H_1)$, then $0 = \langle H_1 u, u\rangle = \langle Hu, \Pi_1 v\rangle$ for all $v \in X$, whence

$$Hu = Q'_\tau(\varphi) = H(\partial_\tau \varphi) \qquad \text{for some} \quad \tau \in \mathfrak{g}_\omega.$$

Hence $u - \partial_\tau \varphi \in Z$. Furthermore, $Z \subset X_1$ because if $v \in Z$ then

$$0 = \langle Hv, \partial_\sigma \varphi\rangle = \langle H(\partial_\sigma \varphi), v\rangle = \langle Q'_\sigma(\varphi), v\rangle$$

for all $\sigma \in \mathfrak{g}_\omega$. Therefore the decomposition

$$u = (u - \partial_\tau \varphi) + (\partial_\tau \varphi) \qquad \text{with} \quad \partial_\tau \varphi \in X_1$$

shows that

$$Z(H_1) = Z + Y_1. \tag{3.8}$$

Moreover, if $v \in Z \cap Y_1$, then $v = \partial_\tau \varphi$ for some τ, so that $0 = Hv = H(\partial_\tau \varphi) = Q'_\tau(\varphi) = J^{-1} T_\tau \varphi$ and $T_\tau \varphi = 0$. This means that

$$z_0 = \dim(Z \cap Y_1). \tag{3.9}$$

By (3.6), (3.8), and (3.9), we have proved

$$z(H_1) = z(H) + z(d'') - z_0. \tag{3.2}$$

Recall that $\langle Hy, y \rangle = 0$ for all $y \in Y_1$, by (3.5) and the definition of Y_1. Let W be any subspace of X such that

$$\langle Hu, u \rangle \leqslant 0 \qquad \text{for all} \quad u \in W + Y_1.$$

We show that $W \subset X_1 + Y$.

In order to do so, let $w \in W$ and $y \in Y_1$. Then

$$\begin{aligned} 0 &\geqslant \langle H(w + y), w + y \rangle \\ &= \langle Hw, w \rangle + 2\langle Hy, w \rangle, \end{aligned}$$

since $\langle Hy, y \rangle = 0$. Since w may be multiplied by a small constant of either sign, it follows that $\langle Hy, w \rangle = 0$. This means $H(Y_1) \subset W^\perp$. (In this argument the $\perp$ denotes orthogonality with respect to $\langle\ ,\ \rangle$.) By (3.3), $H(Y_1) \subset X_1^\perp$. By (3.5), $H(Y_1) \subset Y^\perp$. Thus

$$H\colon Y_1 \mapsto (X_1 \cap Y \cap W)^\perp.$$

The kernel of this mapping is $Z \cap Y_1 = Z \cap Y$, which we showed above has dimension z_0. Hence

$$\dim Y_1 \leqslant z_0 + \operatorname{codim}(X_1 \cap Y \cap W). \tag{3.10}$$

On the other hand, H maps all of Y onto $X_1^\perp$ with the same kernel $Z \cap Y$ by (3.5). (Remember that $X_1^\perp = \{Q'_\sigma(\varphi)\}$.) Hence

$$\dim Y = z_0 + \operatorname{codim} X_1. \tag{3.11}$$

But, of course (for any subspaces),

$$\dim Y - \operatorname{codim} X_1 = \dim(X_1 \cap Y) - \operatorname{codim}(X_1 + Y). \tag{3.12}$$

By (3.10), (3.11), and (3.12) we deduce that

$$\operatorname{codim}(X_1 + Y) \leqslant \operatorname{codim}(X_1 + Y + W).$$

Therefore, $W \subset X_1 + Y$.

This means that a maximal non-positive subspace of $H \mid (X_1 + Y)$ cannot be enlarged to a non-positive subspace of H. In other words,

$$n(H) + z(H) = n(H \mid (X_1 + Y)) + z(H \mid (X_1 + Y)). \tag{3.13}$$

By (3.13), (3.7), and (3.2) we conclude that

$$\begin{aligned} n(H) + z(H) &= n(H_1) + z(H_1) + p(d'') \\ &= n(H_1) + z(H) + z(d'') - z_0 + p(d''). \end{aligned}$$

This proves (3.1).

4. Stability

We showed in Section 3 that $p(d'') \leqslant n(H)$. If $p(d'') = n(H)$, then (3.1) implies that $z(d'') = z_0 = 0$ and the Stability Theorem of Section 1 asserts that φ is G_ω-stable.

The following more general theorem permits the function d to be degenerate. Let

$$\mathscr{K}_\omega = \{\tau \in \mathfrak{g}_\omega \mid \partial_\tau \varphi_\omega \in X_1\}$$

and let $d_{\mathscr{K}} : \mathscr{K}_\omega \mapsto \mathbb{R}$ be the function $\tau \mapsto d(\omega + \tau)$. In the case that d is non-degenerate, we have $z(d'') = 0$ and $\mathscr{K}_\omega = \{0\}$.

THEOREM 4.1. *The bound state $T(e^{i\omega})\varphi_\omega$ is G_ω-stable if both of the following conditions hold*:

(a) $n(H) = p(d'') + z(d'') - z_0$.

(b) *$d_{\mathscr{K}}$ is strictly convex in a neighborhood of $\tau = 0$.*

If $z(d'') = 0$, then $z_0 = 0$, $\mathscr{K}_\omega = \{0\}$, and (b) is trivial, so that Theorem 4.1 generalizes the Stability Theorem.

Note that the Hessian of $d_{\mathscr{K}}$ always vanishes at ω because

$$\partial_\tau \partial_\sigma d(\omega) = -\langle Q'_\sigma(\varphi_\omega), \partial_\tau \varphi_\omega \rangle = 0$$

for $\tau, \sigma \in \mathscr{K}_\omega$. Therefore (b) is simply an assumption about the higher order behavior of $d_{\mathscr{K}}$ at the point in question.

LEMMA 4.2. *Assume* (a), (b), *and* $z_0 = 0$. *There is a smooth map*

$$X \supset V' \ni u \mapsto (g, \Omega, a) \in G_\omega \times (\omega + \mathscr{K}_\omega) \times \mathfrak{g}_\omega,$$

where V' *is a neighborhood of* φ, *with the following property*: *Let*

$$y = T(g)u - \varphi_\Omega - I^{-1}Q'_a(\varphi_\Omega). \tag{4.1}$$

Then

$$(y, T_\sigma \varphi_\Omega) = (y, \partial_\Sigma \varphi) = \langle Q'_\sigma(\varphi_\Omega), y\rangle = 0 \tag{4.2}$$

for all $\sigma \in \mathfrak{g}_\omega$ *and* $\Sigma \in \mathscr{K}_\omega$.

Proof. With y defined by (4.1), we use the Implicit Function Theorem to solve (4.2) for the unknown parameters $(\hat{g}, \Omega, a)$. The Jacobian matrix of the mapping defined by (4.2), evaluated at $g = e$, $a = 0$, $\Omega = \omega$, and $u = \varphi = \varphi_\omega$, is

$$\begin{pmatrix} (T_\tau \varphi, T_\sigma \varphi) & -(\partial_\Sigma \varphi, T_\sigma \varphi) & -\langle Q'_\tau(\varphi), T_\sigma \varphi\rangle = 0 \\ (T_\Sigma \varphi, \partial_\Lambda \varphi) & -(\partial_\Sigma \varphi, \partial_\Lambda \varphi) & -\langle Q'_\tau(\varphi), \partial_\Lambda \varphi\rangle = 0 \\ \langle T_\tau \varphi, Q'_\sigma(\varphi)\rangle = 0 & -\langle \partial_\Sigma \varphi, Q'_\sigma(\varphi)\rangle = 0 & -(I^{-1}Q'_\tau(\varphi), I^{-1}Q'_\sigma(\varphi)) \end{pmatrix}$$

(written in blocks), where τ and σ run through a basis of $\mathfrak{g}_\omega$ and Σ and Λ run through a basis of $\mathscr{K}_\omega$. The Jacobian is nonsingular because the vectors involved are linearly independent. To be more precise, we now provide the details. The off-diagonal corners vanish because of (2.7). The other two entries indicated vanish because Λ and Σ belong to $\mathscr{K}_\omega$. The last diagonal entry is nonsingular because otherwise $I^{-1}Q'_\tau(\varphi) = 0$ for some $0 \neq \tau \in \mathfrak{g}_\omega$. In that case $T_\tau \varphi = 0$, which would contradict the assumption that $z_0 = 0$. Thus the whole Jacobian is nonsingular if and only if $\{T_\sigma \varphi, \partial_\Sigma \varphi\}$ is linearly independent, where σ runs over a basis of $\mathfrak{g}_\omega$ and Σ over a basis of $\mathscr{K}_\omega$. If this set were linearly dependent, then $T_\sigma \varphi + \partial_\Sigma \varphi = 0$ for some $(\sigma, \Sigma) \neq (0, 0)$. Then

$$0 = HT_\sigma \varphi + H\,\partial_\Sigma \varphi = J^{-1}T_\Sigma \varphi,$$

so that $\Sigma = 0$ and $\sigma = 0$ and we would have a contradiction.

LEMMA 4.3. *Either* (i) $T(g)\varphi = \varphi$ *for some* $g \in G_\omega$, $g \neq e$, *or* (ii) $T(g_n)\varphi \to \varphi$, $g_n \in G_\omega$ *implies* $g_n \to e$ *in* G_ω.

Proof. The set of critical points of $L = E - Q_\omega$ in a neighborhood of φ is isomorphic to the null space Z of $H = L''(\varphi)$. But $T(g)\varphi$ is a critical point of L for all $g \in G_\omega$. Therefore there is a neighborhood N of φ in X and a neighborhood M of e in G_ω such that

$$\{u \in N \mid L'(u) = 0\} = \{T(g)\varphi \mid g \in M\}.$$

If (ii) is false, there is a sequence g_n and a neighborhood M of e with $g_n \notin N$ and $T(g_n)\varphi \in N$. Fixing n, we have shown that there exist $h_n \in M$ with $T(g_n)\varphi = T(h_n)\varphi$. So $T(g_n^{-1}h_n)\varphi = \varphi$, which means that (i) is valid.

LEMMA 4.4. *For $\varepsilon > 0$ define the "tube"*

$$U_\varepsilon = \{u \in X \mid \inf_{g \in G_\omega} \|u - T(g)\varphi_\omega\| < \varepsilon\}.$$

If ε is small enough, the map of Lemma 4.2 *extends to a smooth map*

$$U_\varepsilon \ni u \mapsto (g(u), \Omega(u), a(u)) \tag{4.3}$$

such that (4.1)–(4.2) *is valid and*

$$g(T(h)u) = g(u) \cdot h^{-1} \tag{4.4}$$

for $h \in G_\omega$ and $u \in U_\varepsilon$.

Proof. We first show that there exists a neighborhood V of φ such that (4.4) is valid if $u \in V$ and $T(h)u \in V$. Indeed, let W' be the image of V' under the map $u \mapsto g(u)$. By Lemma 4.3, V' can be chosen so small that

$$T(h)\varphi \in V' \text{ and } h \in G_\omega \text{ imply } h \in W'. \tag{4.5}$$

By shrinking the neighborhood further we may assume that $W'W'^{-1} \subset W'$. If V' contains the ball of radius η and center φ, let V be the ball of radius $\eta/2$ and let $W = g(V)$.

Let $u \in V$ and $T(h)u \in V$. Then $T(g_n)\varphi$,

$$\|T(h)\varphi - \varphi\| \leqslant \|T(h)[\varphi - u]\| + \|T(h)u - \varphi\| < \eta,$$

so that $T(h)\varphi \in V'$. By (4.5), $h \in W'$. Also $g(u) \in W \subset W'$ so that $g(u)h^{-1} \in W'$. Since

$$T(g(u))u = T(g(u)\, h^{-1})\, T(h)u,$$

the uniqueness of the local map in Lemma 4.2 implies that $g(u)h^{-1}$ must be the image of $T(h)u$ under the map $g(\cdot)$. That is, (4.4) holds if $u \in V$ and $T(h)u \in V$. It also follows that $\Omega(T(h)u) = \Omega(u)$ and $a(T(h)u) = a(u)$.

Now let $u \in U_\varepsilon$. We choose ε so small that, for all such u, $T(h)u \in V$ for some $h \in G_\omega$. Define

$$g(u) = g(T(h)u)h, \qquad \Omega(u) = \Omega(T(h)u), \qquad a(u) = a(T(h)u).$$

This definition is independent of the choice of h for the following reason: Let $T(k)u$ also belong to V for some $k \in G_\omega$. Applying the first part of this proof to the group element hk^{-1} and the point $T(h)u$, we have

$$g(T(kh^{-1})\,T(h)u) = g(T(h)u)\,hk^{-1},$$

which means that

$$g(T(k)u)k = g(T(h)u)h.$$

Furthermore, (4.1)–(4.2) is valid in U_ε because it is valid in V.

LEMMA 4.5. *Assume* (a), (b), *and* $z_0 = 0$. *Then* $E(u) \geqslant E(\varphi_\omega)$ *for all*

$$u \in U_\varepsilon \qquad \textit{with} \quad Q_\sigma(u) = Q_\sigma(\varphi_\omega), \forall \sigma \in \mathfrak{g}_\omega. \tag{4.6}$$

If $E(u) = E(\varphi_\omega)$ *then* $u = T(g)\varphi_\omega$ *for some* $g \in G_\omega$.

Proof. Let u satisfy (4.6). Denote $g = g(u)$, $\Omega = \Omega(u)$, and $a = a(u)$ as in Lemma 4.4. Then, for all $\sigma \in \mathfrak{g}_\omega$,

$$\begin{aligned} Q_\sigma(\varphi_\omega) &= Q_\sigma(u) = Q_\sigma(T(g)u) \\ &= Q_\sigma(\varphi_\Omega) + \langle Q'(\varphi_\omega), T(g)\,u - \varphi_\Omega \rangle + O(\|T(g)u - \varphi_\Omega\|^2). \end{aligned}$$

Denoting $z = T(g)u - \varphi_\Omega = y + I^{-1}Q'_a(\varphi_\Omega)$, where y is given by (4.1), we have

$$Q_\sigma(\varphi_\omega) = Q_\sigma(\varphi_\Omega) + (I^{-1}Q'_\sigma(\varphi_\Omega), I^{-1}Q'_a(\varphi_\Omega)) + O(\|z\|^2)$$

for all $\sigma \in \mathfrak{g}_\omega$. Choosing $\sigma = a/|a|$ and noting that $Q'_a(\varphi_\Omega) \neq 0$ unless $a = 0$,

$$|a| \leqslant c\,|Q_\sigma(\varphi_\omega) - Q_\sigma(\varphi_\Omega)| + c\,\|z\|^2. \tag{4.8}$$

Recall that $Q_\sigma(\varphi_\omega) - Q_\sigma(\varphi_\Omega) = -\partial_\sigma\,d(\omega) + \partial_\sigma\,d(\Omega)$. On the other hand, a Taylor expansion around φ_Ω gives

$$\begin{aligned} E(u) - Q_\Omega(u) &= E(T(g)u) - Q_\Omega(T(g)u) \\ &= d(\Omega) + \tfrac{1}{2}\langle H_\Omega z, z\rangle + o(\|z\|^2). \end{aligned}$$

We substitute $z = y + I^{-1}Q'_a(\varphi_\Omega)$ to obtain

$$d(\Omega) + \tfrac{1}{2}\langle H_\Omega\, y, y\rangle + O(|a|\;\|y\| + |a|^2) + o(\|z\|^2).$$

But y satisfies (4.2) by Lemma 4.4, and

$$\langle H_\Omega\, y, y\rangle \geqslant \delta\,\|y\|^2.$$

To prove the last inequality, we note that, by Theorem 3.1, $n(H_1)=0$ and, by (3.8), $Z(H_1)=Z+Y_1$. Therefore $\langle H_1 y, y\rangle > 0$ for all y orthogonal to $Z+Y_1$. Therefore $\langle Hy, y\rangle > 0$ for all y orthogonal to $T_\sigma \varphi_\omega$, $\partial_\Sigma \varphi_\omega$, and $I^{-1}Q'_\sigma(\varphi_\omega)$ for all $\sigma \in \mathfrak{g}_\omega$ and all $\Sigma \in \omega + \mathscr{K}_\omega$. The assertion follows by continuity.

Putting these estimates together, we have

$$\begin{aligned} E(u) - Q_\Omega(u) &\geqslant d(\Omega) + \tfrac{1}{2}\delta \, \|y\|^2 + O(|a| \, \|y\| + |a|^2) + O(\|z\|^2) \\ &\geqslant d(\Omega) + \tfrac{1}{3}\delta \, \|z\|^2 + O(|a| \, \|z\| + |a|^2) + o(\|z\|^2) \\ &\geqslant d(\Omega) + \tfrac{1}{4}\delta \, \|z\|^2 - c\,|a|^2 \\ &\geqslant d(\Omega) + \tfrac{1}{5}\delta \, \|z\|^2 - c\,|\partial_\sigma d(\Omega) - \partial_\sigma d(\omega)|^2 \end{aligned}$$

by (4.8). We combine this estimate with the identity

$$E(\varphi_\omega) - Q_\Omega(\varphi_\omega) = d(\omega) + Q_{\omega - \Omega}(\varphi_\omega) = d(\omega) + \partial_{\Omega - \omega} d(\omega)$$

to obtain

$$\begin{aligned} E(u) - E(\varphi_\omega) \geqslant \tfrac{1}{5}\delta \, \|z\|^2 + d(\Omega) - d(\omega) - \partial_{\Omega - \omega} d(\omega) \\ - c\,|\partial_\sigma d(\Omega) - \partial_\sigma d(\omega)|^2, \end{aligned}$$

where $\Omega = \Omega(u) \in \omega + \mathscr{K}_\omega$ and $\sigma = a/|a|$.

Let $x = \Omega - \omega$ and $f(x) = d(\omega + x) - d(\omega) - x \cdot d'(\omega)$. Then f is a strictly convex function of $x \in \mathscr{K}_\omega$ near $x = 0$, and $f(0) = f'(0) = f''(0) = 0$.

LEMMA 4.6. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be strictly convex and C^2 in a neighborhood of the origin. Let $f(0) = f'(0) = f''(0) = 0$, where $f' = \nabla f$ is the gradient and $f'' = H$ is the Hessian. Then $f \geqslant 0$ in a neighborhood of the origin and $|f'|^2 = o(f)$ as $x \to 0$.*

Proof. The positivity follows directly from

$$f(x) = \int_0^1 f''(tx)x \cdot x(1-t)\, dt \geqslant 0.$$

If $f(x) = 0$ for some x in a neighborhood, then $f''(tx)x \cdot x = 0$ for $t \in [0, 1]$, hence $f(tx) = 0$ for $t \in [0, 1]$, contradicting the strict convexity. So $f(x) > 0$ for $x \neq 0$ (in a neighborhood). Hence,

$$f'(x) \cdot x = \int_0^1 f''(tx)x \cdot x \, dt > 0 \qquad \text{for} \quad x \neq 0.$$

Let $\varepsilon > 0$ and choose a neighborhood $N = \{x \mid f(x) < \delta\}$ of the origin in which $|f''(x)| < \varepsilon/2$. Let $x_0 \in N$ and define the curve

$$\frac{dx}{dt} = -\frac{f'(x)}{|f'(x)|^2}, \qquad x(0) = x_0.$$

Along the curve, $df/dt = -1$. Hence $x(t) \in N$ for $t > 0$ and $f(x(T)) = f(x_0) - T = 0$, where $T = f(x_0)$. Hence $x(T) = 0$. Now let $K = \frac{1}{2}|f'|^2 - \varepsilon f$. Then

$$\frac{dK}{dt} = K' \cdot \frac{dx}{dt} = (f'' - \varepsilon) f' \cdot \frac{-f'}{|f'|^2}$$
$$= (\varepsilon - f'') \frac{f'}{|f'|} \cdot \frac{f'}{|f'|} \geqslant \frac{\varepsilon}{2} > 0.$$

Hence

$$K(x_0) = K(x_0) - K(0) = K(x(0)) - K(x(T)) < 0;$$

whence $|f'(x_0)|^2 \leqslant 2\varepsilon f(x_0)$.

The proof of Lemma 4.5 is completed as follows: We have

$$E(u) - E(\varphi_\omega) \geqslant \tfrac{1}{5}\delta \, \|z\|^2 + f(x) - c\, |f'(x)|^2.$$

By Lemma 4.6, $E(u) \geqslant E(\varphi_\omega)$. If $E(u) = E(\varphi_\omega)$, then $z = 0$ and $f(x) = 0$. Hence $x = 0$ and $\Omega = \omega$. Recalling the definition of z, this means $T(g)u = \varphi_\omega$, where $g = g(u)$.

Proof of Theorem 4.1 *if* $z_0 = 0$. If φ is unstable, then there exists a sequence of initial data $u_n(0)$ and $\eta > 0$ such that $\|u_n(0) - \varphi\| \to 0$, but $\sup_t \inf_g \|u_n(t) - T(g)\varphi\| \geqslant \eta$, where $u_n(t)$ is a solution with initial data u_n. By the continuity in t we can pick the first time t_n such that

$$\inf_{g \in G} \|u_n(t_n) - T(g)\varphi\| = \eta,$$

the solution existing at least in the time interval $[0, t_n]$. Then $E(u_n(t_n)) = E(u_n(0)) \to E(\varphi)$ and

$$Q_\sigma(u_n(t_n)) = Q_\sigma(u_n(0)) \to Q_\sigma(\varphi) \qquad \text{for all} \quad \sigma \in \mathfrak{g}_\omega.$$

Choose a sequence $\{v_n\}$ so that

$$\|v_n - u_n(t_n)\| \to 0 \qquad \text{and} \qquad Q_\sigma(v_n) = Q_\sigma(\varphi), \qquad \forall \sigma \in \mathfrak{g}_\omega.$$

By the continuity of E, $E(v_n) \to E(\varphi)$. Choosing η sufficiently s
may apply Lemma 4.5 to obtain

$$E(v_n) - E(\varphi) \geqslant \frac{\delta}{5} \|z_n\| + \frac{1}{2} f(x_n),$$

where $z_n = T(g_n) v_n - \varphi_{\Omega_n}$, $x_n = \Omega_n - \omega$, $g = g(v_n)$, and $\Omega_n = \Omega(v$
$\|z_n\| \to 0$ and $f(x_n) \to 0$. Thus $x_n \to 0$, which means $\Omega_n \to \omega$.
$\|T(g_n) v_n - \varphi\| \to 0$ so that

$$\inf_g \|u_n(t_n) - T(g)\varphi\| \to 0.$$

This contradicts the assumption.

Proof in Case $z_0 \neq 0$. In this case

$$\mathscr{K}_\omega^0 = \{\sigma \in \mathfrak{g}_\omega \mid T_\sigma \varphi = 0\} = \{\sigma \in \mathfrak{g}_\omega \mid \partial_\sigma \varphi \in Z\}$$

is a nontrivial subspace of $\mathscr{K}_\omega$. Choose a basis $\{\sigma_1, ..., \sigma_m\}$ of g
$\{\sigma_1, ..., \sigma_p\}$ is a basis of $\mathscr{K}_\omega^0$ and $\{\sigma_1, ..., \sigma_p, ..., \sigma_q\}$ is a basis of $\mathscr{K}$
$0 \leqslant p \leqslant q \leqslant m$. We claim that Lemma 4.2 is valid except that (Ω
linear combination of $\sigma_{p+1}, ..., \sigma_q$; a is a linear combination of σ_{p}
and g is the exponential of a linear combination of $\sigma_{p+1}, ...,$
orthogonality conditions (4.2) are unchanged. For $0 \leqslant j \leqslant p$, the c
(4.2) are redundant because $T_{\sigma_j} \varphi = 0$, $Q'_{\sigma_j}(\varphi) = J^{-1} T_{\sigma_j} \varphi = 0$, and
Therefore the conditions (4.2) comprise $q + 2m - 3p$ independent co

It follows that the $(q + 2m - 3p) \times (q + 2m - 3p)$ Jacobian deter
non-zero, for the same reasons as before. In this Jacobian the pa
σ and τ run over the span of $\{\sigma_{p+1}, ..., \sigma_m\}$ and the parameters
run over the span of $\{\sigma_{p+1}, ..., \sigma_q\}$. This proves the analog of Le
There is no change in the rest of the proof of Theorem 4.1.

Remark. If, in Theorem 4.1, $d_{\mathscr{K}}$ is convex but not strictly con
the bound state is stable in the weaker sense that

$$\inf_{t>0} \inf_{\Omega} \inf_{g \in G_\omega} \|u(t) - T(g)\varphi_\Omega\| < \varepsilon$$

for $\|u(0) - \varphi\| < \delta$, where Ω runs over the set

$$\{\Omega - \omega \in \mathscr{K}_\omega \mid Q_{\Omega - \omega}(\varphi_\Omega) = Q_{\Omega - \omega}(\varphi_\omega)\}.$$

To prove this, we note that in Lemma 4.5 we have the equalit
$E(\varphi_\omega)$ only if $Q_{\Omega-\omega}(\varphi_\Omega) = Q_{\Omega-\omega}(\varphi_\omega)$. At the end of the
Theorem 4.1, we have $\|z_n\| \to 0$, so that $\|v_n - T(g_n^{-1}) \varphi_{\Omega_n}\| \to 0$ an

$$\inf_\Omega \inf_g \|u_n(t_n) - T(g)\varphi_\Omega\| \to 0.$$

5. The Generator JH

The instability is based on the spectrum of the linearized operator JH.

THEOREM 5.1. *Let $d''(\omega)$ be non-singular, $n(H)-p(d'')$ be odd, and X be be separable. Then JH has at least one pair of real non-zero eigenvalues $\pm\lambda$.*

From Section 3 we recall that X_1 denotes the vectors orthogonal to $Q'_\sigma(\varphi)$ for all $\sigma\in\mathfrak{g}_\omega$, that Π_1 is the orthogonal projection of X onto X_1, that $H_1=\Pi_1^* H\Pi_1$, that $Z=\{T_\sigma\varphi\}$ is the kernel of H, and that $Y=\{\partial_\sigma\varphi\}$. We also define

$$X_2=\{u\in X \mid \langle Q'_\sigma(\varphi), u\rangle=0=(T_\sigma\varphi, u), \forall\sigma\in\mathfrak{g}_\omega\}. \tag{5.1}$$

LEMMA 5.2. *We have the (non-orthogonal) direct sums*

$$X=X_1+Y=X_2+Z+Y.$$

Proof. Although this lemma essentially follows from the proof of Theorem 3.1, we repeat the argument. By assumption, $-\partial_\sigma\partial_\tau d(\omega)=\langle Q'_\sigma(\varphi), \partial_\tau\varphi\rangle$ is non-singular. Given $u\in X$, we choose $\tau\in\mathfrak{g}_\omega$ uniquely so that $\langle Q'_\sigma(\varphi), \partial_\tau\varphi\rangle=\langle Q'_\sigma(\varphi), u\rangle$, $\forall\sigma\in\mathfrak{g}_\omega$. Then $u-\partial_\tau\varphi\in X_1$. So we have a unique representation of $u=(u-\partial_\tau\varphi)+(\partial_\tau\varphi)$ as the sum of elements of X_1 and Y.

Second, since $\langle Q'_\sigma(\varphi), T_\tau\varphi\rangle=0$ for all $\sigma, \tau\in\mathfrak{g}_\omega$, we have $Z\subset X_1$. So we simply decompose X_1 into Z and its orthogonal complement X_2 in X_1.

We define

$$\mathscr{P}: X\to X_2 \tag{5.2}$$

to be the (non-orthogonal) projection of X onto X_2 defined by Lemma 5.2.

LEMMA 5.3. (a) *H restricted to X_2+Y is one–one.*

(b) *The image of the restriction contains $J^{-1}(X_2)$.*

(c) *$\mathscr{P}H^{-1}J^{-1}$ maps X_2 into X_2 in a one–one manner.*

Proof. (a) The restriction is one–one because X_2+Y does not meet the kernel Z of H.

(b) If $w\in W_2$, then

$$\langle J^{-1}w, T_\sigma\varphi\rangle=-\langle w, J^{-1}T_\sigma\varphi\rangle=-\langle w, Q'_\sigma(\varphi)\rangle=0$$

so that $J^{-1}w$ is orthogonal to Z. Hence the equation $Hy = J^{-1}w$ has a solution $y \in X$. By Lemma 5.2 we may decompose

$$y = v + \partial_\tau \varphi + T_\sigma \varphi \qquad \text{with} \quad v \in X_2.$$

Then $H(v + \partial_\tau \varphi) = Hy = J^{-1}w$. This proves (b).

(c) It follows from (a) and (b) that $H^{-1}J^{-1}$ maps X_2 into $X_2 + Y$. If the projection $\mathscr{P}$ is applied, we get a map from X_2 into X_2. If $z \in X_2$ belongs to the kernel of $\mathscr{P}H^{-1}J^{-1}$, then

$$H^{-1}J^{-1}z = \partial_\tau \varphi + T_\sigma \varphi$$

for some τ and σ by definition of $\mathscr{P}$. Then $z = JH\,\partial_\tau \varphi = T_\tau \varphi$ by (3.3). Since $z \in X_2$, $z = 0$. So the map is one–one.

LEMMA 5.4. *Restricted to $X_2 \times X_2$, the symmetric bilinear form $\langle Hu, v\rangle$ has no kernel and its negative index is odd.*

Proof. First, let u belong to the kernel, by which we mean that $u \in X_2$ and $\langle Hu, v\rangle = 0$ for all $v \in X_2$. Let $w = v + \partial_\tau \varphi + T_\sigma \varphi$ be any element of X decomposed according to Lemma 5.2. Then

$$\langle Hu, w\rangle = \langle Hu, v\rangle + \langle u, Q'_\tau(\varphi)\rangle + \langle u, HT_\sigma \varphi\rangle = 0$$

because of (3.3) and $u \in X_2$. Thus $Hu = 0$. Since $u \in X_2$, $u = 0$.

Second, by the assumption of Theorem 5.1 and by (3.1), H_1 has an odd-dimensional negative space $N(H_1)$. That is, $N(H_1) \subset X_1$, the quadratic form $\langle Hu, u\rangle$ is negative on $N(H_1)$, and $N(H_1)$ is maximal with this property. For $y \in N(H_1)$, we decompose $y = (y - T_\tau \varphi) + T_\tau \varphi$, where $y - T_\tau \varphi \in X_2$ as in Lemma 5.2. Replacing y by $y - T_\tau \varphi$, we get a negative subspace in X_2 of the same dimension as that of $N(H_1)$.

In preparation for the next lemma we make the following definitions. By assumption J^{-1} is continuous from X into $D(J) \subset X^*$.

DEFINITION.

$$f(v) = \mathscr{P}H^{-1}J^{-1}v - av, \tag{5.3}$$

where $v \in X_2$ and $a = a(v) = (\mathscr{P}H^{-1}J^{-1}v, v)$. Thus $f\colon X_2 \mapsto X_2$.

DEFINITION.

$$\mathscr{C} = \{v \in X_2 \mid (v, v) = 1 \text{ and } \langle Hv, v\rangle = 0\}. \tag{5.4}$$

Thus $\mathscr{C}$ is the intersection of a cone with a sphere and therefore looks topologically like $S^{k-1} \times S^\infty$, the product of two spheres.

LEMMA 5.5. *f is a tangent vector field on $\mathscr{C}$.*

Proof. The surface $\mathscr{C}$ has codimension two with normal vectors Iv and Hv. So we must check that $(f(v), v) = \langle f(v), Iv \rangle = 0$ and $\langle f(v), v \rangle = 0$ for all $v \in \mathscr{C}$. Now

$$(f(v), v) = (\mathscr{P}H^{-1}J^{-1}v, v) - a(v, v) = 0$$

by the choice of a. Next, we decompose

$$H^{-1}J^{-1}v = \mathscr{P}H^{-1}J^{-1}v + \partial_\tau \varphi$$

as in Lemma 5.3. Then

$$\begin{aligned}\langle Hv, f(v) \rangle &= \langle Hv, H^{-1}J^{-1}v \rangle - \langle Hv, \partial_\tau \varphi \rangle - a\langle Hv, v \rangle \\ &= \langle v, J^{-1}v \rangle - \langle v, Q'_\tau(\varphi) \rangle = 0,\end{aligned}$$

since J is skew, $v \in X_2$, and $\langle Hv, v \rangle = 0$.

Proof of Theorem 5.1. Let N denote a maximal negative subspace of the bilinear form $\langle Hu, u \rangle$ in X_2. By Lemma 5.4 its dimension k is odd. Since X is separable, we can find positive subspaces $P^{(n)}$ of finite dimension which, together with N, fill out X_2. That is, there is an increasing sequence of subspaces $P^{(n)}$ of X_2 of odd dimension n on which $\langle Hu, u \rangle$ is positive such that

$$\bigcup X^{(n)} \text{ is dense in } X_2,$$

where $X^{(n)} = N + P^{(n)}$. We may choose $P^{(n)}$ orthogonal to N, and $P^{(n)}$ and N invariant under the operator H. Let $\Pi^{(n)}$ denote the orthogonal projection of X_2 onto $X^{(n)}$. Consider the mapping $f_n = \Pi^{(n)} f \Pi^{(n)}$ restricted to $\mathscr{C}^{(n)} = \mathscr{C} \cap X^{(n)}$.

For each n, $\{v \in X^{(n)} \mid \langle Hv, v \rangle = 0\}$ is the cone given by a quadratic form of k negative and n positive eigenvalues, and $\mathscr{C}^{(n)}$ is the intersection of this cone with the unit sphere $(v, v) = 1$. The mapping

$$\mathscr{C}^{(n)} \ni v = v^- + v^+ \mapsto \left[\frac{v^-}{\|v^-\|}, \frac{v^+}{\|v^+\|}\right] \in S^{k-1} \times S^{n-1}$$

is a homeomorphism of $\mathscr{C}^{(n)}$ with the product of two even-dimensional spheres, where v^- belongs to a negative subspace of H and v^+ to a positive subspace. Thus $\mathscr{C}^{(n)}$ has non-vanishing Euler characteristic and the tangent vector field f_n must vanish at some $y_n \in \mathscr{C}^{(n)}$. Thus $(f(y_n), w) = 0$ for all $w \in X^{(n)}$. That is, there is a real scalar $a_n = a(y_n)$ such that

$$(\mathscr{P}H^{-1}J^{-1}y_n, w) = a_n(y_n, w) \tag{5.5}$$

for all $w \in X^{(n)}$. Since $\|y_n\| = 1$ and $a_n = (\mathscr{P}H^{-1}J^{-1}y_n, y_n)$ are bounded, we may pick a subsequence such that

$$a_n \to a \text{ and } y_n \rightharpoonup y \text{ weakly in } X_2.$$

The limit satisfies Eq. (5.5) with the subscripts n removed, valid for all $w \in X^{(n)}$. By density it is true for all $w \in X_2$. Hence

$$\mathscr{P}H^{-1}J^{-1}y = ay, \qquad y \in X_2. \tag{5.6}$$

We claim that $y \neq 0$. Indeed we split $y_n = y_n^- + y_n^+$, where $y_n^- \in N$ and $y_n^+ \in P^{(n)}$. Since N is finite-dimensional, $y_n^- \to y^-$ strongly. Let $y^+ = y - y^-$, so that $y_n^+ \rightharpoonup y^+$ weakly and y^+ belongs to a positive subspace P of H in X_2. Since H is strictly positive on P by Assumption 3, we have

$$\begin{aligned}\langle Hy, y\rangle &= \langle Hy^-, y^-\rangle + \langle Hy^+, Hy^+\rangle + 2\langle Hy^-, y^+\rangle \\ &\leqslant \liminf \{\langle Hy_n^-, y_n^-\rangle + \langle Hy_n^+, Hy_n^+\rangle + 2\langle Hy_n^-, y_n^+\rangle\} \\ &= \liminf \langle Hy_n, y_n\rangle = 0,\end{aligned}$$

since $y_n \in \mathscr{C}^{(n)}$. If $\langle Hy, y\rangle \neq 0$ then $y \neq 0$, which we wanted to prove. On the other hand, if $\langle Hy, y\rangle = 0$, the above inequality is an equality, so that $y_n \to y$ strongly. In that case, $\|y\| = 1$. So in either case, $y \neq 0$.

Now we rewrite (5.6) as

$$H^{-1}J^{-1}y = ay + \partial_\tau \varphi$$

for some $\tau \in \mathfrak{g}_\omega$. Thus by (3.3),

$$J^{-1}y = aHy + H(\partial_\tau \varphi) = aHy + J^{-1}T_\tau \varphi.$$

Hence $aHy \in D(J)$ and $y = aJHy + T_\tau \varphi$. Since $y \neq 0$ and y is orthogonal in X to $T_\tau \varphi$, it follows that $aJHy = y - T_\tau \varphi \neq 0$. Therefore, $a \neq 0$ and

$$JH(y - T_\tau \varphi) = JHy = a^{-1}(y - T_\tau \varphi). \tag{5.7}$$

Thus $y - T_\tau \varphi$ is an eigenvector of JH with the eigenvalue $\lambda = a^{-1}$. The proof of Theorem 5.1 is now completed with the following lemma.

LEMMA 5.6. *The spectrum of JH is symmetric with respect to both the real and imaginary axes.*

Proof. Complexify X and X^* and extend H and J so that $\overline{Hu} = H\bar{u}$ and $\overline{Ju} = J\bar{u}$. Then $H^* = H$ and $J^* = -J$ for the extended operators. Now $\overline{JHu} = JH\bar{u}$ implies that the spectrum of JH is invariant under complex conjugation. Also $JH = JHJJ^{-1} = -J(JH^*)J^{-1}$ is similar to the negative of its adjoint. So the spectrum of JH is also invariant under the map $\lambda \to -\bar{\lambda}$.

Remark 5.7. In the proof of Theorem 5.1, the convergence $y_n \to y$ in X is strong. Indeed,

$$\begin{aligned}\langle y, Hy\rangle &= \langle y, \mathscr{P}^* Hy\rangle = \langle \lambda \mathscr{P} H^{-1} J^{-1} y, \mathscr{P}^* Hy\rangle \\ &= \lambda \langle J^{-1} y, y\rangle = 0\end{aligned}$$

by (5.6). Under this condition we proved above that $y_n \to y$ strongly in X.

According to Lemma 5.6, the eigenvalues of JH appear in quadruplets (or pairs on the axes). The next theorem asserts that there can be at most $n(H)$ such quadruplets (or pairs) off the imaginary axis. When dealing with the spectrum of JH, we have complexified the space X and extended the operators J and H to complex-linear operators on the complexification. By $\langle\ ,\ \rangle$ we denote the extended form which is bilinear with respect to the reals.

THEOREM 5.8. *The number of eigenvalues of JH in the half-closed quarter plane* $Q = \{\operatorname{Re} \lambda < 0,\ \operatorname{Im} \lambda \geq 0\}$ *is at most* $n(H)$, *the number of negative eigenvalues of H. The essential spectrum of JH lies on the imaginary axis.*

Proof. We shall only use the basic properties of J and H, including Assumption 3, on the complexified space X. Let

$$JHy_j = \lambda_j y_j \qquad (j = 1, ..., n(H) + 1),$$

where $y_j \neq 0$ and $\lambda_j \in Q$. We show that $y_1, ..., y_{n(H)+1}$ are linearly dependent. We have

$$\langle JHy_j, H\bar{y}_k\rangle = \langle \lambda_j y_j, H\bar{y}_k\rangle$$

and, because $JH\bar{y}_k = \bar{\lambda}_k \bar{y}_k$,

$$\langle JH\bar{y}_k, Hy_j\rangle = \bar{\lambda}_k \langle y_k, Hy_j\rangle.$$

Since J is skew-hermitian,

$$0 = (\lambda_j + \bar{\lambda}_k)\langle Hy_j, \bar{y}_k\rangle.$$

Since $\operatorname{Re}(\lambda_j + \bar{\lambda}_k) < 0$, we have

$$0 = \langle Hy_j, \bar{y}_k\rangle \qquad (j, k = 1, ..., n(H) + 1). \tag{5.8}$$

By Assumption 3 we have the spectral decomposition $X = N + Z + P$ for the operator H, which we may assume is orthogonal. Let $n_1, ..., n_m$ be orthogonal eigenvectors of H in N, where $m = n(H)$. We decompose

$$y_j = a_{j1} n_1 + \cdots + a_{jm} n_m + z_j + p_j$$

with $z_j \in Z$ and $p_j \in P$. We choose scalars $c_1, ..., c_{m+1}$ to be a nontrivial solution of the scalar equations

$$\sum_{j=1}^{m+1} c_j a_{jk} = 0 \qquad (k = 1, ..., m).$$

Then

$$\sum c_j y_j = \sum c_j a_{jk} n_k + \sum c_j (z_j + p_j) = \sum c_j (z_j + p_j).$$

By (5.8) we have

$$0 = \left\langle H\left(\sum c_j y_j\right), \sum \bar{c}_j \bar{y}_j \right\rangle = \left\langle H\left(\sum c_j p_j\right), \sum \bar{c}_j \bar{p}_j \right\rangle.$$

Since H is strictly positive on P, we have $\sum c_j p_j = 0$. Hence $\sum c_j y_j = \sum c_j z_j \in Z$ so that

$$\sum c_j \lambda_j y_j = JH \sum c_j y_j = 0.$$

Since the coefficients are not trivial, the eigenvectors $y_1, ..., y_{m+1}$ are linearly dependent. (The case of generalized eigenvectors y_j is left for the reader.) This proves the first assertion of Theorem 5.8.

On the subspace P we define the new inner product $[u, v] = \langle Hu, v \rangle$ by Assumption 3. This inner product is equivalent to the inner product (u, v) of X. The operator JH satisfies

$$[JHu, v] = \langle HJHu, v \rangle = -\langle Hu, JHu \rangle = -[u, JHv]$$

for $u, v \in P$ with $Hu, Hv \in D(J)$. Hence $PJHP$ has purely imaginary spectrum, where P is the projection onto the subspace P. Therefore the essential spectrum of JH, which differs from $PJHP$ by an operator of finite rank, is also purely imaginary.

Remark. On the finite-dimensional level, the idea of Theorem 5.1 is contained in the following elementary fact. If J and H are nonsingular $n \times n$ real matrices, J is skew, H is symmetric, and H has an odd number of negative eigenvalues, then JH has a real eigenvalue. (Indeed, the eigenvalues come in complex-conjugate pairs. Since J has purely imaginary eigenvalues, $\det J > 0$. By assumption, $\det H < 0$. Thus $\det JH < 0$. So JH must have a real negative eigenvalue.)

6. Instability

Our purpose is to deduce instability from the "linearized instability" of Theorem 5.1. The only non-standard feature is the presence of the group of symmetries in the definition of instability.

THEOREM 6.1. *Consider an equation*

$$\frac{dx}{dt} = Ax + f(x), \qquad x(t) \in X, \tag{6.1}$$

where X is a Hilbert space. Let $f: X \to X$ be a locally Lipschitz mapping such that $\|f(x)\| \leqslant k \|x\|^2$ for $\|x\| \leqslant l$, for some constants $k > 0$ and $l > 0$. Let A be a linear operator which generates a strongly continuous semigroup $\exp(tA)$ *on X. Assume that A has an eigenvalue λ with* $\operatorname{Re} \lambda > 0$ *and that*

$$\|e^{tA}\| \leqslant be^{\mu t} \qquad \text{for some } \mu < 2 \operatorname{Re} \lambda. \tag{6.2}$$

Then the zero solution is (nonlinearly) unstable for (6.1).

A more general conclusion is the following. Let $Ay = \lambda y$, $0 \neq y \in X$, $\operatorname{Re} \lambda > 0$. Let M be any closed subspace not containing y. Then there exist constants ε_0, δ_0, and c_1 such that for any $0 < \varepsilon < \varepsilon_0$ and any $0 < \delta < \delta_0$, there exists $t > 0$ such that the solution of (6.1) with the initial condition $x(0) = \delta y$ satisfies

$$|(x(t), y^\perp)| > \varepsilon \qquad \text{and} \qquad \|x(t)\| < c_1 \varepsilon, \tag{6.3}$$

where $y^\perp$ is the projection of y onto $M^\perp$.

Proof. The last statement implies the instability of the zero solution because we may fix $\varepsilon > 0$ and choose δ arbitrarily small. Thus every neighborhood of the origin contains initial data which launch solutions which exit from a fixed neighborhood of the origin. To prove the theorem, let $\|y\| = 1$, $c_1 = 6 \|y^\perp\|^{-2}$, $\eta = c_1 \varepsilon$, $\alpha = \operatorname{Re} \lambda$, $\sigma = 2\alpha - \mu > 0$, and $T = \alpha^{-1} \log(\eta / 2\delta)$. Let $x(0) = \delta y$ and $x(t)$ be the solution of (6.1). We claim that if δ and η are sufficiently small, then the solution exists and satisfies

$$\|x(t)\| \leqslant 2\delta\, e^{\alpha t} \leqslant l \qquad \text{for} \quad 0 \leqslant t \leqslant T. \tag{6.4}$$

To prove this, we note that $2\delta\, e^{\alpha T} = \eta \leqslant l$, the last inequality by choice of η. If (6.4) were false, let t be the first time $\leqslant T$ when $\|x(t)\| = 2\delta\, e^{\alpha t}$. By (6.1) and (6.2) we have

$$\begin{aligned}
\|x(t)\| &\leqslant \delta\, \|e^{tA} y\| + \int_0^t \|e^{(t-s)A}\|\, \|f(x(s))\|\, ds \\
&\leqslant \delta\, e^{\alpha t} + bk \int_0^t e^{\mu(t-s)} [2\delta\, e^{\alpha s}]^2\, ds \\
&= \delta\, e^{\alpha t} + bk4\delta^2\sigma^{-1}[e^{2\alpha t} - e^{\mu t}] \\
&< \delta\, e^{\alpha t}[1 + bk4\sigma^{-1}\,\delta\, e^{\alpha t}] \leqslant \delta\, e^{\alpha t}[1 + 2bk\sigma^{-1}\eta] \\
&\leqslant 2\delta\, e^{\alpha t} \leqslant \eta = c_1 \varepsilon
\end{aligned}$$

for sufficiently small η. This contradiction proves (6.4) and $\|x(t)\| < \eta \leqslant l$.

We now choose $t = \alpha^{-1} \log(\eta/3\delta)$ so that $\delta\, e^{\alpha t} = \eta/3$. Then the same estimation gives

$$\|x(t) - \delta\, e^{\alpha t} y\| < bk4\delta^2\sigma^{-1}\, e^{2\alpha t} = \delta\, e^{\alpha t}[4bk\eta/3\sigma] \leqslant \frac{\delta}{2} e^{\alpha t}\, \|y^{\perp}\|$$

for sufficiently small η. Therefore

$$|(x(t), y^{\perp})| > \frac{1}{2}\delta\, e^{\alpha t}\, \|y^{\perp}\|^2 = \frac{\eta}{6}\, \|y^{\perp}\|^2 = \varepsilon.$$

THEOREM 6.2. *Let $d''(\omega)$ be nonsingular, $n(H) - p(d'')$ be odd, and X be separable. Then $T(\exp t\omega)\varphi$ is G_ω-unstable.*

Proof. Let $M = Z = \{T_\sigma \varphi \mid \sigma \in \mathfrak{g}_\omega\}$. Let $A = JH$. By Theorem 5.1, A has a positive real eigenvalue λ, with an eigenvector y. Let $\lambda < \mu < 2\lambda$. For any solution $u(t)$ of (2.10), let

$$x(t) = T(e^{-t\omega})\, u(t) - \varphi.$$

Then

$$\begin{aligned}
\frac{dx}{dt} &= T(e^{-t\omega})[JE'(u(t)) - T_\omega u(t)] \\
&= T(e^{-t\omega})\, JT^*(e^{+t\omega})[E'(\varphi + x(t)) - Q'_\omega(\varphi + x(t))] \\
&= J\{E'(\varphi) + E''(\varphi)\, x(t) + O(\|x(t)\|^2) - Q'_\omega(\varphi) - Q''_\omega(\varphi)\, x(t)\} \\
&= Ax(t) + O(\|x(t)\|^2).
\end{aligned}$$

By Theorem 6.1, for all sufficiently small ε and δ, we may find $t>0$, such that if $x(0)=\delta y$ then

$$\|x(t)\| < c_1 \varepsilon \qquad \text{and} \qquad (x(t), y^\perp) > \varepsilon.$$

In some neighborhood of φ, the orbit $\{T(g)\varphi \mid g \in G_\omega\}$ is a smooth manifold by Lemma 4.3. We choose $h \in G_\omega$ so that

$$\Theta = \inf_{g \in G_\omega} \|u(t) - T(g)\varphi\| = \inf_{g \in G_\omega} \|x(t) + \varphi - T(g)\varphi\| = \|x(t) + \varphi - T(h)\varphi\|.$$

Then $\|x(t) + \varphi - T(h)\varphi\| \leqslant \|x(t) + \varphi - T(e)\varphi\| = \|x(t)\| < c_1\varepsilon$, so that $\|\varphi - T(h)\varphi\| < 2c_1\varepsilon$. Hence there exists $\sigma \in \mathfrak{g}_\omega$, such that

$$T(h)\varphi = \varphi + T_\sigma \varphi + O(\varepsilon^2),$$

where $|\sigma| = O(\varepsilon)$. Therefore,

$$\begin{aligned} \Theta &\geqslant (x(t) + \varphi - T(h)\varphi, y^\perp) \\ &= (x(t), y^\perp) - (T_\sigma\varphi, y^\perp) - O(\varepsilon^2) > \varepsilon - O(\varepsilon^2) > \varepsilon/2 \end{aligned}$$

for small ε because $y^\perp$ is orthogonal to $M = Z$. Fix ε and consider δ to be arbitrarily small. We have

$$\inf_g \|u(t) - T(g)\varphi\| > \varepsilon/2$$

with $\|u(0) - \varphi\| = \delta$, where $u(\cdot)$ and t depend on δ. This means G_ω-instability.

It is natural to ask whether the instability is simply due to the solitary wave staying close to the orbit of the full group G but away from the centralizer G_ω. The next theorem shows that this is not the case for certain groups, for instance, abelian or nilpotent ones.

THEOREM 6.3. *In addition to the conditions of Theorem* 6.2, *assume that*

(i) $[\tau, \omega] \neq \lambda\tau$ *for all* $0 \neq \tau \in \mathfrak{g}$, $0 \neq \lambda \in \mathbb{R}$, *and*

(ii) *the alternative of Lemma* 4.3 *is valid for the full group* G.

Then $T(\exp t\omega)\varphi$ *is* G*-unstable.*

Proof. The proof of Theorem 6.2 works perfectly for the whole group G provided the eigenvector y of JH does not belong to M, where $M = \{T_\sigma\varphi \mid \sigma \in \mathfrak{g}\}$. Assume, on the contrary, that $y \in M$. Then

$$JHy = \lambda y, \qquad \lambda \neq 0, \qquad y \neq 0, \qquad y = T_\sigma\varphi,$$

for some $\sigma \in \mathfrak{g}$. By Lemma 2.2,

$$T_{[\sigma,\omega]}\varphi = JQ'_{[\sigma,\omega]}\varphi = JHT_\sigma\varphi = \lambda T_\sigma\varphi = T_{\lambda\sigma}\varphi,$$

so that $T_{[\sigma,\omega]-\lambda\sigma}\varphi = 0$. Now define the linear map $\Lambda : \mathfrak{g} \to \mathfrak{g}$ by $\Lambda(\tau) = [\tau, \omega] - \lambda\tau$. By assumption (i), Λ is one–one. Hence Λ is onto. Let $\mathfrak{h} = \{\alpha \in \mathfrak{g} \mid T_\alpha\varphi = 0\}$. What we showed above was that $\Lambda(\sigma) \in \mathfrak{h}$. For all $\tau \in \mathfrak{h}$, we have (by the same calculation as above),

$$T_{[\tau,\omega]}\varphi = JHT_\tau\varphi = 0.$$

So $\Lambda : \mathfrak{h} \to \mathfrak{h}$. Therefore $\sigma \in \mathfrak{h}$ and $y = T_\sigma\varphi = 0$, which is a contradiction.

7. Coupled Wave Equations

Consider the system of 3-wave equations

$$u_{tt} - u_{xx} + f(u) = 0, \tag{7.1}$$

$x \in \mathbb{R}$, $u(x, t) \in \mathbb{R}^3$, where f has the form $f(u) = g(|u|^2)u$. Consider a "solitary wave" of the form

$$u = e^{tS}\varphi(x + \omega_0 t), \tag{7.2}$$

where $\omega_0 \in \mathbb{R}$ and S is a non-zero 3×3 skew matrix. We write $Sy = \omega \wedge y$, where $\omega \in \mathbb{R}^3$. Let v be any unit vector orthogonal to ω.

THEOREM 7.1. *Suppose that $f(u) = u - |u|^2 u$ and $|\omega|^2 + \omega_0^2 < 1$.*

(a) *Then there exists a solution of* (7.1) *of the form* (7.2) *with*

$$\varphi(x) = \eta(x)\, e^{\alpha x S} v,$$

where $\alpha = \omega_0(1 - \omega_0^2)^{-1}$ and $\eta(x)$ is a positive function decreasing exponentially as $|x| \to \infty$.

(b) *If $|\omega|$ is sufficiently small, this solution is G-unstable. If $|\omega|$ is sufficiently close to $(1 - \omega_0^2)^{1/2}$, it is G_ω-stable.*

Thus the rotational motion stabilizes the solitary wave. To prove the theorem, we write (7.1) in Hamiltonian form (1.1) on the space $X = H^1(\mathbb{R}; \mathbb{R}^3) \times L^2(\mathbb{R}; \mathbb{R}^3)$ with $J = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$,

$$E(\mathbf{u}) = \int_{-\infty}^{\infty} \left[\tfrac{1}{2}|\partial u|^2 + \tfrac{1}{2}|v|^2 + F(u)\right] dx,$$

$F(0)=0$, $F'=f$, $\mathbf{u}=$ the pair $[u, v]$, and $\partial = d/dx$. The equation is well-posed in the space X, locally in time. The group is $G = SO(3) \times \mathbb{R}$, which acts as rotations in u and translations in x, namely,

$$\mathbf{u}(x) \mapsto R\mathbf{u}(x+a) \qquad \text{for} \quad R \in SO(3),\, a \in \mathbb{R}.$$

Let $S_0 = \partial$ and let $\{S_1, S_2, S_3\}$ be a basis of the skew 3×3 matrices chosen so that $\omega \wedge y = \omega_1 S_1 y + \omega_2 S_2 y + \omega_3 S_3 y$. The corresponding "charges" (from (2.5) and (2.6)) are

$$Q_i(\mathbf{u}) = \int S_i u \cdot v\, dx, \qquad Q_0(\mathbf{u}) = \int \partial u \cdot v\, dx,$$

where $\mathbf{u} = [u, v]$ and $i = 1, 2, 3$.

A *solitary wave* has the form $[e^{tS}\varphi(x+\omega_0 t), e^{tS}\psi(x+\omega_0 t)]$, where $\boldsymbol{\varphi} =$ the pair $[\varphi, \psi]$. Then

$$E'(\boldsymbol{\varphi}) - \sum_{i=0}^{3} \omega_i Q_i'(\boldsymbol{\varphi}) = 0$$

takes the form

$$-\partial^2\varphi + f(\varphi) + (\omega_0\, \partial + S)^2 \varphi = 0 \tag{7.3}$$

with $\psi = (\omega_0\, \partial + S)\varphi$. We can eliminate the first derivative terms in (7.3) as follows. Assuming $\omega_0 \neq 1$, let $\alpha = \omega_0/(1-\omega_0^2)$ and $\zeta(x) = \exp(-x\alpha S)\, \varphi(x)$. Then Eq. (7.3) is converted to

$$\begin{aligned}
\partial^2\zeta &= e^{-x\alpha S}(\partial^2\varphi - 2\alpha S\, \partial\varphi + \alpha^2 S^2 \varphi) \\
&= e^{-x\alpha S}(1-\omega_0^2)^{-1}(f(\varphi) + S^2\varphi + \alpha^2 S^2 \varphi) \\
&= (1-\omega_0^2)^{-1} f(\zeta) + (1-\omega_0^2)^{-2} S^2 \zeta.
\end{aligned}$$

Thus ζ satisfies

$$-(1-\omega_0^2)\, \partial^2\zeta + (1-\omega_0^2)^{-1} S^2 \zeta + f(\zeta) = 0. \tag{7.4}$$

As stated above, we look for $\zeta(x) = \eta(x) v$, where $\eta(x)$ is a scalar function. Then $S^2\zeta = -|\omega|^2\zeta$, so that (7.4) reduces to the scalar equation

$$-(1-\omega_0^2)\, \partial^2\eta - |\omega|^2 (1-\omega_0^2)^{-1}\eta + f(\eta) = 0. \tag{7.5}$$

Assuming that $f'(0) > 0$, that $F(s)$ is somewhere negative, that $f(s_0) \neq 0$ for the smallest positive root of $F(s_0) = 0$, and that

$$|\omega|^2 (1-\omega_0^2)^{-1} < f'(0), \tag{7.6}$$

we know from Lemma 6.1 of [1] that there is a unique positive even solution $\eta(x)$ of (7.5) which decays exponentially such that $\eta(0)=s_0$. This proves part (a) of the theorem.

Now we analyze the spectrum of

$$H=E''(\boldsymbol{\varphi})-\sum_{i=0}^{3}\omega_i Q_i''(\boldsymbol{\varphi})=\begin{pmatrix}-\partial^2+f'(\varphi) & \omega_0\,\partial+S\\ -\omega_0\,\partial-S & 1\end{pmatrix}. \tag{7.7}$$

We show that $n(H)=1$. Writing $\mathbf{w}=$ the pair $[w, y]$, we have

$$\begin{aligned}\langle H\mathbf{w}, \mathbf{w}\rangle &= \langle[-\partial^2+f'(\varphi)]\,w, w\rangle-2\langle(\omega_0\,\partial+S)\,w, y\rangle+\langle w, y\rangle\\ &=\langle Lw, w\rangle+\langle w-(\omega_0\,\partial+S)\,y, \quad w-(\omega_0\,\partial+S)y\rangle,\end{aligned} \tag{7.8}$$

where $L=-\partial^2+f'(\varphi)+(\omega_0\,\partial+S)^2$ and $\langle\ \rangle$ is the L^2 inner product. We also write the L^2 norm as $\langle w\rangle=\langle w, w\rangle^{1/2}$. Because of (7.8), it suffices to show that $n(L)=1$.

Differentiating $f(s)=g(|s|^2)s$ for $s=(s_1, s_2, s_3)\in\mathbb{R}^3$, we have

$$\partial f_i/\partial s_j=g(|s|^2)\,\delta_{ij}+2g'(|s|^2)\,s_i s_j.$$

Therefore,

$$f'(\varphi)=g(\eta^2)I+2g'(\eta^2)\,\eta^2\Pi,$$

where Π is the orthogonal projection of $\mathbb{R}^3$ onto the line generated by $e^{x\alpha S}v$. Thus, for any function $w\in H^1(\mathbb{R}, \mathbb{R}^3)$,

$$\langle Lw, w\rangle=\int\{|\partial w|^2+g(\eta^2)\,|w|^2+2g'(\eta^2)\,\eta^2\,|\Pi w|^2+(\omega_0\,\partial+S)^2 w\cdot w\}\,dx.$$

Again we can eliminate the first-order derivative by the transformation $z(x)=\exp(-x\alpha S)\,w(x)$. Then a simple calculation, using the orthogonality of $\exp(-x\alpha S)$, yields

$$\begin{aligned}\langle Lw, w\rangle=\int\{&(1-\omega_0^2)\,|\partial z|^2-(1-\omega_0^2)^{-1}\,|Sz|^2\\ &+g(\eta^2)\,|z|^2+2g'(\eta^2)\,\eta^2\,|\Pi_v z|^2\}\,dx,\end{aligned}$$

where Π_v is the projection onto v itself. We split this expression into three components. Let m be a unit vector orthogonal to both $\hat{\omega}=\omega/|\omega|$ and v. Then

$$|Sz|^2=-S^2z\cdot z=+|\omega|^2\{(v\cdot z)^2+(m\cdot z)^2\}$$

and $|\Pi_v z|^2=(v\cdot z)^2$, so that the whole expression becomes

$$\begin{aligned}\langle Lw, w\rangle&=\langle A(v\cdot z), v\cdot z\rangle+\langle B(m\cdot z), m\cdot z\rangle\\ &\quad+\langle[B+(1-\omega_0^2)\,|\omega|^2](\hat{\omega}\cdot z), \hat{\omega}\cdot z\rangle,\end{aligned} \tag{7.9}$$

where A and B act on scalar functions as

$$A = -(1-\omega_0^2)\,\partial^2 - (1-\omega_0^2)\,|\omega|^2 + g(\eta^2) + 2g'(\eta^2)\eta^2,$$
$$B = -(1-\omega_0^2)\,\partial^2 - (1-\omega_0^2)\,|\omega|^2 + g(\eta^2).$$

From (7.5) it follows that $A\,\partial\eta = 0$ and $B\eta = 0$. Since η has no nodes, it is the ground state for B, and B has no negative eigenvalues. Since $\partial\eta$ has one node, it is the second eigenfunction for A; hence A has exactly one negative eigenvalue. Thus $1 = n(A) = n(L) = n(H)$.

Next we compute $d(\omega_0, \omega_1, \omega_2, \omega_3)$,

$$d = E(\boldsymbol{\varphi}) - \sum_{i=0}^{3} \omega_i Q_i(\boldsymbol{\varphi})$$
$$= \int \{\tfrac{1}{2}|\partial\varphi|^2 + F(\varphi) - \tfrac{1}{2}|(\omega_0\,\partial + S)\varphi|^2\}\,dx.$$

Using the same transformation as before and the orthogonality of $\exp(x\alpha S)$, this reduces to

$$d = \int \{(1-\omega_0^2)\,\tfrac{1}{2}|\partial\zeta|^2 - (1-\omega_0^2)^{-1}\,\tfrac{1}{2}|S\zeta|^2 + F(\zeta)\}\,dx$$

$$d = \int \{(1-\omega_0^2)\,\tfrac{1}{2}|\partial\eta|^2 - |\omega|^2(1-\omega_0^2)^{-1}\tfrac{1}{2}\eta^2 + F(\eta)\}\,dx. \qquad (7.10)$$

We must compute the Hessian of $d(\omega_0, ..., \omega_3)$ restricted to $\mathfrak{g}_\omega$. The latter consists of those elements of the Lie algebra which commute with $\omega_0 S_0 + \cdots + \omega_3 S_3 = \omega_0\,\partial + S$. This means we only consider matrices which commute with the given S. Therefore, what we must compute is the Hessian of the function $a, b \mapsto d(a, b\hat{\omega}_1, b\hat{\omega}_2, b\hat{\omega}_3)$ near $a = \omega_0$, $b = |\omega|$. This Hessian is denoted d''. We have already proven that

$$1 = n(H) = n(H_1) + p(d'') + (z(d'') - z_0), \qquad (7.11)$$

so that $p(d'') = 0$ or 1. To determine which case occurs, we specialize to the pure power case.

Let $f(s) = m^2 s - |s|^{p-1}s$, where $m > 0$ and $p > 1$. Rescale by $\eta(x) = \delta\xi(\lambda x)$, where δ and λ are constants. Then (7.5) becomes

$$(1-\omega_0^2)\,\delta\lambda^2\,\partial^2\xi + [m^2 - (1-\omega_0^2)^{-1}\,|\omega|^2]\,\delta\xi - \delta^p\,|\xi|^{p-1}\xi = 0.$$

Then $\xi(x)$ satisfies

$$\partial^2\xi + \xi - \xi^p = 0$$

and is independent of ω_0 and $|\omega|$, provided λ and δ are chosen so that

$$(1-\omega_0^2)\,\lambda^2=(1-\omega_0^2)^{-1}\alpha=\delta^{p-1}, \tag{7.12}$$

where $\alpha=m^2-m^2\omega_0^2-|\omega|^2>0$. Furthermore, in terms of $\xi(x)$, (7.10) takes the form

$$d=\delta^2\lambda^{-1}\int\{(1-\omega_0^2)\,\lambda^2\tfrac{1}{2}|\partial\xi|^2-\delta^{p-1}\xi^{p+1}/(p+1)$$
$$+\alpha(1-\omega_0^2)^{-1}\tfrac{1}{2}\xi^2\}\,dx$$
$$=\delta^{p+1}\lambda^{-1}\int\{\tfrac{1}{2}|\partial\xi|^2-\xi^{p+1}/(p+1)+\tfrac{1}{2}\xi^2\}\,dx.$$

Thus,

$$d(\omega_0,|\omega|)=c(1-\omega_0^2)^{-\mu}(m^2-m^2\omega_0^2-|\omega|^2)^{\mu+1/2}, \tag{7.13}$$

where $\mu=2/(p-1)$ and c is a constant.

Finally, we specialize to the case $f(u)=u-|u|^2u$ with $p=3$, $\mu=1$, and $m=1$. We calculate three limiting cases. If $\omega_0\to 0$, then

$$d''(\omega_0,|\omega|)\to(1-\omega_0^2+|\omega|^2)^{-1/2}\begin{pmatrix}-(1-|\omega|^2)(1+2\,|\omega|^2) & 0\\ 0 & -3(1-|\omega|^2)\end{pmatrix}.$$

Thus for ω_0 small, d'' has two negative eigenvalues and $n(H)-p(d'')=1$ and, by Theorem 6.3, the solitary wave is *G-unstable*. The same conclusion is true for $|\omega|$ small and ω_0 fixed. On the other hand, if $|\omega|\to(1-\omega_0^2)^{1/2}$ with $\omega_0\neq 0$ fixed, then $(1-\omega_0^2-|\omega|^2)^{1/2}\,\partial^2 d/\partial\,|\omega|^2\to 3(1-\omega_0^2)>0$. So for $|\omega|$ near $(1-\omega_0^2)^{1/2}$, $p(d'')=1$, so that by the Stability Theorem the solitary wave is *G_ω-stable.*

8. Harmonic Map into a Sphere

Consider a harmonic map $\varphi\colon S^1\times\mathbb{R}\mapsto N$ between $S^1\times\mathbb{R}$ equipped with the metric $d\vartheta^2-dt^2$ and a complete Riemannian manifold (N,g). It satisfies the equation

$$\partial_t^2\varphi^a-\partial_\vartheta^2\varphi^a-\Gamma^a_{bc}(\partial_t\varphi^b\,\partial_t\varphi^c-\partial_\vartheta\varphi^b\,\partial_\vartheta\varphi^c)=0,$$

where Γ^a_{bc} are the Christoffel symbols. Taking the case $N=S^2$, we can embed $S^2\subset\mathbb{R}^3$ and use the coordinates $u=(u_1,u_2,u_3)$ on $\mathbb{R}^3$ to write

$$\partial_t^2u-\partial_\vartheta^2u+(|\partial_tu|^2-|\partial_\vartheta u|^2)u=0, \tag{8.1}$$

where $|u(\vartheta, t)| = 1$ and $u(0, t) = u(2\pi, t)$. The symmetry group is that of the target space together with the translations in S^1; that is, $G = SO(3) \times S^1$.

Consider the solitary wave

$$u(\vartheta, t) = e^{(k\vartheta + \omega t)A} v, \tag{8.2}$$

where A is a skew 3×3 matrix, v is a unit vector such that $A^2 v = -v$, ω is a real number, and k is an integer.

THEOREM 8.1. *If $\omega^2 < k^2$ the solitary wave is G-unstable. If $\omega^2 > k^2$ it is G_ω-stable.*

This problem does not fit our setting because the function u lies on a curved manifold, not a linear space. The definition of stability is, however, the same as before (2.13), using the norm in $X = H^1(S^1; \mathbb{R}^3) \times L^2(S^1; \mathbb{R}^3)$. One way around our difficulty would be to modify our setting to allow manifolds. But it is easier to avoid that task as follows. We consider the perturbed equation

$$\partial_t^2 u - \partial_\vartheta^2 u + \Lambda(|u|^2 - 1)u = 0, \tag{8.3}$$

$u(2\pi, t) = u(0, t)$ with $u(\vartheta, t) \in \mathbb{R}^3$, where Λ is a large parameter. If the initial datum $u(\vartheta, 0)$ has unit magnitude, the solution of (8.3) converges strongly in X as $\Lambda \to \infty$ to the solution of (8.1). Both equations have the same symmetry group. We see easily that (8.2) is a solution of (8.3) if

$$A^2 v = -v \qquad \text{and} \qquad |v|^2 = 1 + (\omega^2 - k^2)/\Lambda. \tag{8.4}$$

As $\Lambda \to \infty$, this solitary wave converges to the solitary wave solution of (8.1) for the same k and ω. The stability or instability of (8.3) does not imply the stability of (8.1). However, the linearized operator H will retain the same spectral properties in the limit of $\Lambda \to \infty$ and therefore we will be able to deduce Theorem 8.1 from the corresponding theorem for (8.3) for large Λ.

The analysis of Eq. (8.3) is almost identical with that of Section 7. The variable ϑ pays the role of x, $\exp(k\vartheta A)v$ the role of $\varphi(x)$, ωA the role of S, and $\omega_0 = 0$. Furthermore, $f(\varphi) = \Lambda(|\varphi|^2 - 1)\varphi$, so that

$$f'(\varphi) = \Lambda(|\varphi|^2 - 1) + 2\Lambda\langle\varphi, \ \rangle\varphi = \omega^2 - k^2 + 2\Lambda\langle\varphi, \ \rangle\varphi.$$

Therefore, (7.7) takes the form

$$H = \begin{pmatrix} -\partial^2 + \omega^2 - k^2 + 2\Lambda\langle\varphi, \ \rangle\varphi & \omega A \\ -\omega A & 1 \end{pmatrix}.$$

Writing $\mathbf{w} = [w, y]$ as in (7.8), we have

$$\langle H\mathbf{w}, \mathbf{w}\rangle = \langle(-\partial^2 + \omega^2 - k^2)\, w, w\rangle + 2\Lambda\langle e^{k\vartheta A}v, w\rangle + 2\omega\langle Aw, y\rangle + \langle y\rangle^2$$
$$= \langle Lw, w\rangle + \langle \omega Aw + y\rangle^2,$$

where

$$\langle Lw, w\rangle = \langle(-\partial^2 + \omega^2 - k^2)\, w, w\rangle - \omega^2\langle Aw\rangle^2 + 2\Lambda\langle e^{k\vartheta A}v, w\rangle^2.$$

We define $p_1(\vartheta) = e^{k\vartheta A}v$ and $p_2(\vartheta) = Ap_1(\vartheta)$ and let p_3 be a constant vector of length $|v|$ orthogonal to both v and Av. For each ϑ and k, these form an orthogonal basis of $\mathbb{R}^3$ such that $Ap_1 = p_2$, $Ap_2 = -p_1$, and $Ap_3 = 0$. Indeed, $Ap_2 = e^{k\vartheta A}A^2v = -e^{k\vartheta A}v = -p_1$, while for some unit vector a we have $Av = a \wedge v$ and $p_3 = |v|\, a$ so that $Ap_3 = a \wedge |v|\, a = 0$. Also $|p_1| = |p_2| = |p_3| = |v|$.

Expanding $w(\vartheta) = \alpha_1(\vartheta)\, p_1(\vartheta) + \alpha_2(\vartheta)\, p_2(\vartheta) + \alpha_3(\vartheta)\, p_3$, we therefore have $Aw = \alpha_1 p_2 - \alpha_2 p_1$ and $|Aw|^2 = (\alpha_1^2 + \alpha_2^2)\, |v|^2$. Furthermore,

$$\begin{aligned}\langle(-\partial^2 - k^2)w, w\rangle &= \langle \partial w\rangle^2 - k^2\langle w\rangle^2\\ &= |v|^2\{\langle \partial\alpha_1 - k\alpha_2\rangle^2 + \langle \partial\alpha_2 + k\alpha_1\rangle^2\\ &\quad + \langle\partial\alpha_3\rangle^2 - k^2(\langle\alpha_1\rangle^2 + \langle\alpha_2\rangle^2 + \langle\alpha_3\rangle^2)\}\\ &= |v|^2\{\langle\partial\alpha_1\rangle^2 + \langle\partial\alpha_2\rangle^2 + \langle\partial\alpha_3\rangle^2\\ &\quad + 4k\langle\alpha_1, \partial\alpha_2\rangle - k^2\langle\alpha_3\rangle^2\}.\end{aligned}$$

Hence,

$$\begin{aligned}\langle Lw, w\rangle &= |v|^2\{\langle\partial\alpha_1\rangle^2 + \langle\partial\alpha_2\rangle^2 + \langle\partial\alpha_3\rangle^2\\ &\quad + 4k\langle\alpha_1, \partial\alpha_2\rangle + (\omega^2 - k^2)\langle\alpha_3\rangle^2 + 2\Lambda\langle\alpha_1\rangle^2\}\\ &= |v|^2\{\langle\partial\alpha_1\rangle^2 + \langle\partial\alpha_2 + 2k\alpha_1\rangle^2 + \langle\partial\alpha_3\rangle^2\\ &\quad + (\omega^2 - k^2)\langle\alpha_3\rangle^2 + (2\Lambda - 4k^2)\langle\alpha_1\rangle^2\}.\end{aligned}$$

Supposing that $\omega^2 > k^2$ and $\Lambda > 2k^2$, we therefore have $\langle H\mathbf{w}, \mathbf{w}\rangle \geqslant 0$. We have $\langle H\mathbf{w}, \mathbf{w}\rangle = 0$ only if $\alpha_1 = \alpha_3 = \partial\alpha_2 = 0$ and $y = -\omega Aw$, which means that

$$\mathbf{w} = \alpha_2\begin{bmatrix} p_2 \\ -\omega p_1\end{bmatrix} = \alpha_2\begin{bmatrix} Ap_1 \\ A(\omega p_2)\end{bmatrix} = T_A \boldsymbol{\varphi}_\omega.$$

Thus the kernel of H is exactly the set Z defined in Section 2, since the commutator subgroup is the one-parameter group generated by A. Except for this kernel, the spectrum of H is bounded away from zero, independently of Λ for Λ large. So the same is true for (8.1) and the stability follows as in Section 4.

In case $\omega^2 < k^2$, we let $l = (k^2 - \omega^2)^{1/2}$ and notice that

$$JH\begin{bmatrix} p_3 \\ -lp_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} -\partial^2 + \omega^2 - k^2 + 2A\langle p_1, \ \rangle p_1 & \omega A \\ -\omega A & 1 \end{bmatrix}\begin{bmatrix} p_3 \\ -lp_3 \end{bmatrix}$$

$$= \begin{bmatrix} -\omega A & 1 \\ \partial^2 - \omega^2 + k^2 - 2A\langle p_1, \ \rangle p_1 & -\omega A \end{bmatrix}\begin{bmatrix} p_3 \\ -lp_3 \end{bmatrix}$$

$$= \begin{bmatrix} -lp_3 \\ (-\omega^2 + k^2) p_3 \end{bmatrix} = -l\begin{bmatrix} p_3 \\ -lp_3 \end{bmatrix},$$

because $Ap_3 = \partial^2 p_3 = \langle p_1, p_3 \rangle = 0$. Thus JH has a negative eigenvalue with an eigenvector independent of A. So we have instability for large A as well as in the limiting case $A \to \infty$.

9. Coupled Schrödinger Equations

Consider the system of m equations

$$K\frac{\partial u}{\partial t} = -\Delta u + f(u), \tag{9.1}$$

where $u(x, t) \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, K is an invertible, skew real $m \times m$ matrix and f: $\mathbb{R}^m \mapsto \mathbb{R}^m$. Consider a solitary wave of the form

$$u = e^{tS}\varphi(x + ct), \tag{9.2}$$

where $c \in \mathbb{R}^n$ and S is a skew $m \times m$ matrix. We assume that

$$S \text{ commutes with } K \tag{9.3}$$

and that

$$M = \tfrac{1}{4}|c|^2 K^2 - KS \text{ is positive definite.} \tag{9.4}$$

We assume that $f(Ru) = Rf(u)$ for $R \in SO(m)$, $f(0) = f'(0) = 0$, $f'(s) = o(|s|^{4/(n-2)})$ as $|s| \to \infty$, and the primitive F is somewhere negative (where $F' = f$, $F(0) = 0$). Let $G = \mathbb{R}^n \times SO(m)$ acting as translations in x and rotations in u.

Theorem 9.1. *Under these conditions there exists a family of solitary waves* (9.2) *which are solutions of* (9.1) *and which depend on the smallest eigenvalue* μ *of the matrix* M. *Such a solitary wave* (9.2) *is* G_ω*-stable if*

$$\frac{\partial}{\partial \mu}\int_{\mathbb{R}^n} |\varphi(x)|^2 \, dx > 0, \tag{9.5}$$

and G_ω*-unstable if this quantity is negative.*

EXAMPLE. The nonlinear Schrödinger equation

$$i\frac{\partial u}{\partial t} = -\Delta u + f(u), \qquad u(x, t) \in \mathbb{C}, \tag{9.6}$$

is a special case of (9.1). Indeed, splitting (9.6) into its real and imaginary parts, we find $m = 2$ and

$$K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad J = K^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

If we choose $S = \omega_0 K$, then $M = (\omega_0 - |c|^2/4)I$. The condition (9.4) requires $\mu = \omega_0 - |c|^2/4 > 0$. The solitary wave (9.2) takes the form

$$u = \eta(x + ct) \exp i\{-\tfrac{1}{2}c \cdot x + (\omega_0 - \tfrac{1}{2}|c|^2)t\}, \tag{9.7}$$

where η is the positive radial solution of

$$-\Delta\eta + (\omega_0 - |c|^2/4)\eta + f(\eta) = 0. \tag{9.8}$$

This immediately yields the following result.

COROLLARY 9.2. *The solution* (9.7) *is G-stable if*

$$\frac{\partial}{\partial \mu} \int \eta^2(x)\, dx > 0$$

and unstable if it is < 0.

In fact, the general system (9.1) is just a coupled system of Schrödinger equations. Indeed, the change of variables $u \to Au$ for a nonsingular matrix A leads to the same equation with K replaced by a similar matrix. So we may assume that K is composed of $m/2$ 2×2 blocks of the form $[\begin{smallmatrix} 0 & -\alpha_j \\ \alpha_j & 0 \end{smallmatrix}]$ $(j = 1, ..., m/2)$. Thus (9.1) is just a coupled system of $m/2$ Schrödinger equations.

Proof of Theorem 9.1. Substituting (9.2) into (9.1), we find that φ must satisfy the elliptic equation

$$-\Delta\varphi + f(\varphi) - K(S + c \cdot \nabla)\varphi = 0. \tag{9.9}$$

We eliminate the first-order term by the substitution

$$\varphi(x) = e^{-c \cdot xK/2}\psi(x). \tag{9.10}$$

Then (9.9) takes the form

$$-\Delta\psi + (c \cdot \nabla) K\psi - \tfrac{1}{4}|c|^2 K^2\psi + f(\psi) - KS\psi - K(c \cdot \nabla)\psi + \tfrac{1}{2}|c|^2 K^2\psi = 0,$$

which simplifies to

$$-\Delta\psi + f(\psi) + M\psi = 0 \tag{9.11}$$

with M defined by (9.4). Let $\mu > 0$ be the smallest eigenvalue of the matrix M and let $v \in \mathbb{R}^m$ be a corresponding unit eigenvector ($Mv = \mu v$). Since M commutes with K, v is also an eigenvector of K. Now we let $\psi(x) = \eta(x)v$, where $\eta: \mathbb{R}^n \mapsto \mathbb{R}$ is the ground state of the equation

$$-\Delta\eta + f(\eta) + \mu\eta = 0. \tag{9.12}$$

Then ψ depends on c and S only through the quantity μ.

To set up the problem abstractly, we let $X = H^1(\mathbb{R}^n; \mathbb{R}^m)$, $J = K^{-1}$ and

$$E(u) = \int_{\mathbb{R}^n} \{\tfrac{1}{2}|\nabla u|^2 + F(u(x))\}\, dx,$$

where we write $f(u) = g(|u|^2)u$, $\mathscr{G}' = g$, $\mathscr{G}(0) = 0$, and $F(u) = \frac{1}{2}\mathscr{G}(|u|^2)$. The symmetry group is $G = \mathbb{R}^n \times SO(m)$ which acts as $u(x) \mapsto Au(x+a)$, for $a \in \mathbb{R}^n$ and $A \in SO(m)$. We choose any basis $S_1, ..., S_p$ of the skew matrices where $p = m(m-1)/2$ and coordinates $x_1, ..., x_n$ of $\mathbb{R}^n$. Then the "charges" are

$$Q_j(u) = \tfrac{1}{2}\int u \cdot KS_j u\, dx \qquad (j = 1, ..., p)$$

and

$$Q_k(u) = \tfrac{1}{2}\int u \cdot K\,\partial_k u\, dx \qquad (k = 1, ..., n),$$

where $\partial_k = \partial/\partial x_k$.

The matrix S of (9.2) is expanded as $S = \sum_{j=1}^{p} \omega_j S_j$, and then the parameters of our problem are $\omega_1, ..., \omega_p$, $c_1, ..., c_n$. The linearized Hamiltonian is

$$\begin{aligned} H &= E''(\varphi) - \sum_{j=1}^{p} \omega_j Q_j''(\varphi) - \sum_{k=1}^{n} c_k Q_k''(\varphi), \\ H &= -\Delta + f'(\varphi) - K(S + c\cdot\nabla), \end{aligned} \tag{9.13}$$

because K and S commute. For any $y \in X$,

$$\langle Hy, y\rangle = \langle \nabla y\rangle^2 + \langle f'(\varphi)y, y\rangle - \langle K(S + c\cdot\nabla)\,y, y\rangle,$$

where

$$\langle f(\varphi)y, y\rangle = \int \{g(|\varphi|^2)\,|y|^2 + 2g'(|\varphi|^2)(\varphi\cdot y)^2\}\,dx \qquad (9.14)$$

Substituting $y(x) = \exp\{-\frac{1}{2}c\cdot xK\}\,z(x)$, we get

$$\begin{aligned}\langle Hy, y\rangle &= \langle \nabla z\rangle^2 - \langle \nabla z, Kcz\rangle + \tfrac{1}{4}|c|^2\langle Kz\rangle^2 + \langle f'(\psi)z, z\rangle - \langle KSz, z\rangle \\ &\quad - \langle Kc\cdot\nabla z, z\rangle + \tfrac{1}{2}|c|^2\langle K^2 z, z\rangle \\ &= \langle \nabla z\rangle^2 + \langle Mz, z\rangle + \langle f'(\psi)z, z\rangle,\end{aligned}$$

repeatedly using the fact that K is skew and $\exp\{-c\cdot xK/2\}$ is orthogonal. We split

$$z(x) = z_1(x)v + z_2(x), \qquad v\cdot z_2(x) = 0,$$

where $z_1(x)$ is a scalar function. Using $Mv = \mu v$ and the form of the non-linear term in (9.14), we may write

$$\langle Hy, y\rangle = \langle Az_1, z_1\rangle + \langle Bz_2, z_2\rangle + \langle (M-\mu)z_2, z_2\rangle, \qquad (9.15)$$

where

$$A = -\Delta + \mu + g(\eta^2) + 2g'(\eta^2)\eta^2, \qquad B = -\Delta + \mu + g(\eta^2).$$

The third term in (9.15) is non-negative, since μ is the smallest eigenvalue of M. Since $B\eta = 0$ and η has no nodes, the second term in (9.15) is also non-negative. Since $A(\partial\eta/\partial x_j) = 0$ and $\partial\eta/\partial x_j$ has one node, it follows that $n(H) = n(A) = 1$.

Next we calculate

$$\begin{aligned}d(\omega, c) &= E(\varphi) - \sum_{j=1}^{p} \omega_j Q_j(\varphi) - \sum_{k=1}^{n} c_k Q_k(\varphi) \\ &= \int \{\tfrac{1}{2}|\nabla\varphi|^2 + F(\varphi) - \tfrac{1}{2}\varphi\cdot K(S + c\cdot\nabla)\varphi\}\,dx.\end{aligned}$$

Substituting (9.10) again, this simplifies to

$$\begin{aligned}d(\omega, c) &= \int \{\tfrac{1}{2}|\nabla\psi|^2 - \tfrac{1}{2}\nabla\psi\cdot Kc\psi + \tfrac{1}{8}|c|^2\,|K\psi|^2 + F(\psi) - \tfrac{1}{2}\psi\cdot KS\psi \\ &\quad + \tfrac{1}{4}|c|^2\psi\cdot K^2\psi - \tfrac{1}{2}\psi\cdot(c\cdot\nabla)\,K\mu\}\,dx \\ &= \int \{\tfrac{1}{2}|\nabla\psi|^2 + \tfrac{1}{2}\psi\cdot M\psi + F(\psi)\}\,dx \\ &= \tfrac{1}{2}\int \{|\nabla\eta|^2 + \mu\eta^2 + \mathscr{G}(\eta^2)\}\,dx.\end{aligned}$$

Therefore, $d(\omega, c)$ only depends on the single parameter μ. It follows from (9.12) that

$$\frac{\partial d}{\partial \mu} = \frac{1}{2} \int \eta^2 \, dx.$$

We denote the last quantity by q and let $q_\mu = \partial q / \partial \mu$. The centralizer depends on the $(n+1)$ parameters $c_1, ..., c_n$ and $b = |\omega|$. Writing $S = b\hat{S}$, we have

$$0 < \mu = v \cdot Mv = -\tfrac{1}{4} |c|^2 \, |Kv|^2 + b\hat{S}v \cdot Kv.$$

What we call d'' is the $(n+1) \times (n+1)$ matrix with the entries

$$\frac{\partial^2 d}{\partial b^2} = (\hat{S}v \cdot Kv)^2, \qquad \frac{\partial^2 d}{\partial b \, \partial c_k} = -\frac{1}{2} c_k |Kv|^2 (\hat{S}v \cdot Kv) q_\mu$$

and

$$\frac{\partial^2 d}{\partial c_k \, \partial c_l} = \frac{1}{4} c_k c_l \, |Kv|^2 \, q_\mu - \frac{1}{2} \, |Kv|^2 \, \delta_{kl} q.$$

Hence

$$\det(d'') = (\hat{S}v \cdot Kv)^2 q_\mu (-\tfrac{1}{2} |Kv|^2 q)^n,$$

whose sign is $(-1)^n$ times the sign of q_μ.

Since $n(H) = 1$ we have $p(d'') = 0$ or 1, so that d'' has at least n negative eigenvalues $\lambda_1, ..., \lambda_n$. If the other eigenvalue is called λ_0, then

$$(-1)^n \operatorname{sign} \lambda_0 = \operatorname{sign}(\lambda_0 \cdots \lambda_n) = \operatorname{sign}(\det d'') = (-1)^n \operatorname{sign}(q_\mu).$$

If $q_\mu > 0$ then $\lambda_0 > 0$ and $p(d'') = 1 = n(H)$. This is the stable case. If $q_\mu < 0$, then $\lambda_0 < 0$ and $n(H) - p(d'') = 1 - 0 = 1$. This is the unstable case.

10. Optical Wave Guide

We consider the higher modes of the same model as in [6, 1]. The equation is

$$iu_t + u_{xx} + g(x, |u|^2) u = 0, \tag{10.1}$$

where $g(x, |u|^2) = \eta_1 + \alpha \, |u|^2$ for $|x| > d$ and $g(x, |u|^2) = \eta_0 > \eta_1$ for $|x| < d$.

We consider standing waves $u = \varphi(x)\exp(i\omega t)$ with $\eta_1 < \omega < \eta_0$ and $\varphi(x)$ real. Then φ satisfies the equation

$$-\varphi_{xx} - g(x, \varphi^2)\varphi + \omega\varphi = 0. \tag{10.2}$$

As outlined in [6, 1], the solution curve bifurcates at a critical value $\omega = \omega_c$, for the higher mode as well as the fundamental mode.

Here we concentrate on the case when $\omega_c < \omega < \eta_0$. For each integer $n \geqslant 0$ there are one symmetric solution and two antisymmetric solutions with exactly n zeros (nodes). Each of them uses the center (the linear medium) to rotate n times and the homoclinic orbit (the nonlinear medium) to die off at $\pm\infty$. The linearized Hamiltonian is $H = [\begin{smallmatrix} R & 0 \\ 0 & S \end{smallmatrix}]$, where $S = \partial^2 - g(x, \varphi^2) + \omega$ and

$$R = S - \frac{\partial g(x, \varphi^2)}{\partial \varphi}\varphi = \begin{cases} -\partial^2 + \omega - \eta_1 - 3\alpha\varphi^2 & \text{for} \quad |x| > d \\ -\partial^2 + \omega - \eta_0 & \text{for} \quad |x| < d. \end{cases}$$

The number of negative eigenvalues of these operators is

$$n(S) = n, \qquad n(R) = \begin{cases} n+2 & \text{for} \quad \varphi \text{ symmetric} \\ n+1 & \text{for} \quad \varphi \text{ asymmetric.} \end{cases} \tag{10.3}$$

To prove this, note that $S\varphi = 0$, so that the number n of zeros of φ equals the number of negative eigenvalues $n(S)$ of S, by a well-known argument based on the Sturm comparison theorem. As for $n(R)$, note that $\psi = \varphi_x$ satisfies the equation

$$R\psi = (\eta_1 - \eta_0 + 3\varphi^2)\,\psi[\delta(x-d) - \delta(x+d)],$$

where $\delta(\cdot)$ denotes the delta function, because of the discontinuities at $x = \pm d$. For the symmetric solution, $h = \eta_1 - \eta_0 + 3\varphi^2 > 0$ at $x = \pm d$. For the asymmetric solution, $h(d)\,h(-d) < 0$. A comparison argument proves (10.3).

Thus for the symmetric solution, $n(H) = n(R) + n(S) = (n+2) + n = 2n+2$. These solutions correspond to the curve BE in the bifurcation diagram (Fig. 2 of [1]) so that

$$d''(\omega) = \frac{d}{d\omega}\int \frac{1}{2}|\varphi_\omega|^2\,dx > 0.$$

Hence $n(H) - p(d'') = (2n+2) - 1$ is odd. Therefore *the symmetric solutions are unstable.*

For an asymmetric solution (a pair of them for each n), we have $n(H) = (n+1)+n = 2n+1$. They correspond to the branch BC in the bifurcation diagram in [1], along which $d''(\omega) < 0$. Hence $n(H) - p(d'') = (2n+1) - 0$ is again odd. So *the asymmetric solutions are also unstable.*

On the branches AB and CD, we have $d''(\omega) > 0$, $n(R) = n+1$, and $n(S) = n$, so that $n(H) - p(d'') = (n+1) + n - 1 = 2n$ is even. For $n=0$ these branches are stable, as we pointed out in [1], but for $n>0$ our analysis is inconclusive.

11. Errata for [1]

Page 167. In case $d'' \equiv 0$ in a one-sided neighborhood $\mathscr{I}$ of ω, the correct definition of stability is

$$\sup_{0<t<\infty} \inf_{s\in\mathbb{R}} \inf_{\Omega\in\mathscr{I}} \|u(t) - T(s)\,\varphi_\Omega\| < \varepsilon.$$

Page 181. As stated, Theorem 5.4 requires the assumption that d'' not be identically zero in a one-sided neighborhood of ω. The proof is faulty because $\|u-\varphi_\omega\| \neq O(\|y\|)$. For a correct proof, see Section 4 of this paper.

Page 186. In (6.8), L_0 should be replaced by $-\partial_x^2 + f'(\varphi_\omega)$.

Page 187. The proof of Lemma 6.2 is wrong and should be replaced by the following one. We verify Assumption 3B on page 183. A simple calculation gives $\langle H_\omega u, u\rangle = \langle L_\omega u_1, u_1\rangle + \langle u_2 - \omega\,\partial_x u_1\rangle^2$, where $u = [u_1, u_2]$. Let $\mathscr{X}_1$ be a negative eigenvector of L_ω, and $\mathscr{X}_2 = \omega\,\partial_x \mathscr{X}_1$ and $\mathscr{X} = [\mathscr{X}_1, \mathscr{X}_2]$. Then $\langle H\mathscr{X}, \mathscr{X}\rangle = \langle L_\omega \mathscr{X}_1, \mathscr{X}_1\rangle < 0$. Also H_ω has the null vector $[\partial_x\varphi, \omega\,\partial_x^2\varphi]$. For any vector $u \in X$, such that $\langle u_1, \mathscr{X}_1\rangle = \langle u_1, \partial_x\varphi\rangle = 0$, we then have $\langle H_\omega u, u\rangle \geqslant \delta\,\|u\|^2$ for some $\delta > 0$.

Page 194. Figure 3 is incorrectly drawn. The dotted curve should cross the dashed curve only once in each quadrant. The solid curve should start on the dashed curve, switch to the dotted curve and switch back to the dashed curve.

References

1. M. Grillakis, J. Shatah, and W. Strauss, Stability theory of solitary waves in the presence of symmetry, I, *J. Funct. Anal.* **74** (1987), 160–197.
2. P. Blanchard, J. Stubbe, and L. Vazquez, *Ann. Inst. H. Poincaré* **47** (1987), 309–336.
3. M. Grillakis, Linearized instability for nonlinear Schrödinger and Klein–Gordon equations, *Comm. Pure Appl. Math.* **41** (1988), 747–774.
4. M. Grillakis, Analysis of the linearization around a critical point of an infinite-dimensional Hamiltonian system, *Comm. Pure Appl. Math.* **43** (1990), 299–333.

5. C. Jones, An instability mechanism for radially symmetric standing waves of a nonlinear Schrödinger equation, *J. Differential Equations* **71** (1988), 34–62.
6. C. Jones and J. Moloney, Instability of nonlinear waveguide modes, *Phys. Lett. A* **117** (1986), 175–180.
7. Y.-G. Oh, A stability theory for Hamiltonian systems with symmetry, *J. Geom. Phys.* **4** (1987), 163–182.
8. W. Strauss, Stability theory of solitary waves with invariants, *Abstracts Amer. Math. Soc.* **7** (Jan. 1986), 72.