STABILITY INDICES
FOR CONSTRAINED SELF-ADJOINT OPERATORS

TODD KAPITULA AND KEITH PROMISLOW

(Communicated by Yingfei Yi)

ABSTRACT. A wide class of problems in the study of the spectral and or-
bital stability of dispersive waves in Hamiltonian systems can be reduced to
understanding the so-called "energy spectrum", that is, the spectrum of the
second variation of the Hamiltonian evaluated at the wave shape, which is
constrained to act on a closed subspace of the underlying Hilbert space. We
present a substantially simplified proof of the negative eigenvalue count for
such constrained, self-adjoint operators, and extend the result to include an
analysis of the location of the point spectra of the constrained operator rel-
ative to that of the unconstrained operator. The results are used to provide
a new proof of the Jones-Grillakis instability index for generalized eigenvalue
problems of the form $(\mathcal{R} - z\mathcal{S})u = 0$ via a careful analysis of the associated
Krein matrix.

1. INTRODUCTION

It was first conjectured by Boussinesq [3] that the stability of families of disper-
sive waves for Hamiltonian systems can be reduced to the study of the minimizers
of the Hamiltonian constrained by the quantities conserved under the flow. This
idea was first developed into a rigorous stability proof by Benjamin for the soliton
solution of the Korteweg-de Vries equation [1], applied to the ground state solu-
tion of the nonlinear Schrödinger equation by [22], extended to a wide class of wave
equations by [2], and finally broadened to a general class of Hamiltonian systems by
Grillakis, Shatah, and Strauss in their now classic papers [9, 10]. They considered
systems of the form

$$(1.1) \qquad u_t = \mathcal{J} \frac{\delta \mathcal{H}}{\delta u}(u),$$

where $\mathcal{J}$ is a skew operator with respect to the inner product on the Hilbert space
H and $\mathcal{H}$ is the Hamiltonian. For a traveling wave solution ϕ of (1.1) the underlying
idea is to understand the spectrum of the second variation, $\mathcal{L}$, of the Hamiltonian
evaluated at ϕ as an operator acting on a subspace of H orthogonal to the con-
straints induced by the conserved quantities of the Hamiltonian flow. The operator
$\mathcal{L}$ is self-adjoint, and assuming it has at most a finite number of negative eigenval-
ues (counting multiplicity), and that there is a $\mu > 0$ which bounds the essential
spectrum of $\mathcal{L}$ from below, $\sigma_e(\mathcal{L}) \subset [\mu, +\infty)$, then [10] characterized the spectrum

of the constrained operator. In particular, they showed that accounting for the conserved quantities of the underlying system in effect constrains the operator $\mathcal{L}$ to act on a finite co-dimensional space $S^\perp$. If the constrained operator had no negative eigenvalues, then under a nondegeneracy condition the wave is orbitally stable. Furthermore, they showed that if the constrained operator has an odd number of negative eigenvalues, then the wave is unstable with the instability realized via at least one positive real-valued eigenvalue for the operator $\mathcal{JL}$. Interestingly, this instability criterion was predated by the Jones/Grillakis stability index [8, 12], which provides a sharper criterion for the existence of positive real-valued eigenvalues and was arrived at independently by Jones and by Grillakis using markedly different methods. The general problem of relating the spectra of the constrained operator to that of $\mathcal{JL}$ has been visited in [6, 11, 15, 21].

In this paper we present a simplified proof of a stability index which relates the spectrum of $\mathcal{L}$ to that of its constrained form $\mathcal{L}_{S^\perp}$, defined in (2.1). We focus on two issues:

(1) the number of negative eigenvalues of $\mathcal{L}_{S^\perp}$,
(2) the location of the eigenvalues of $\mathcal{L}_{S^\perp}$ relative to those of $\mathcal{L}$.

The first issue was resolved by the index theorem of [5, Section 3], which, for any S perpendicular to $\ker(\mathcal{L})$ constructed a real meromorphic matrix which is singular precisely at the eigenvalues of $\mathcal{L}_{S^\perp}$. We develop herein a concise proof which is functional analytic in nature and is based upon a decomposition of the quadratic form of $\mathcal{L}$ first introduced in the stability proof of [10]. We also provide a novel answer to the second issue via the Eigenvalue Interlacing Theorem.

The paper is organized as follows. In Section 2 we rederive the negative eigenvalue index for the operator $\mathcal{L}$ when constrained to act on $S^\perp$ (Index Theorem). We then compare the location of the negative eigenvalues of $\mathcal{L}$ with those of $\mathcal{L}$ constrained to $S^\perp$, developing the Eigenvalue Interlacing Theorem. In Section 3 we use the Index Theorem in combination with the recent formulation of the Krein matrix [13] in order to present a new and relatively simple proof of the Jones/Grillakis instability index [8, 12].

2. Negative eigenvalues for a constrained self-adjoint operator

Consider a self-adjoint operator $\mathcal{L}$ with its domain a Hilbert space H with inner product $\langle \cdot, \cdot \rangle$. We assume that there is a $\delta > 0$ such that $\mathcal{L}$ has a finite number of eigenvalues (counting multiplicity) for $\lambda < \delta$. We denote by the negative ($\mathrm{n}(\mathcal{L})$), zero ($\mathrm{z}(\mathcal{L})$), and positive indices ($\mathrm{p}(\mathcal{L})$) the total number (counting multiplicity) of negative, zero, and positive eigenvalues respectively of $\mathcal{L}$ acting on H. For a subspace $\tilde{H} \subset H$ we denote by $\mathrm{n}_{\tilde{H}}(\mathcal{L})$ the dimension of the maximal subspace of $\tilde{H}$ for which the bilinear form $\langle w, \mathcal{L}w \rangle < 0$ for all $w \in \tilde{H}$, with analogous definitions for $\mathrm{z}_{\tilde{H}}(\mathcal{L})$ and $\mathrm{p}_{\tilde{H}}(\mathcal{L})$. It is clear that $\mathrm{m}(\mathcal{L}) = \mathrm{m}_H(\mathcal{L})$ for m = n, z, p; however, for a given subspace $\tilde{H}$, which operator characterizes $\mathrm{m}_{\tilde{H}}(\mathcal{L})$ for m = n, z, p?

We will address this question in the following context. Let $S \subset H$ be a finite-dimensional subspace which is orthogonal to the kernel of $\mathcal{L}$, i.e., $S \perp \ker(\mathcal{L})$. Let $P : H \mapsto S$ be the orthogonal projection onto S, and let $Q := \mathrm{id} - P$ be the complimentary projection. From the variational characterization of eigenvalues it is easy to see that

$$\mathrm{m}_{S^\perp}(\mathcal{L}) = \mathrm{m}(\mathcal{L}_{S^\perp}),$$

for m = n, z, p, where the *constrained operator*

$$(2.1) \qquad \mathcal{L}_{S^\perp} := Q\mathcal{L} : S^\perp \mapsto S^\perp$$

is the operator induced by the bilinear form of $\mathcal{L}$ when it is constrained to act on $S^\perp$. This is different than the operator $\mathcal{L}\big|_{S^\perp}$ obtained by merely restricting the domain of $\mathcal{L}$ in that the constrained operator maps into $S^\perp$. Moreover, since Q is the identity on $S^\perp$, the constrained operator $\mathcal{L}_{S^\perp}$ is readily seen to be self-adjoint.

A key goal is to characterize the location of the point spectrum of the constrained operator $\mathcal{L}_{S^\perp}$ in terms of the point spectrum of the original operator $\mathcal{L}$. As a first step we characterize $\ker(\mathcal{L}_{S^\perp})$. We define the inverse of $\mathcal{L}$ on its range by the map $\mathcal{L}^{-1} : [\ker(\mathcal{L})]^\perp \mapsto [\ker(\mathcal{L})]^\perp$, which uniquely inverts $\mathcal{L}$ on $[\ker(\mathcal{L})]^\perp$. There are two opportunities for an element $s^\perp \in S^\perp$ to lie in $\ker(\mathcal{L}_{S^\perp})$: (a) $s^\perp \in \ker(\mathcal{L})$, or (b) $\mathcal{L}s^\perp \in S$. For the first option, since $\ker(\mathcal{L}) \subset S^\perp$, it follows that $Q\ker(\mathcal{L}) = \ker(\mathcal{L})$ so that $\ker(\mathcal{L}) \subset \ker(\mathcal{L}_{S^\perp})$. For the second option, since $\mathcal{L}$ is invertible on S, we obtain an element of the kernel precisely when we have $s \in S$ with $\mathcal{L}^{-1}s \in S^\perp$ since for such an s,

$$\mathcal{L}_{S^\perp}\mathcal{L}^{-1}s = Q\mathcal{L}\mathcal{L}^{-1}s = Qs = 0.$$

This motivates the introduction of the subspaces

$$(2.2) \qquad S_1 := \{s \in S : \mathcal{L}^{-1}s \in S^\perp\}, \quad T_1 := \{\mathcal{L}^{-1}s : s \in S_1\} \subset S^\perp,$$

as well as the orthogonal complement of S_1 relative to S, $S_1^c := S\backslash S_1$, which yields the decomposition $S = S_1 \oplus S_1^c$. Since $\mathcal{L}$ is one-to-one on S_1, we have $k := \dim(T_1) = \dim(S_1)$. To see that T_1 is precisely the difference between $\ker(\mathcal{L})$ and $\ker(\mathcal{L}_{S^\perp})$ we consider an orthonormal basis, $\{\phi_1, \ldots, \phi_k\}$ for S_1 and $\{\phi_{k+1}, \ldots, \phi_m\}$ for S_1^c, and introduce the $m \times m$ Hermitian matrix

$$(2.3) \qquad \boldsymbol{D}_{ij} = \langle \phi_i, \mathcal{L}^{-1}\phi_j \rangle.$$

We observe from the definition of S_1 that $\langle \mathcal{L}^{-1}\phi_i, \phi_j \rangle = 0$ for $i = 1, \ldots, k$ and $j = 1, \ldots, m$, since $\mathcal{L}^{-1}\phi_i$ is orthogonal to all of S. As a consequence $\boldsymbol{D}$ has a block-diagonal form

$$(2.4) \qquad \boldsymbol{D} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_1 \end{pmatrix},$$

where $\boldsymbol{D}_1 \in \mathbb{C}^{(m-k)\times(m-k)}$. Moreover, $\boldsymbol{D}_1$ is nonsingular; indeed, if it were not, then we could use the kernel of $\boldsymbol{D}_1$ to construct an element $\psi \in S_1^c$ for which $\mathcal{L}^{-1}\psi \perp S$, in contradiction to the maximality of S_1. As a consequence it follows that $\boldsymbol{c} \in \ker(\boldsymbol{D})$ and $\phi \in \ker(\mathcal{L}_{S^\perp}) \cap (\ker \mathcal{L})^\perp$ if and only if $\sum c_i s_i \in S_1$. Thus we have the decomposition

$$(2.5) \qquad \ker(\mathcal{L}_{S^\perp}) = \ker(\mathcal{L}) \oplus T_1.$$

Generically, $k = \dim(T_1) = 0$, so for an appropriate perturbation of the space S we have $\ker(\mathcal{L}_{S^\perp}) = \ker(\mathcal{L})$. After the perturbation k, the zero eigenvalues of $\mathcal{L}_{S^\perp}$ will become nonzero, either positive or negative, and consequently they must be accounted for in the stability index. The following is our main result:

Theorem 2.1 (Index Theorem). *Suppose that $S \subset \ker(\mathcal{L})^\perp$ is an m-dimensional subspace and the Hermitian matrix $\boldsymbol{D} \in \mathbb{C}^{m\times m}$ is defined as in (2.3). The difference in the negative eigenvalue count of $\mathcal{L}$ and $\mathcal{L}_{S^\perp}$ is determined through the negative and zero eigenvalue count of $\boldsymbol{D}$, i.e.,*

$$(2.6) \qquad \mathrm{n}(\mathcal{L}_{S^\perp}) = \mathrm{n}(\mathcal{L}) - \mathrm{n}(\boldsymbol{D}) - \mathrm{z}(\boldsymbol{D});$$

moreover,

$$(2.7) \qquad\qquad z(\mathcal{L}_{S^\perp}) = z(\mathcal{L}) + z(\boldsymbol{D}).$$

The constrained operator loses $n(\boldsymbol{D}) + z(\boldsymbol{D})$ *negative eigenvalues and gains* $z(\boldsymbol{D})$ *eigenvalues in its kernel.*

Remark 2.2. The Index Theorem was first proved at this level of generality in [5, Lemma 3.4]; however, the proof, while motivated by [10, Section 3], is new.

Proof. The result (2.7) follows from (2.5) and the fact that $z(\boldsymbol{D}) = \dim(S_1)$. To prove (2.6) we first consider the case when $z(\boldsymbol{D}) = 0$, so that $\boldsymbol{D}$ is nonsingular. The key step is to show that any $h \in H$ can be written as

$$(2.8) \qquad\qquad h = \mathcal{L}^{-1}s + s^\perp,$$

where $s \in S$ and $s^\perp \in S^\perp$. To determine s we write $Ph = \sum a_i \phi_i$ and $s = \sum b_i \phi_i$ and project (2.8) with P,

$$\sum a_i \phi_i = P \sum b_i \mathcal{L}^{-1} \phi_i.$$

Taking the inner product with ϕ_j for $j = 1, \ldots, m$ yields a system for the unknown $\boldsymbol{b}$ with unique solution

$$\boldsymbol{b} = \boldsymbol{D}^{-1} \boldsymbol{a}.$$

The element s so defined annihilates Ph, and we define $s^\perp = Qh - Q\mathcal{L}^{-1}s$. Consequently, for a given $h \in H$ we have (at least) two decompositions: the eigenfunction expansion and (2.8).

The two decompositions for h allow us to make the following argument. Using the decomposition (2.8) to write $\mathcal{L}h = s + \mathcal{L}s^\perp$ and exploiting the orthogonality of S and $S^\perp$ yield the identity

$$(2.9) \qquad \langle \mathcal{L}h, h \rangle = \langle s + \mathcal{L}s^\perp, \mathcal{L}^{-1}s + s^\perp \rangle = \langle s, \mathcal{L}^{-1}s \rangle + \langle s^\perp, \mathcal{L}s^\perp \rangle.$$

Set $d_1 = n_S(\mathcal{L}^{-1})$ and $d_2 = n_{S^\perp}(\mathcal{L})$, and let S_{d_1} (resp. $S_{d_2}^\perp$) be a space corresponding to $n_S(\mathcal{L}^{-1})$ (resp. $n_{S^\perp}(\mathcal{L})$), i.e.,

$$S_{d_1} = \mathrm{span}\{s_1, \ldots, s_{d_1}\}, \quad S_{d_2}^\perp = \mathrm{span}\{s_1^\perp, \ldots, s_{d_2}^\perp\}.$$

From the basis elements of these spaces then form the subspace of H given by

$$H_{d_1+d_2} = \mathrm{span}\{\mathcal{L}^{-1}s_1, \ldots, \mathcal{L}^{-1}s_{d_1}, s_1^\perp, \ldots, s_{d_2}^\perp\} \quad \left(= \mathcal{L}^{-1}S_{d_1} + S_{d_2}^\perp\right).$$

It follows from (2.9) that for all $h \in H_{d_1+d_2}$ there is an $s \in S_{d_1}$ and an $s^\perp \in S_{d_2}^\perp$ such that

$$\langle \mathcal{L}h, h \rangle = \langle s, \mathcal{L}^{-1}s \rangle + \langle s^\perp, \mathcal{L}s^\perp \rangle < 0.$$

Moreover, $\dim(H_{d_1+d_2}) = d_1 + d_2$, since if not there exists $s \in S_{d_1}$ such that $\mathcal{L}^{-1}s \in S^\perp$, which contradicts $z(\boldsymbol{D}) = 0$. We deduce that

$$n_H(\mathcal{L}) \geq n_S(\mathcal{L}^{-1}) + n_{S^\perp}(\mathcal{L}).$$

On the other hand, if $H_d = \mathrm{span}\{h_1, \ldots, h_d\}$ is a space corresponding to $n_H(\mathcal{L})$, where $d = n_H(\mathcal{L})$, then for $i = 1, \ldots, d$ we write each basis element as $h_i = \mathcal{L}^{-1}s_i + s_i^\perp$ and consequently form the spaces

$$S_d = \mathrm{span}\{s_1, \ldots, s_d\}, \quad S_d^\perp = \mathrm{span}\{s_1^\perp, \ldots, s_d^\perp\}.$$

Writing $h = \sum a_j h_j$, we use (2.9) to expand

$$\langle \mathcal{L}h, h \rangle = \sum_{i,j=1}^{d} a_i a_j \left(\langle s_i, \mathcal{L}^{-1} s_j \rangle + \langle s_i^{\perp}, \mathcal{L} s_j^{\perp} \rangle \right) = \boldsymbol{a} \cdot (\boldsymbol{A} + \boldsymbol{B}) \boldsymbol{a},$$

for appropriately defined $d \times d$ symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. Since the negative index is subadditive on symmetric matrices,

$$\mathrm{n}_H(\mathcal{L}) = \mathrm{n}(\boldsymbol{A} + \boldsymbol{B}) \leq \mathrm{n}(\boldsymbol{A}) + \mathrm{n}(\boldsymbol{B}) \leq \mathrm{n}_S(\mathcal{L}^{-1}) + \mathrm{n}_{S^{\perp}}(\mathcal{L}).$$

We can then conclude equality, i.e.,

$$\mathrm{n}_H(\mathcal{L}) = \mathrm{n}_S(\mathcal{L}^{-1}) + \mathrm{n}_{S^{\perp}}(\mathcal{L}).$$

Now, writing $s = \sum a_j \phi_j$ yields

$$\langle s, \mathcal{L}^{-1} s \rangle = \boldsymbol{a} \cdot \boldsymbol{D} \boldsymbol{a}, \quad \boldsymbol{a} = (a_1, \ldots, a_m)^{\mathrm{T}};$$

hence, $\mathrm{n}_S(\mathcal{L}^{-1}) = \mathrm{n}(\boldsymbol{D})$. In conclusion, we have

$$(2.10) \qquad \mathrm{n}(\mathcal{L}) = \mathrm{n}(\boldsymbol{D}) + \mathrm{n}_{S^{\perp}}(\mathcal{L}),$$

which establishes (2.6) in the case $\mathrm{z}(\boldsymbol{D}) = 0$.

The general case $k \equiv \mathrm{z}(\boldsymbol{D}) \geq 1$ follows by analytically perturbing the finite-dimensional space S to push the kernel of $\boldsymbol{D}$ negative. The analyticity of the eigenvalues of $\mathcal{L}$ and $\boldsymbol{D}$ under analytic perturbation of S follows from classical results; see Kato [17] for example. It is sufficient to perturb only the basis elements of S_1 as $\phi_j(\epsilon) = \phi_j + \epsilon \phi_j^1$, where $j = 1, \ldots, k$ and $0 \leq \epsilon \ll 1$. The perturbations $\{\phi_j^1\}$ are constrained to lie in $\ker(\mathcal{L})^{\perp}$. The projection operator is analytic in ϵ with expansion

$$P(\epsilon) = P(0) + \epsilon P'(0) + \mathcal{O}(\epsilon^2), \quad P'(0)f = \sum_{j=1}^{m} \left(\langle f, \phi_j \rangle \phi_j^1 + \langle f, \phi_j^1 \rangle \phi_j \right).$$

The matrix $\boldsymbol{D} = \boldsymbol{D}(\epsilon)$ given by

$$(\boldsymbol{D}(\epsilon))_{ij} = \langle \phi_i(\epsilon), \mathcal{L}^{-1} \phi_j(\epsilon) \rangle$$

will also be analytic in ϵ, and the k zero eigenvalues of $\boldsymbol{D}(0)$ will each have an asymptotic expansion $\lambda_j = \epsilon \lambda_j^1 + \mathcal{O}(\epsilon^2)$, where λ_j^1 is an eigenvalue of the Hermitian matrix $\boldsymbol{M} \in \mathbb{C}^{k \times k}$ given by

$$\boldsymbol{M}_{ij} = -\langle \phi_i, [P'(0)\mathcal{L}P(0) + P(0)\mathcal{L}P'(0)]\phi_j \rangle = -\langle \phi_i, [P'(0)\mathcal{L} + \mathcal{L}P'(0)]\phi_j \rangle.$$

The task is to choose the perturbations $\{\phi_j^1\}$ such that $\{\lambda_j^1\}$ are strictly negative. We impose the condition

$$(2.11) \qquad \{\phi_j^1\} \subset S^{\perp} \cap (\ker(\mathcal{L}))^{\perp}$$

and observe from the orthonormality of the set $\{\phi_j\}$ that

$$P'(0)\phi_i = \sum_{j=1}^{m} \left(\langle \phi_i, \phi_j \rangle \phi_j^1 + \langle \phi_i, \phi_j^1 \rangle \phi_j \right) = \phi_i^1.$$

In particular, since $P'(0)$ is self-adjoint we have the expression

$$\boldsymbol{M}_{ij} = -\langle \phi_i, (P'(0)\mathcal{L} + \mathcal{L}P'(0))\phi_j \rangle = -2\langle \phi_i^1, \mathcal{L}\phi_j \rangle$$

if, in addition to the constraints (2.11), we impose

$$(2.12) \qquad \langle \phi_i^1, \mathcal{L}\phi_j \rangle = \frac{1}{2}\delta_{ij},$$

for $i, j = 1, \ldots, k$ and where δ_{ij} is the Kronecker delta. Recalling the definition of the subspace T_1 given in (2.2), we may find $\{\phi_j^1\}$ from the span of T_1 which satisfy (2.12) for which $\boldsymbol{M} = -\boldsymbol{I}_{k \times k}$ and $\lambda_j^1 = -1$ for $j = 1, \ldots, k$.

To complete the proof, for $\epsilon > 0$ we have $\mathrm{z}(\boldsymbol{D}(\epsilon)) = 0$, and from (2.10) we conclude that

$$\mathrm{n}(Q(\epsilon)\mathcal{L}Q(\epsilon)) = \mathrm{n}(\mathcal{L}) - \mathrm{n}(\boldsymbol{D}(\epsilon)) \qquad (Q(\epsilon) = \mathrm{id} - P(\epsilon)).$$

However from the ϵ−continuity of the eigenvalues of $\boldsymbol{D}$ we have

$$\mathrm{n}(\boldsymbol{D}(\epsilon)) = \mathrm{n}(\boldsymbol{D}(0)) + \mathrm{z}(\boldsymbol{D}(0)).$$

That is, the kernel of $\boldsymbol{D}(0)$ generates $\mathrm{z}(\boldsymbol{D}(0))$ negative eigenvalues of $\boldsymbol{D}(\epsilon)$ for $\epsilon > 0$. The result (2.6) follows in the general case.   □

In the remainder of this section we apply the Index Theorem to locate the eigenvalues of $\mathcal{L}_{S^\perp}$ relative to the eigenvalues of $\mathcal{L}$. The quantity $\mathrm{n}(\mathcal{L} - \lambda_2\,\mathrm{id}) - \mathrm{n}(\mathcal{L} - \lambda_1\,\mathrm{id})$ counts the number of eigenvalues of $\mathcal{L}$ on the interval $[\lambda_1, \lambda_2]$, which in conjunction with the index (2.6) generates precise statements about $\sigma_p(\mathcal{L}_{S^\perp})$. Enumerate the point spectrum of $\mathcal{L}$ and $\mathcal{L}_{S^\perp}$ according to multiplicity

$$\sigma_p(\mathcal{L}) = \{\lambda_0 \leq \lambda_1 \leq \cdots\}, \quad \sigma_p(\mathcal{L}_{S^\perp}) = \{\lambda_0^\perp \leq \lambda_1^\perp \leq \cdots\}.$$

For real-valued λ define $\boldsymbol{D}(\lambda) \in \mathbb{R}^{m \times m}$ via

$$(2.13) \qquad \boldsymbol{D}_{ij}(\lambda) = \langle \phi_i, (\mathcal{L} - \lambda\,\mathrm{id})^{-1}\phi_j \rangle$$

(see [22, Lemma E.1] or [5, Lemma 3.4]). For $\lambda \in \mathbb{R}$ to the left of the essential spectra the matrix $\boldsymbol{D}(\lambda)$ is real meromorphic and self-adjoint. Note that $\boldsymbol{D}(\lambda)$ is regular at $\lambda = 0$ by assumption; however, it will generically be the case that $\boldsymbol{D}(\lambda)$ will have a nonzero (but not necessarily full rank) residue at $\lambda = \lambda_j \neq 0$. Set

$$(2.14) \qquad \boldsymbol{D}_j = \lim_{\lambda \to \lambda_j} (\lambda_j - \lambda)\boldsymbol{D}(\lambda);$$

i.e., $\boldsymbol{D}_j$ is the negative of the residue of $\boldsymbol{D}(\lambda)$ at $\lambda = \lambda_j$.

Since $\boldsymbol{D}(\lambda)$ is self-adjoint and real-meromorphic, the eigenvalues, denoted by $d_j(\lambda)$ for $j = 1, \ldots, m$, as well as the associated eigenvectors, denoted by $\boldsymbol{v}_j(\lambda)$, are also real meromorphic (e.g., see [17]). Now,

$$\boldsymbol{D}_{ij}'(\lambda) = \langle (\mathcal{L} - \lambda\,\mathrm{id})^{-1}\phi_i, (\mathcal{L} - \lambda\,\mathrm{id})^{-1}\phi_j \rangle,$$

so that for any $\boldsymbol{c} \in \mathbb{R}^m$ and any $\lambda \notin \sigma_p(\mathcal{L})$ we may introduce $\phi = \sum c_i \phi$ to find

$$\boldsymbol{c} \cdot \boldsymbol{D}'(\lambda)\boldsymbol{c} = \langle (\mathcal{L} - \lambda\,\mathrm{id})^{-1}\phi, (\mathcal{L} - \lambda\,\mathrm{id})^{-1}\phi \rangle > 0.$$

In other words, at the regular points of $\boldsymbol{D}$, the matrix $\boldsymbol{D}'(\lambda)$ is positive-definite, which implies that its eigenvalues are strictly increasing functions of λ,

$$(2.15) \qquad d_i'(\lambda) = \frac{\boldsymbol{v}_i(\lambda) \cdot \boldsymbol{D}'(\lambda)\boldsymbol{v}_i(\lambda)}{|\boldsymbol{v}_i(\lambda)|^2} > 0,$$

for the values of λ at which they are regular.

Definition 2.3. We will denote by z_j the number of eigenvalues of $\boldsymbol{D}(\lambda)$ which are regular at $\lambda = \lambda_j$ and take the value 0. We will denote by n_j the number of eigenvalues of $\boldsymbol{D}(\lambda)$ which are regular at $\lambda = \lambda_j$ and are strictly negative.

Theorem 2.4 (Eigenvalue Interlacing Theorem)**.** *The dimension of the kernel of* $\mathcal{L}_{S^\perp} - \lambda_j \, \mathrm{id}$ *is given by*

$$(2.16) \qquad \mathrm{z}\left(\mathcal{L}_{S^\perp} - \lambda_j \, \mathrm{id}\right) = \mathrm{z}\left(\mathcal{L} - \lambda_j \, \mathrm{id}\right) + \mathrm{z}_j - \mathrm{rank}(\boldsymbol{D}_j).$$

The number of eigenvalues of $\mathcal{L}_{S^\perp}$ *in the open interval* $(\lambda_j, \lambda_{j+1})$ *is given by*

$$(2.17) \qquad \mathrm{n}\left(\mathcal{L}_{S^\perp} - \lambda \, \mathrm{id}\right)\Big|_{\lambda_j^+}^{\lambda_{j+1}^-} = \mathrm{rank}(\boldsymbol{D}_j) + \mathrm{n}_j - (\mathrm{n}_{j+1} + \mathrm{z}_{j+1}).$$

Finally, the number of eigenvalues for $\mathcal{L}_{S^\perp}$ *in the interval* $[\lambda_j, \lambda_{j+1})$ *is given by*

$$(2.18) \qquad \mathrm{n}\left(\mathcal{L}_{S^\perp} - \lambda \, \mathrm{id}\right)\Big|_{\lambda_j}^{\lambda_{j+1}^-} = (\mathrm{z}_j + \mathrm{n}_j) - (\mathrm{n}_{j+1} + \mathrm{z}_{j+1}).$$

Proof. Since $(\mathcal{L} - \lambda \, \mathrm{id})|_{S^\perp} = \mathcal{L}_{S^\perp} - \lambda \, \mathrm{id}$, the quantity $\mathrm{n}(\mathcal{L}_{S^\perp} - \lambda \, \mathrm{id})$ counts the number of point spectra of $(\mathcal{L} - \lambda \, \mathrm{id})_{S^\perp}$ to the left of λ. Let $\epsilon > 0$ be sufficiently small. Applying (2.6) to $\mathcal{L} - (\lambda_j \pm \epsilon) \, \mathrm{id}$ yields

$$(2.19) \qquad \mathrm{n}\left(\mathcal{L}_{S^\perp} - \lambda \, \mathrm{id}\right)\Big|_{\lambda_j - \epsilon}^{\lambda_j + \epsilon} = \mathrm{z}(\mathcal{L} - \lambda_j \, \mathrm{id}) - (\mathrm{n} + \mathrm{z})\left(\boldsymbol{D}(\lambda)\right)\Big|_{\lambda_j - \epsilon}^{\lambda_j + \epsilon}.$$

Since the point spectra are isolated, for $\epsilon > 0$ sufficiently small we have

$$(2.20) \qquad \mathrm{n}\left(\mathcal{L}_{S^\perp} - \lambda \, \mathrm{id}\right)\Big|_{\lambda_j - \epsilon}^{\lambda_j + \epsilon} = \mathrm{z}\left(\mathcal{L}_{S^\perp} - \lambda_j \, \mathrm{id}\right).$$

Since the eigenvalues of $\boldsymbol{D}(\lambda)$ are strictly increasing in λ (see (2.15)), and since $m - \mathrm{rank}(\boldsymbol{D}_j)$ of the eigenvalues will have a removable singularity at $\lambda = \lambda_j$ (e.g., see the argument in either of [5, Lemma 3.4] or [13, p. 13], in which it is shown that $\boldsymbol{D}_j$ is positive semi-definite, but is not necessarily full rank), we find the equalities

$$(\mathrm{n} + \mathrm{z})(\boldsymbol{D}(\lambda_j - \epsilon)) = \mathrm{z}_j + \mathrm{n}_j, \quad (\mathrm{n} + \mathrm{z})(\boldsymbol{D}(\lambda_j + \epsilon)) = \mathrm{n}_j + \mathrm{rank}(\boldsymbol{D}_j)$$

(see Figure 1 for an illustrative example). Subtracting these two equalities yields

$$(2.21) \qquad (\mathrm{n} + \mathrm{z})\left(\boldsymbol{D}(\lambda)\right)\Big|_{\lambda_j - \epsilon}^{\lambda_j + \epsilon} = \mathrm{rank}(\boldsymbol{D}_j) - \mathrm{z}_j.$$

Substitution of (2.20) and (2.21) into (2.19) yields the result of (2.16). Since λ_j and λ_{j+1} are successive, distinct eigenvalues of $\mathcal{L}$,

$$\mathrm{n}\left(\mathcal{L} - (\lambda_{j+1} - \epsilon) \, \mathrm{id}\right) = \mathrm{n}\left(\mathcal{L} - (\lambda_j + \epsilon) \, \mathrm{id}\right),$$

and applying the index (2.6) yields

$$\mathrm{n}\left(\mathcal{L}_{S^\perp} - \lambda \, \mathrm{id}\right)\Big|_{\lambda_j + \epsilon}^{\lambda_{j+1} - \epsilon} = -(\mathrm{n} + \mathrm{z})\left(\boldsymbol{D}(\lambda)\right)\Big|_{\lambda_j + \epsilon}^{\lambda_{j+1} - \epsilon}.$$

However,

$$(\mathrm{n} + \mathrm{z})(\boldsymbol{D}(\lambda_j + \epsilon)) = \mathrm{n}_j + \mathrm{rank}(\boldsymbol{D}_j), \quad (\mathrm{n} + \mathrm{z})(\boldsymbol{D}(\lambda_{j+1} - \epsilon)) = \mathrm{z}_{j+1} + \mathrm{n}_{j+1},$$

and taking the limit $\epsilon \to 0^+$ we obtain (2.17).

Finally, combining the statements of (2.16) and (2.17) yields the statement of (2.18) via

$$\mathrm{z}\left(\mathcal{L}_{S^\perp} - \lambda_j \, \mathrm{id}\right) + \mathrm{n}\left(\mathcal{L}_{S^\perp} - \lambda \, \mathrm{id}\right)\Big|_{\lambda_j^+}^{\lambda_{j+1}^-} = \mathrm{n}\left(\mathcal{L}_{S^\perp} - \lambda_{j+1} \, \mathrm{id}\right) - \mathrm{n}\left(\mathcal{L}_{S^\perp} - \lambda_j \, \mathrm{id}\right). \quad \square$$

Remark 2.5. For the case $\lambda = 0$, the Eigenvalue Interlacing Theorem can be considered a generalization of (2.7), where the assumption $\mathrm{rank}(\boldsymbol{D}(0)) = 0$ was made.
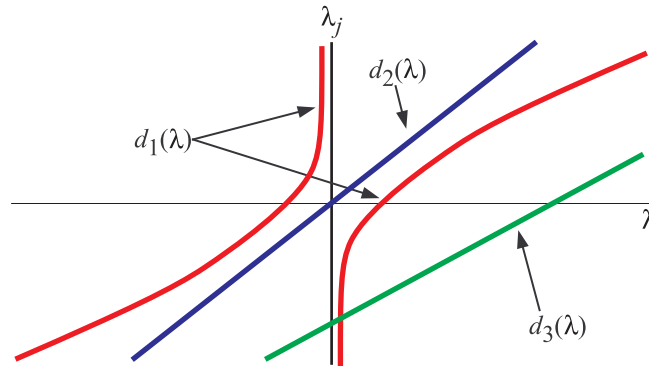
FIGURE 1. (color online only) A sketch of the eigenvalues $d_i(\lambda)$ of $\boldsymbol{D}(\lambda)$ near $\lambda = \lambda_j$ in the case that $m = 3$ with $z_j = n_j = \mathrm{rank}(\boldsymbol{D}_j) = 1$. Here we see that $(n + z)(\boldsymbol{D}(\lambda_j - \epsilon)) = 2$ and $(n + z)(\boldsymbol{D}(\lambda_j + \epsilon)) = 2$.

Remark 2.6. It has been observed in [13, p. 13] that $\mathrm{rank}(\boldsymbol{D}_j) \leq z(\mathcal{L} - \lambda_j \, \mathrm{id})$; consequently, when $m \geq 2$ it will generically be the case that $\boldsymbol{D}_j$ will not have a full rank.

The Eigenvalue Interlacing Theorem can be refined in the case $m = 1$. If $\boldsymbol{D}_j \neq 0$, then it will necessarily be true that $z_j = n_j = 0$. This then yields:

Corollary 2.7. *Suppose that $m = 1$, and further suppose that $\boldsymbol{D}_j, \boldsymbol{D}_{j+1} \neq 0$. If λ_j is an eigenvalue with geometric multiplicity ℓ for $\mathcal{L}$, then it will be an eigenvalue with multiplicity $\ell - 1$ for $\mathcal{L}_{S^\perp}$,*

$$z\left(\mathcal{L}_{S^\perp} - \lambda_j \, \mathrm{id}\right) = z(\mathcal{L} - \lambda_j \, \mathrm{id}) - 1.$$

Moreover the missing eigenvalue moves to the right, to lie in the interval $(\lambda_j, \lambda_{j+1})$, and

$$\mathrm{n}\left(\mathcal{L}_{S^\perp} - \lambda \, \mathrm{id}\right)\Big|_{\lambda_j^+}^{\lambda_{j+1}^-} = 1.$$

In particular, if all the eigenvalues of $\mathcal{L}$ are simple, and $\boldsymbol{D}_j \neq 0$ for any j such that $\lambda_j \neq 0$, then the eigenvalues of the reduced operator interlace those of $\mathcal{L}$ (see Figure 2).

3. APPLICATION: THE JONES/GRILLAKIS INSTABILITY INDEX

The dynamics of a Hamiltonian system of the form (1.1) in a neighborhood of a critical point can be related to an eigenvalue problem of the form

$$(3.1) \qquad\qquad \mathcal{J}\mathcal{L}u = \lambda u,$$

where as in Section 1 $\mathcal{J}$ is a skew-symmetric operator with respect to the inner product on the Hilbert space H with bounded inverse. The self-adjoint operator $\mathcal{L}$ arises from the second variation of the Hamiltonian at the critical point (see [13, Introduction] for further details), and as in Section 2 it will be assumed that there is a $\delta > 0$ such that $\mathcal{L}$ has a finite number of eigenvalues (counting multiplicity) for $\lambda < \delta$. The spectrum, $\sigma(\mathcal{L})$, can be related to the curvature of the energy surface,
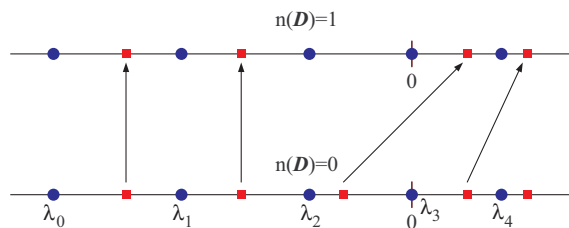
FIGURE 2. (color online only) The spectrum of $\mathcal{L}$ (blue circles) and $\mathcal{L}_{S^\perp}$ (red squares) for two choices, S_0 and S_1, of the constraint space. In both cases, $\mathrm{n}(\mathcal{L}) = 3$, $\mathrm{z}(\mathcal{L}) = 1$ and all the eigenvalues are geometrically simple with $\boldsymbol{D}_j \neq 0$ for all j except for $\lambda = 0$ (a removable singularity by assumption). In the top figure $\mathrm{n}(\boldsymbol{D}) = 1$, and in the bottom figure $\mathrm{n}(\boldsymbol{D}) = 0$ so that by the Index Theorem, $\mathrm{n}(\mathcal{L}_{S^\perp}) = \mathrm{n}(\mathcal{L}) - \mathrm{n}(\boldsymbol{D}) = 2$ or 3 in the bottom and top figures respectively. Moreover by the Eigenvalue Interlacing Theorem there is one eigenvalue of $\mathcal{L}_{S^\perp}$ in each interval $(\lambda_j, \lambda_{j+1})$ except in the interval (λ_2, λ_3) in the case $\mathrm{n}(\boldsymbol{D}) = 1$. The arrows indicate the motion of the eigenvalues of $\mathcal{L}_{S^\perp}$ as S_0 homotopies to S_1.

that is, the level set of the Hamiltonian, at the critical point. If $\sigma(\mathcal{L}) \subset \mathbb{R}^+$, with the kernel of $\mathcal{L}$ corresponding to underlying symmetries of the evolution equation (1.1), then the critical point is a minimizer of the Hamiltonian, and the energy surfaces of proximal values of the Hamiltonian are closed and remain localized near the critical point. Any orbit originating near the critical point is trapped on its corresponding energy surface and the critical point is dynamically stable. If $\sigma(\mathcal{L}) \not\subset \mathbb{R}^+$, then the critical point is not a minimizer. It may be a constrained minimizer if other conserved quantities of the system are accounted for, or it may be dynamically unstable under the flow.

If the imaginary part of $\mathcal{J}\mathcal{L}$ satisfies $\mathrm{Im}(\mathcal{J}\mathcal{L}) = 0$, then it was shown in [16] that defining $\mathcal{L}_+ := \mathcal{L}$ and $\mathcal{L}_- := -\mathcal{J}\mathcal{L}\mathcal{J}$ the eigenvalue problem (3.1) can be made equivalent to the canonical case

$$
(3.2) \qquad \mathcal{J} = \begin{pmatrix} 0 & \mathrm{id} \\ -\,\mathrm{id} & 0 \end{pmatrix}, \quad \mathcal{L} = \begin{pmatrix} \mathcal{L}_+ & 0 \\ 0 & \mathcal{L}_- \end{pmatrix}.
$$

If $\mathrm{Im}(\mathcal{J}\mathcal{L}) \neq 0$, then one must be careful when comparing the two problems; in particular, the reduction of the four-fold eigenvalue symmetry for (3.1) when $\mathrm{Im}(\mathcal{J}\mathcal{L}) = 0$ to a two-fold symmetry when $\mathrm{Im}(\mathcal{J}\mathcal{L}) \neq 0$ can break the equivalence. In the sequel we assume that (3.1) is in the canonical form of (3.2). Furthermore, it will be assumed that:

Assumption 3.1. The kernel of $\mathcal{J}\mathcal{L}$ satisfies the following conditions:
   (1) $\ker(\mathcal{L}_+) \perp \ker(\mathcal{L}_-)$,
   (2) $\dim[\mathrm{gker}(\mathcal{J}\mathcal{L})] = 2\dim[\ker(\mathcal{L})]$.
A simple condition which ensures (b) is given in [15, Lemma 3.1].

In contrast to [11, Assumption 2.1], we do not assume that $\mathcal{L}$ is invertible. However, the invertibility assumption can be recovered by considering the eigenvalue problem in an appropriate subspace. The condition (a) of Assumption 3.1 ensures

that the orthogonal projection $\Pi : H \mapsto [\ker(\mathcal{L}_+) \oplus \ker(\mathcal{L}_-)]^\perp$ is well-defined. It was shown in [15, Section 3] that for nonzero eigenvalues the system (3.1) is equivalent to

$$(3.3) \qquad -\Pi \mathcal{L}_+ \Pi u = \lambda v, \quad \Pi \mathcal{L}_- \Pi v = \lambda u.$$

The operators $\Pi \mathcal{L}_\pm \Pi$ are naturally self-adjoint and due to Assumption 3.1(b) they are nonsingular on the range of Π. Moreover, the nonzero spectrum associated with either (3.2) or (3.3) has the four-fold symmetry $\{\pm\lambda, \pm\overline{\lambda}\}$. This motivates the introduction of the operators

$$(3.4) \qquad \mathcal{R} := \Pi \mathcal{L}_+ \Pi, \quad \mathcal{S}^{-1} := \Pi \mathcal{L}_- \Pi, \quad z := -\lambda^2 \quad (-\pi/2 < \arg \lambda \le \pi/2),$$

which reduces the system (3.3) to the equivalent eigenvalue problem

$$(3.5) \qquad (\mathcal{R} - z\mathcal{S})u = 0,$$

under the mapping $z = -\lambda^2$, as illustrated in Figure 3. Eigenvalues with positive real part and nonzero imaginary part are mapped in a one-to-one fashion to eigenvalues with nonzero imaginary part, and the four-fold symmetry is reduced to the reflection symmetry $\{z, \overline{z}\}$. In particular, the system (3.3) has an unstable eigenvalue λ with $\operatorname{Re}\lambda > 0$ iff the system (3.5) has an eigenvalue z with $z < 0$ or with $\operatorname{Im} z \ne 0$. We enumerate the unstable eigenvalues, denoting the number of negative real eigenvalues of (3.5) by k_r, and the number of eigenvalues with $\operatorname{Im} z < 0$ by k_c. The subscripts r and c refer to the original system which has a real pair $\pm\lambda$ or a complex quad $\{\pm\lambda, \pm\overline{\lambda}\}$ for nonzero k_r or k_c respectively; see Figure 3.
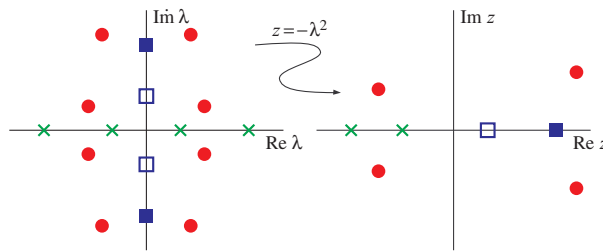

FIGURE 3. (color online only) Six sets of eigenvalues and their images under the map. The circles (red) denote two quads of complex eigenvalues under the four-fold symmetry $\{\pm\lambda, \pm\overline{\lambda}\}$ and their images, $k_\mathrm{c} = 2$. The crosses (green) denote two pairs of real eigenvalues $\{\pm\lambda\}$ and their images on the negative real axis, $k_\mathrm{r} = 2$. The boxes (blue) denote two pairs of purely imaginary eigenvalues $\{\pm\lambda\}$ and their images on the positive real axis. The filled square has a positive Krein signature, while the empty square has a negative Krein signature, so that $k_\mathrm{i}^- = 1$. If the positive and negative Krein eigenvalues were to collide under the tuning of a parameter, they would generically form a Jordon block of zero Krein signature and bifurcate into a quad of complex eigenvalues.

Of particular interest for bifurcation problems are the eigenvalues λ of (3.3) which lie on the imaginary axis; these correspond to eigenvalues $z \in \mathbb{R}^+$ of (3.5).

To each such imaginary eigenvalue we associated a positive and a negative Krein index, defined through the corresponding $z \in \mathbb{R}^+$ as

$$k_i^{\pm}(z) = \mathrm{n}(\mp \mathcal{S}_{E_z}),$$

where E_z is the eigenspace associated to z. If the tuning of a parameter leads to the collision of a negative and a positive Krein index eigenvalue, then the resulting bifurcation generically leads to unstable eigenvalues. For $z \in \mathbb{R}^+$, which are geometrically and algebraically simple with eigenfunction u, we obtain

$$(3.6) \qquad k_i^-(z) = \begin{cases} 0, & \langle u, \mathcal{S}u \rangle > 0, \\ 1, & \langle u, \mathcal{S}u \rangle < 0. \end{cases}$$

If $\langle u, \mathcal{S}u \rangle = 0$, then the eigenspace has a nontrivial Jordan block structure [7, Theorem 2.3]. The total negative Krein index of the system is defined by the sum of the individual indices

$$k_i^- = \sum_{z \in \sigma(\mathcal{S}^{-1}\mathcal{R}) \cap \mathbb{R}^+} k_i^-(z).$$

The imaginary eigenvalues of negative Krein index are the "swing producers" of instability; removing an eigenvalue of negative Krein index generically produces instability. Indeed, it has recently been shown through different proofs in [4, 11, 16] that

$$(3.7) \qquad k_r + 2k_c + 2k_i^- = \mathrm{n}(\mathcal{R}) + \mathrm{n}(\mathcal{S}).$$

Thus, spectral stability of a critical point requires

$$2k_i^- = \mathrm{n}(\mathcal{R}) + \mathrm{n}(\mathcal{S}),$$

and any decrease in the negative Krein index which is not balanced by a decrease in the negative indices of the operators $\mathcal{R}$ and $\mathcal{S}$ leads to instability. It follows immediately from (3.7) that $k_r \geq 1$ if $\mathrm{n}(\mathcal{R}) + \mathrm{n}(\mathcal{S})$ is odd (see [10] for a similar result for (3.1)). However, a significantly stronger result was first proven independently by [12] and by [8]. Jones established this result from dynamical systems arguments, while Grillakis used a detailed analysis of a contour integral involving the operator $\mathcal{K}(z) := \mathcal{R} - z\mathcal{S}$. The instability criterion (3.8) has seen substantial applications; see [14, 18, 23, 24] and the references therein.

Theorem 3.2 (Jones/Grillakis Instability Theorem). *The number of unstable eigenvalues of positive real part for the problem* (3.1) *is bounded below by*

$$(3.8) \qquad k_r \geq |\mathrm{n}(\mathcal{R}) - \mathrm{n}(\mathcal{S})|,$$

where the operators $\mathcal{R}$ and $\mathcal{S}$ are as defined in (3.4).

We furnish a new proof of this classical result via an analysis of the *Krein eigenvalues*, the eigenvalues of the Krein matrix. These are real meromorphic functions of the spectral parameter z with the properties that (a) a zero corresponds to an eigenvalue for the system (3.5), and (b) for positive z the sign of the derivative at a simple zero is the negative of the Krein signature of the eigenvalue; see [19] for a development in the scalar setting and [13] for the matrix formulation. We construct the Krein matrix by projecting off the negative space of the operator $\mathcal{S}$. More specifically, let $N(\mathcal{S})$ represent the maximal negative subspace of $\mathcal{S}$ such that

$\langle u, \mathcal{S}u \rangle < 0$ for all $u \in N(\mathcal{S})$, and set $S := N(\mathcal{S})$ so that $d := \dim(S) = \mathrm{n}(\mathcal{S})$. Let P be the projection onto $S^\perp$, and define the constrained operators

$$(3.9) \qquad \mathcal{R}_2 := P\mathcal{R}, \quad \mathcal{S}_2 := P\mathcal{S},$$

which map $S^\perp$ into itself. Since $\mathcal{S}\big|_{S^\perp}$ has no kernel it follows that $\mathcal{S}_2 > 0$ on $S^\perp$. We also define the conjugated operator

$$\widetilde{\mathcal{R}} := \mathcal{S}_2^{-1/2} \mathcal{R}_2 \mathcal{S}_2^{-1/2},$$

acting on $S^\perp$, where it satisfies $\mathrm{n}(\widetilde{\mathcal{R}}) = \mathrm{n}(\mathcal{R}_2)$ and $\mathrm{z}(\widetilde{\mathcal{R}}) = \mathrm{z}(\mathcal{R}_2)$. For $\{s_1, \ldots, s_d\}$ a basis of S and denoting the negative eigenvalues of $\mathcal{S}$ by $\lambda_1, \ldots, \lambda_d$, we introduce the $d \times d$ matrices $\boldsymbol{R}$, $\boldsymbol{S}$, and $\boldsymbol{C}(z)$ by

$$(3.10) \qquad \begin{aligned} \boldsymbol{R}_{ij} &:= \langle s_i, \mathcal{R}s_j \rangle, \quad \boldsymbol{S} := \mathrm{diag}(\lambda_1, \ldots, \lambda_d), \\ \boldsymbol{C}(z)_{ij} &:= \langle \mathcal{S}_2^{-1/2} (\widetilde{\mathcal{R}} - z\,\mathrm{id})^{-1} \mathcal{S}_2^{-1/2} P\mathcal{R}s_i, P\mathcal{R}s_j \rangle. \end{aligned}$$

The meromorphic $d \times d$ *Krein matrix* is now defined by

$$(3.11) \qquad \boldsymbol{K}(z) := \boldsymbol{R} - z\boldsymbol{S} - \boldsymbol{C}(z).$$

The matrix $\boldsymbol{K}(z)$ is symmetric; in particular, it is Hermitian for $z \in \mathbb{R}$. For $z \in \mathbb{R}$ the *Krein eigenvalues*, $\{r_i(z)\}_{i=1}^d$, are the real meromorphic eigenvalues of $\boldsymbol{K}(z)$, with potential singularities arising through $\boldsymbol{C}$ at the eigenvalues of the self-adjoint matrix $\widetilde{\mathcal{R}}$.

Lemma 3.3. *Define the matrix $\boldsymbol{D}$ by*

$$\boldsymbol{D}_{ij} := \langle s_i, \mathcal{R}^{-1}s_j \rangle.$$

Then $\mathrm{p}(\boldsymbol{K}(0)) = \mathrm{p}(\boldsymbol{D})$.

Proof. We examine the matrix $\boldsymbol{C}$ at $z = 0$,

$$\boldsymbol{C}(0)_{ij} = \langle \mathcal{S}_2^{-1/2} \widetilde{\mathcal{R}}^{-1} \mathcal{S}_2^{-1/2} P\mathcal{R}s_i, P\mathcal{R}s_j \rangle = \langle \mathcal{R}_2^{-1} P\mathcal{R}s_i, P\mathcal{R}s_j \rangle.$$

First suppose that $\boldsymbol{D}$ is nonsingular. Using the definition $\mathcal{R}_2 = P\mathcal{R}P$, we may rewrite the left entry in the inner product above as

$$(3.12) \qquad \mathcal{R}_2^{-1} P\mathcal{R}s_i = s_i + \sum_{k=1}^d a_k^i \mathcal{R}^{-1}s_k.$$

In order to enforce the condition $\mathcal{R}_2^{-1} P\mathcal{R}s_i \in S^\perp$, the vector $\boldsymbol{a}_i = (a_1^i, \ldots, a_d^i)^{\mathrm{T}}$ must be chosen so that for $m = 1, \ldots, d$,

$$0 = \langle \mathcal{R}_2^{-1} P\mathcal{R}s_i, s_m \rangle = \langle s_i, s_m \rangle + \sum_{k=1}^d a_k^i \langle \mathcal{R}^{-1}s_k, s_m \rangle,$$

or equivalently

$$\boldsymbol{0} = \boldsymbol{e}_i + \boldsymbol{D}\boldsymbol{a}_i,$$

where $\boldsymbol{e}_i \in \mathbb{R}^d$ is the ith unit vector. In other words,

$$\boldsymbol{a}_i = -\boldsymbol{D}^{-1}\boldsymbol{e}_i \quad \Longrightarrow \quad a_k^i = -(\boldsymbol{D}^{-1})_{ik},$$

so that (3.12) can be rewritten as

$$(3.13) \qquad \mathcal{R}_2^{-1} P\mathcal{R}s_i = s_i - \sum_{k=1}^d (\boldsymbol{D}^{-1})_{ik} \mathcal{R}^{-1}s_k.$$

Substituting the expression of (3.13) into that for $\boldsymbol{C}(0)_{ij}$ and using the fact that $\langle s, Pu \rangle = 0$ for any $s \in S$ yield

$$\boldsymbol{C}(0)_{ij} = \langle \mathcal{R}_2^{-1} P\mathcal{R}s_i, P\mathcal{R}s_j \rangle$$

$$= -\sum_{k=1}^{d}(\boldsymbol{D}^{-1})_{ik}\langle \mathcal{R}^{-1}s_k, P\mathcal{R}s_j \rangle$$

$$= -\sum_{k=1}^{d}(\boldsymbol{D}^{-1})_{ik}\langle \mathcal{R}^{-1}s_k, \mathcal{R}s_j \rangle + \sum_{\ell=1}^{d}\sum_{k=1}^{d}(\boldsymbol{D}^{-1})_{ik}\langle \mathcal{R}^{-1}s_k, s_\ell \rangle\langle \mathcal{R}s_j, s_\ell \rangle,$$

i.e.,

$$\boldsymbol{C}(0)_{ij} = -\sum_{k=1}^{d}(\boldsymbol{D}^{-1})_{ik}\langle s_k, s_j \rangle + \sum_{\ell=1}^{d}\sum_{k=1}^{d}(\boldsymbol{D}^{-1})_{ik}\boldsymbol{D}_{k\ell}\boldsymbol{R}_{j\ell} = -(\boldsymbol{D}^{-1})_{ij} + \boldsymbol{R}_{ij}.$$

For $\boldsymbol{D}$ nonsingular, the result now follows from the definition of the Krein matrix,

$$\boldsymbol{K}(0) = \boldsymbol{R} - \boldsymbol{C}(0) = \boldsymbol{D}^{-1}.$$

Now suppose that $z(\boldsymbol{D}) \geq 1$. We would like to characterize the pole of $\boldsymbol{C}$ at $z = 0$. From (2.5) and the Index Theorem we have $z(\mathcal{R}_2) = z(\boldsymbol{D})$, and moreover the kernel can be characterized as

$$\ker(\mathcal{R}_2) = \{\mathcal{R}^{-1}s : s \in S \text{ and } \mathcal{R}^{-1}s \in S^\perp\}.$$

Now, (3.12) requires that $P\mathcal{R}s_i \in \ker(\mathcal{R}_2)^\perp$, i.e.,

$$0 = \langle P\mathcal{R}s_i, \mathcal{R}^{-1}s \rangle = \langle \mathcal{R}s_i, \mathcal{R}^{-1}s \rangle - \sum_{k=1}^{d}\langle \mathcal{R}s_i, s_k \rangle\langle s_k, \mathcal{R}^{-1}s \rangle.$$

If $\mathcal{R}^{-1}s \in \ker(\mathcal{R}_2)$, then the above simplifies to

$$0 = \langle s_i, s \rangle.$$

Since $s \in S_1$ (see (2.2)), this yields a contradiction unless $s_i \in S_1^c$. Consequently, we can conclude that

$$\text{rank}\left(\lim_{z \to 0} z\boldsymbol{K}(z)\right) = \text{rank}\left(\lim_{z \to 0} z\boldsymbol{C}(z)\right) = z(\boldsymbol{D}).$$

Considering only $s \in S_1^c$ corresponds to projecting off of the kernel of $\boldsymbol{D}$, as in the decomposition (2.4) and reduces us to the invertible case. Since $\text{p}\left(\boldsymbol{D}\big|_{S_1^c}\right) = \text{p}(\boldsymbol{D})$ we recover the result $\text{p}(\boldsymbol{K}(0)) = \text{p}(\boldsymbol{D})$. $\qquad\square$

From (3.11) we readily see that

$$\lim_{z \to -\infty} \frac{1}{z}\boldsymbol{K}(z) = -\boldsymbol{S},$$

so that we may enumerate the Krein eigenvalues to satisfy

(3.14) $$\lim_{z \to -\infty} \frac{r_i(z)}{z} = -\lambda_i > 0.$$

In particular, each of the Krein eigenvalues is negative for $z < 0$ sufficiently large. Moreover, each zero-crossing of a Krein eigenvalue which occurs for $z < 0$ generates a negative eigenvalue of (3.5), while from the definition of the Krein matrix it is clear that its poles are a subset of the eigenvalues of $\widetilde{\mathcal{R}}$. In [13, Section 2.3] it was

shown that each pole of the Krein matrix is simple with a positive definite residue
given by

$$\text{(3.15)} \qquad \text{Res}_{z_0} = \lim_{z \to z_0} (z - z_0) K(z) \geq 0.$$

In particular, since the poles of the Krein matrix are simple, the poles of each
Krein eigenvalue are distinct and a higher rank residue of the Krein matrix occurs
only when poles of distinct Krein eigenvalues coincide. Moreover, it was shown
that the rank of the residue is precisely the geometric multiplicity of the eigenvalue
of $\widetilde{\mathcal{R}}$ minus the number of eigenvalues whose eigenfunctions lie in $S^\perp$. Any such
eigenvalue of $\widetilde{\mathcal{R}}$ is also an eigenvalue of (3.5) and hence contributes to the count of
k_r. As we aim for a lower bound on k_r, without loss of generality we may assume
that all residues of K have maximal rank.

From Lemma 3.3, precisely

$$\text{(3.16)} \qquad \text{p}(\boldsymbol{K}(0)) = \text{p}(\boldsymbol{D}) = \text{n}(\mathcal{S}) - \text{n}(\boldsymbol{D}) - \text{z}(\boldsymbol{D})$$

of the Krein eigenvalues are positive for $z = 0$. Denoting the poles of the Krein
eigenvalues $r_i(z)$ by $\{\mu_i^0 < \mu_i^1 < \mu_i^2 < \cdots \}$, then from (3.15) for each $\ell \in \mathbb{N}_0$ we
have the limits

$$\text{(3.17)} \qquad \lim_{z \to (\mu_i^\ell)^\pm} r_i(z) = \pm\infty$$

(see Figure 4). In particular, the ith Krein eigenvalue tends to $-\infty$ as it approaches
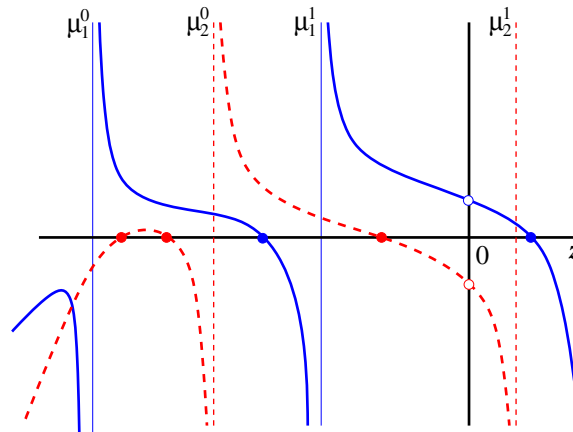

FIGURE 4. (color online only) A sketch of the graph of the Krein
eigenvalues for $z \in \mathbb{R}$ when $d = 2$. The solid (blue) curve corre-
sponds to the graph of $r_1(z)$, and the dashed (red) curve is the
graph of $r_2(z)$. The zero crossings are indicated by solid balls, the
location of the Krein eigenvalue at $z = 0$ by open balls. This car-
toon illustrates the situation when the poles of the Krein matrix are
simple and have nonzero residue, and the poles of the Krein eigen-
values are interlaced, i.e., $\mu_0^1 < \mu_2^0 < \mu_1^1 < \mu_2^1$. The three negative
poles of the Krein eigenvalues and one positive Krein eigenvalue at
$z = 0$ require at least two zero-crossings for $z < 0$; see (3.20).

one of its poles, μ_i^ℓ, from the left, and takes a positive value to the right of the pole, so that successive poles of the same Krein eigenvalue generate at least one zero-crossing. By the Index Theorem and the definition (3.9) of $\mathcal{R}_2$ we see that the number of poles is precisely

$$(3.18) \qquad \mathrm{n}(\mathcal{R}_2) = \mathrm{n}\big(\mathcal{R}_{N(S)^\perp}\big) = \mathrm{n}(\mathcal{R}) - \mathrm{n}(\boldsymbol{D}) - \mathrm{z}(\boldsymbol{D}).$$

If there are no zero-crossings of the Krein eigenvalues on the set $z < 0$, then it must be that the number of poles is precisely equal to the number of positive Krein eigenvalues at $z = 0$. Each integer increment in the discrepancy between the number of positive eigenvalues at $z = 0$ and the number of poles must correspond to at least one additional zero-crossing. In light of the equalities (3.16) and (3.18) we may rephrase this argument in terms of the lower bound

$$(3.19) \qquad k_{\mathrm{r}} \geq [\mathrm{p}(\boldsymbol{D}) - \mathrm{n}(\mathcal{R}_2)] = \mathrm{n}(\mathcal{S}) - \mathrm{n}(\mathcal{R}).$$

On the other hand, if the number of poles exceeds $\mathrm{p}(\boldsymbol{D})$, then either the excess pole is a second pole of one of the Krein eigenvalues, see the solid line of Figure 4, or one of the positive Krein eigenvalues must have crossed zero to become negative at $z = 0$; see the dashed line of Figure 4. In either case, the excess poles generate at least one zero-crossing a piece, so that

$$(3.20) \qquad k_{\mathrm{r}} \geq [\mathrm{n}(\mathcal{R}_2) - \mathrm{p}(\boldsymbol{D})] = \mathrm{n}(\mathcal{R}) - \mathrm{n}(\mathcal{S}).$$

Taken together, the two bounds (3.19) and (3.20) yield the Jones/Grillakis instability condition (3.8).

Remark 3.4. The asymptotic property of (3.17) guarantees that $r_i(z)$ will have an odd number of zeros in the interval $(\mu_i^\ell, \mu_i^{\ell+1})$. The lower bound of (3.19) is achieved by assuming that the odd number is actually one. Furthermore, the property of (3.14) ensures that $r_i(z)$ will have an even number of zeros in the interval $(-\infty, \mu_i^0)$. Again, the lower bound is achieved by assuming that the even number is zero. See [20] for an example in which both stability and instability are obtained for $\mathrm{n}(\mathcal{R}) = \mathrm{n}(\mathcal{S})$.

Acknowledgments

References

[1] T. Benjamin, The stability of solitary waves, *Proc. Roy. Soc. (London) Ser. A*, 328, 153–183, 1972. MR0338584 (49:3348)

[2] J. Bona, P. Souganidis, and W. Strauss, Stability and instability of solitary waves of Korteweg-de Vries type, *Proc. R. Soc. London A*, 411, 395–412, 1987. MR897729 (88m:35128)

[3] J. Boussinesq, Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond, *J. Math. Pures Appl.*, 17, 55–108, 1872.

[4] M. Chugunova and D. Pelinovsky, Count of eigenvalues in the generalized eigenvalue problem. *J. Math. Phys.*, 51 (5), 052901, 2010. MR2666968

[5] S. Cuccagna, D. Pelinovsky, and V. Vougalter, Spectra of positive and negative energies in the linearized NLS problem, *Comm. Pure Appl. Math.*, 58, 1–29, 2005. MR2094265 (2005k:35374)

[6] B. Deconinck and T. Kapitula, On the spectral and orbital stability of spatially periodic stationary solutions of generalized Korteweg-de Vries equations, preprint.

[7] M. Grillakis, Analysis of the linearization around a critical point of an infinite-dimensional Hamiltonian system, *Comm. Pure Appl. Math.*, 43, 299–333, 1990. MR1040143 (91d:58231)

[8] M. Grillakis, Linearized instability for nonlinear Schrödinger and Klein-Gordon equations, *Comm. Pure Appl. Math.*, 46, 747–774, 1988. MR948770 (89m:35192)

[9] M. Grillakis, J. Shatah, and W. Strauss, Stability theory of solitary waves in the presence of symmetry, I, *Journal of Functional Analysis*, 74, 160–197, 1987. MR901236 (88g:35169)

[10] M. Grillakis, J. Shatah, and W. Strauss, Stability theory of solitary waves in the presence of symmetry, II, *Journal of Functional Analysis*, 94, 308–348, 1990. MR1081647 (92a:35135)

[11] M. Hărăguş and T. Kapitula, On the spectra of periodic waves for infinite-dimensional Hamiltonian systems, *Physica D*, 237 (20), 2649–2671, 2008. MR2514124 (2010c:37167)

[12] C.K.R.T. Jones, Instability of standing waves for nonlinear Schrödinger-type equations, *Ergod. Th. & Dynam. Sys.*, 8*, 119–138, 1988. MR967634 (90d:35267)

[13] T. Kapitula, The Krein signature, Krein eigenvalues, and the Krein Oscillation Theorem, Indiana U. Math. J., 2010 (to appear).

[14] T. Kapitula, P. Kevrekidis, and B. Malomed, Stability of multiple pulses in discrete systems, *Phys. Rev. E*, 63 (036604), 2001.

[15] T. Kapitula, P. Kevrekidis, and B. Sandstede, Counting eigenvalues via the Krein signature in infinite-dimensional Hamiltonian systems, *Physica D*, 195 (3–4), 263–282, 2004. MR2089513 (2006k:37167)

[16] T. Kapitula, P. Kevrekidis, and B. Sandstede, Addendum: Counting eigenvalues via the Krein signature in infinite-dimensional Hamiltonian systems, *Physica D*, 201 (1–2), 199–201, 2005. MR2118638 (2006k:37168)

[17] T. Kato, *Perturbation Theory for Linear Operators.* Reprint of the 1980 edition. Springer-Verlag, Berlin, 1995. MR1335452 (96a:47025)

[18] S. Lafortune and J. Lega, Spectral stability of local deformations of an elastic rod: Hamiltonian formalism, *SIAM J. Math. Anal.*, 36 (6), 1726–1741, 2005. MR2178219 (2006g:35188)

[19] Y. Li and K. Promislow, Structural stability of non-ground state traveling waves of coupled nonlinear Schrödinger equations, *Physica D*, 124 (1–3), 137–165, 1998. MR1662538 (99j:35204)

[20] Y. Li and K. Promislow, The mechanism of the polarization mode instability in birefringent fiber optics, *SIAM J. Math. Anal.*, 31 (6), 1351–1373, 2000. MR1766556 (2001f:35375)

[21] D. Pelinovsky, Inertia law for spectral stability of solitary waves in coupled nonlinear Schrödinger equations, *Proc. Royal Soc. London A*, 461, 783–812, 2005. MR2121936 (2006d:37142)

[22] M. Weinstein, Modulational stability of ground states of nonlinear Schrödinger equations, *SIAM J. Math. Anal.*, 16 (3), 472–491, 1985. MR783974 (86i:35130)

[23] A. Yew, Stability analysis of multipulses in nonlinearly coupled Schrödinger equations, *Indiana U. Math. J.*, 49 (3), 1079–1124, 2000. MR1803222 (2001m:35033)

[24] A. Yew, B. Sandstede, and C.K.R.T. Jones, Instability of multiple pulses in coupled nonlinear Schrödinger equations, *Phys. Rev. E*, 61 (5), 5886–5892, 2000. MR1791733 (2001h:78041)

Department of Mathematics and Statistics, Calvin College, Grand Rapids, Michigan 49546

E-mail address: `tmk5@calvin.edu`

Department of Mathematics, Michigan State University, East Lansing, Michigan 48824

E-mail address: `kpromisl@math.msu.edu`