A CANONICAL FORM FOR PAIRS CONSISTING OF A HERMITIAN FORM AND A SELF-ADJOINT ANTILINEAR OPERATOR

DAVID SYKES AND IGOR ZELENKO

ABSTRACT. Motivated by a problem in local differential geometry of Cauchy–Riemann (CR) structures of hypersurface type, we find a canonical form for pairs consisting of a nondegenerate Hermitian form and a self-adjoint antilinear operator, or, equivalently, consisting of a nondegenerate Hermitian form and a symmetric bilinear form. This generalizes the only previously known results on simultaneous normalization of such pairs, namely, the results of [2] on simultaneous diagonalization of these pairs in the case where the Hermitian form is positive definite and of [11], where a criterion for simultaneous diagonalization is given.

1. INTRODUCTION

In the present paper we find canonical forms for pairs consisting of a nondegenerate Hermitian form $\ell$ on a complex $n$-dimensional vector space $W$ and an antilinear operator $A : W \rightarrow W$ that is self-adjoint with respect to the form $\ell$. By a canonical form, as usual, we mean a specified choice of matrices representing elements of any such pair, chosen from among matrix representations in all possible bases of $W$. Our main result is formulated in Theorem 2.2. Recall that a map $A : W \rightarrow W$ is called an antilinear operator if

$$A(\lambda v + w) = \bar{\lambda} A(v) + A(w) \quad \forall v, w \in W, \lambda \in \mathbb{C},$$

and an antilinear operator $A$ is called self-adjoint with respect to the form $\ell$ or, shortly, $\ell$-self-adjoint if

$$\ell(Av, w) = \ell(Aw, v) \quad \forall v, w \in W.$$  \hfill (1.1)

Our original motivation for this work comes from the local differential geometry of certain Cauchy–Riemann (CR) structures, more precisely of real hypersurfaces of a complex space having uniformly degenerate Levi form with one dimensional kernel. As was shown recently in [14], the basic invariant of such structures at a point is given exactly by a pair of the algebraic objects under consideration. For more details see section 3 below.

Our main result, Theorem 2.2, also gives canonical forms for pairs consisting of a nondegenerate Hermitian form and a symmetric bilinear form because the set of these pairs is in bijective correspondence with the one we originally considered. Indeed, to the pair $(\ell, A)$ we can assign the pair $(\ell, \ell')$, where

$$\ell'(v, w) := \ell(w, Av)$$

is a symmetric bilinear form by (1.1). From the nondegeneracy of $\ell$ it follows that the assignment of $(\ell, A)$ to $(\ell, \ell')$ defines the bijection between the two sets of pairs under consideration.

Surprisingly, when we encountered the necessity of finding the canonical forms for pairs $(\ell, A)$ in the course of our study in CR geometry, we were not able to find the desired results in the literature.

\textit{Date: September 23, 2019.}

\textit{2010 Mathematics Subject Classification.} 15A21, 15A24, 15B05, 15A22, 32V40.

\textit{Key words and phrases.} antilinear operators, indefinite Hermitian forms, canonical forms, uniformly Levi degenerate CR structures, pencils.

I. Zelenko is supported by Simons Foundation Collaboration Grant for Mathematicians 524213.
The only results in this direction that we found are those addressing the problem of simultaneous diagonalization [2, 11] and those giving canonical forms for a single antilinear operator [7, 8, 10] and, more generally, for a single semi-linear operator [1, 13] or for a square matrix under \( \varphi \)-equivalence [9]. In [2, Theorem 7], it shown that \( \ell \) and \( A \) can be simultaneously diagonalized if \( \ell \) is positive definite, and, in [11, Theorem 2.1], the pairs \((\ell, A)\) admitting a simultaneous diagonalization are classified. Perhaps, the main difficulty here is that the matrix representations for a Hermitian form and an antilinear operator transform differently under a change of the basis (see formulas (2.2) below). It also cannot be reduced to the study of canonical forms of pairs of other objects, wherein the matrix representations of each component of the new pairs transforms in the same way under a basis change. An example of the latter reduction is the set of pairs consisting of a nondegenerate Hermitian form \( \ell \) and an \( \ell \)-self-adjoint linear operator that was treated in [6, Theorem 5.1.1] where a canonical form for such pairs is given, which we will refer to as the Gohberg–Lancaster–Rodman form. Although the matrix representations of each component in such pairs transform differently under a basis change, using a process similar to the one in the previous paragraph, we can obtain a bijective correspondence between the set of such pairs and the set of pairs of Hermitian forms (i.e., a pair of the same type of objects), one of which is nondegenerate. In our case, however, such a reduction is not possible and the problem of finding canonical forms cannot be totally reduced to the study of certain classes of matrix pencils, as was classically done using Weierstrass–Kronecker normal forms for matrix pencils (see, for example, [5] and [17]).

To prove Theorem 2.2, we develop in section 4 a geometric version of the construction of the canonical form for a single antilinear operator of [8] (which was formulated in [10, Theorem 3.1], proved in [8], and stated for completeness in Remark 2.3 below) and combine it with a simultaneous normalization of the Hermitian form, which is comparable in certain respects to the method of [6, subsection 5.3] for obtaining the Gohberg–Lancaster–Rodman form, mentioned in the previous paragraph. By a geometric version we are referring to the study of flags of subspaces analogous to the generalized eigenspaces in the standard theory of linear operators as opposed to the algebraic version in [1, 7, 13] based on the theory of invariant factors and manipulations with matrices as in [8, 9]. Our Theorem 2.2 is related to the Hong–Horn canonical form of [10, Theorem 3.1] for a single antilinear operator in the same way that the Gohberg–Lancaster–Rodman form in [6, Theorem 5.1.1] is related to the classical Jordan normal form for linear operators.

In section 5, for completeness we sketch an alternative approach to the considered problem that leads to an equivalent canonical form, Theorem 5.1. This approach was in fact our original one before we found the more natural and apparently more simple approach leading to Theorem 2.2. The idea in an alternative approach is as follows: Since \( A^2 \) is an \( \ell \)-self-adjoint linear operator whenever \( A \) is an \( \ell \)-self-adjoint antilinear operator, one can first bring the pair \((\ell, A^2)\) to the Gohberg–Lancaster–Rodman form and then find a canonical form for \( A \) with minimal changes in the form of \( \ell \). This requires solving a certain nonlinear matrix equation, which turned out to be feasible.

2. The Canonical Form

As in the introduction, \( \ell \) denotes a nondegenerate Hermitian form and \( A \) denotes an antilinear operator on an \( n \)-dimensional complex space \( W \). Unless otherwise stated, \( A \) is assumed to be \( \ell \)-self-adjoint (see (1.1) for the definition).

Choosing a basis \( \{e_1, \ldots, e_n\} \) of \( W \), one can represent the form \( \ell \) and the antilinear operator \( A \) by \( n \times n \) matrices \( H = (H_{i,j}) \) and \( C = (C_{i,j}) \) via a standard construction, requiring, for all \( i, j \in \{1, \ldots, n\} \), that

\[
H_{i,j} = \ell(e_j, e_i) \quad \text{and} \quad A(e_i) = \sum_{k=1}^{n} C_{k,i} e_k.
\]
The conditions that $\ell$ is a nondegenerate Hermitian form and $A$ is an $\ell$-self-adjoint antilinear operator are equivalent to

\begin{equation}
H^* = H, \quad \det H \neq 0, \quad \text{and} \quad (HC)^T = HC,
\end{equation}

respectively.

If one chooses another basis $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$, letting $\tilde{H}$ and $\tilde{C}$ be the matrices representing the form $\ell$ and the operator $A$ in this new basis and letting $M = (M_{i,j})$ be the transition matrix from the new basis to the old one, (i.e., $e_j = \sum_{i=1}^{n} M_{i,j} \tilde{e}_i$) then

\begin{equation}
\tilde{H} = (M^{-1})^* H M^{-1} \quad \text{and} \quad \tilde{C} = M CM^{-1}.
\end{equation}

Our goal is to find a basis in which the matrix representation of the form $\ell$ and operator $A$ has a particularly simple form. In other words, if we define an action of the matrix group $GL_n(\mathbb{C})$ on the pairs $(H, C)$ of $n \times n$ matrices satisfying (2.1) by the mapping

\begin{equation}
(M, (H, C)) \mapsto \left( (M^{-1})^* H M^{-1}, M CM^{-1} \right), \quad M \in GL_n(\mathbb{C}),
\end{equation}

then our goal is to choose a representative in each orbit of this action in a canonical way. This canonical representative is usually called the \textit{canonical} or \textit{normal} form of the pair $(\ell, A)$.

We let $T_k$ be the $k \times k$ matrix whose $(i, j)$ entry is 1 if $j - i = 1$ and zero otherwise, let $S_k$ be the $k \times k$ matrix whose $(i, j)$ entry is 1 if $j + i = k + 1$ and zero otherwise, let $I_k$ be the rank $k$ identity matrix, and let $J_{\lambda,k} = \lambda I_k + T_k$ be the standard $k \times k$ Jordan block corresponding to the eigenvalue $\lambda$.

To succinctly define new matrices constructed from others, we write

\[ M_1 \oplus M_2 \oplus \cdots \oplus M_k = \bigoplus_{i=1}^{k} M_i \]

to denote the block diagonal matrix whose diagonal entries are the matrices $M_1, \ldots, M_k$. For $\lambda \in \mathbb{C}$, we define the $k \times k$ or $2k \times 2k$ matrix $C_{\lambda,k}$ by

\[ C_{\lambda,k} := \begin{cases} J_{\lambda,k} & \text{if } \lambda \in \mathbb{R} \\ \begin{pmatrix} 0 & J_{\lambda^2,k} \\ I_k & 0 \end{pmatrix} & \text{otherwise}, \end{cases} \]

where 0 denotes a matrix of appropriate size with zero in all entries. We define corresponding matrices $H_{\lambda,k}$ by

\[ H_{\lambda,k} := \begin{cases} S_k & \text{if } \lambda \in \mathbb{R} \\ S_{2k} & \text{otherwise}. \end{cases} \]

For a nonnegative integer $k$, we define

\begin{equation}
W^{(k)}_{\lambda} := \text{span}_{\mathbb{C}} \left\{ v \in W : (A^2 - \lambda^2 I)^k v = 0 \text{ or } \left(A^2 - \overline{\lambda^2} I\right)^k v = 0 \right\}.
\end{equation}

Since $A^2$ is linear, we can enumerate its eigenvalues, letting $\lambda_1^2, \ldots, \lambda_{\mu}^2$ be the real eigenvalues of $A^2$ and $\lambda_{\mu+1}^2, \ldots, \lambda_{\gamma}^2$ be the distinct eigenvalues of $A^2$ with positive imaginary part. In the canonical forms below, we assume that each $\lambda_i$ is the principle square root of $\lambda_i^2$. Since the linear operator $A^2$ is $\ell$-self-adjoint, it is easy to show (see, for example, [6, Theorem 4.2.4]) that the space $W$ can be decomposed into pairwise-$\ell$-orthogonal $A^2$-invariant subspaces

\begin{equation}
W = W^{(n)}_{\lambda_1} \oplus W^{(n)}_{\lambda_2} \oplus \cdots \oplus W^{(n)}_{\lambda_{\gamma}}.
\end{equation}
Remark 2.1. In [6], the authors refine this decomposition of $W$, obtaining a canonical form for $(\ell, A^2)$. For an $\ell$-self-adjoint linear operator $B$, their theorem, [6, Theorem 5.1.1], states that the domain of $B$ can be decomposed into $B$-invariant, pairwise $\ell$-orthogonal subspaces such that there exists a basis with respect to which the restrictions of $\ell$ and $B$ to the decomposition’s component subspaces are represented by matrices of the form $\pm S_k$ and $J_{n,k}$ if $\eta \in \mathbb{R}$ or $\pm S_k$ and $J_{n,k} \oplus J_{n,k}$ if $\eta \not\in \mathbb{R}$ (this gives a canonical form for $(\ell, A^2)$).

Note that $W_{\lambda_i}^{(n)}$ is also $A$-invariant. Indeed, if $v \in W_{\lambda_i}^{(n)}$ and $(A^2 - \lambda^2 I)^k v = 0$, then

$$(A^2 - \overline{\lambda}^2 I)^n (Av) = A(A^2 - \lambda^2 I)^n v = 0,$$

which shows that $Av \in W_{\lambda_i}^{(n)}$. Similarly, if $v \in W_{\lambda_i}^{(n)}$ and $(A^2 - \lambda^2 I)^k v = 0$, then

$$(A^2 - \lambda^2 I)^n (Av) = A(A^2 - \overline{\lambda}^2 I)^n v = 0,$$

which shows that $Av \in W_{\lambda_i}^{(n)}$. This completes the proof of $A$-invariance of $W_{\lambda_i}^{(n)}$.

Accordingly, we can normalize $\ell$ and $A$ on the spaces $W_{\lambda_i}$ separately to obtain a general canonical form.

Theorem 2.2. The domain of an $\ell$-self-adjoint antilinear operator $A$ can be decomposed into $A$-invariant, pairwise $\ell$-orthogonal subspaces such that there exists a basis with respect to which the restrictions of $\ell$ and $A$ to the decomposition’s component subspaces are represented by matrices of the form $\pm H_{\lambda,k}$ and $C_{\lambda,k}$ where $\lambda \in \{\lambda_1, \lambda_2, \ldots, \lambda_\gamma\}$ and $k \in \mathbb{N}$. The corresponding block diagonal matrices representing $\ell$ and $A$ are unique up to a permutation of the blocks on the diagonal.

Proof. Since the decomposition in (2.4) is pairwise $\ell$-orthogonal and $A$-invariant, the result is a corollary of Propositions 4.10, 4.11, 4.12, and 4.13. \qed

Remark 2.3. In [10, Theorem 3.1], the authors show that an antilinear operator $A$ can be represented by a matrix in the form of the matrix given in Theorem 2.2 representing the antilinear operator, that is, the domain of $A$ can be decomposed into $A$-invariant subspaces on which $A$ is represented by $C_{\lambda,k}$ where $\lambda \in \{\lambda_1, \lambda_2, \ldots, \lambda_\gamma\}$ and $k \in \mathbb{N}$ (note, this is achieved without the assumption that $A$ is $\ell$-self-adjoint for some Hermitian form $\ell$).

A canonical form for a nonsingular antilinear operator is fully determined by the Jordan matrix representing its square, and we have a similar relationship between Theorem 2.2 and the Gohberg–Lancaster–Rodman form, recorded in the following lemma.

Lemma 2.4. If $A$ is nonsingular then the canonical form for $(\ell, A)$ given in Theorem 2.2 is determined by the Gohberg–Lancaster–Rodman form for $(\ell, A^2)$.

3. Relation to CR Geometry

In this section we demonstrate how the considered pairs of algebraic objects appear naturally in the study of a certain class of CR manifolds of hypersurface type. CR manifolds of hypersurface type are real hypersurfaces in a complex space $\mathbb{C}^{n+1}$ and originally were introduced in order to study the biholomorphic equivalence between domains in $\mathbb{C}^{n+1}$ via their boundaries (see, for example, the monograph [12]). The complex structure of $\mathbb{C}^{n+1}$ induces additional nontrivial structures on a hypersurface $M$. Namely, for every $x \in M$, let $D_x$ be the maximal complex subspace of the tangent space $T_x M$, $D_x = T_x M \cap iT_x M$, where $iT_x M$ is the real subspace in $\mathbb{R}^{2n+2} (\cong \mathbb{C}^{n+1})$ obtained from $T_x M$ via the multiplication by $i$. The collection of hyperplanes $D = \{D_x\}_{x \in M}$ defines a corank 1 subbundle of $TM$, that is, a corank 1 distribution on $M$.

By construction, the multiplication by $i$ restricted to $D(x)$ defines and endomorphism of $D(x)$ that will be denoted by $J_x$. By construction, $J_x^2 = I_x$, where $I_x$ is the identity operator on $D_x$. The operator $J_x$ extends linearly to $\mathbb{C}D_x = D \otimes \mathbb{C} \subset CT_x M = T_x M \otimes \mathbb{C}$, and, since $J_x^2 = -I$, $\mathbb{C}D$ splits
into a direct sum of the $i$-eigenspace and $-i$-eigenspace of $J_x$, denoted by $E_x$ and $\overline{E}_x$, respectively. The collections $E = \{ E_x \}_{x \in E}$ and $\overline{E} = \{ \overline{E}_x \}_{x \in E}$ define subbundles of the complexified tangent bundle $\mathbb{C}T_x M$ and both of these subbundles are involutive, that is, $[E,E] \subset E$ and $[\overline{E},\overline{E}] \subset \overline{E}$, where, for example, by $[E,E]_x$ we mean the linear span of Lie brackets, evaluated at $x$, of any two sections of the bundle $E$. This involutivity comes from the fact that if $(z_1, \ldots, z_{n+1})$ are standard coordinates in $\mathbb{C}^{n+1}$ then at every point

$$E = \mathbb{C}D \cap \text{span} \left( \left\{ \frac{\partial}{\partial z_i} \right\}_{i=1}^{n+1} \right), \quad \overline{E} = \mathbb{C}D \cap \text{span} \left( \left\{ \frac{\partial}{\partial \overline{z}_i} \right\}_{i=1}^{n+1} \right).$$

Keeping all of this in mind, an abstract CR structure of hypersurface type is a triple $(M, D, J)$, where $M$ is an odd dimensional real manifold, $D$ is a corank 1 distribution, and $J : D \to D$ is an operator that preserves each fiber of $D$ (i.e., $JD_x = D_x$ $\forall x \in M$) such that $J$ is linear on each fiber, $J^2 = -I$, and the corresponding subbundles $E$ and $\overline{E}$ of the complexified bundle $\mathbb{C}D$ are involutive.

Now, given a fiber bundle $\pi : P \to M$, let $\Gamma(P)$ be the set of smooth sections of $P$. For $x \in M$, the Hermitian form $\mathcal{L}_x : E_x \to (\mathbb{C}TM)_p/(\mathbb{C}D)_p \cong \mathbb{C}$ known as the Levi form of $(M, D, J)$ is defined, up to a real multiple, as

$$\mathcal{L}_x(X_x, Y_x) := \frac{1}{2i} \left[ X, \overline{Y} \right]_x \quad \forall X, Y \in \Gamma(E).$$

The kernel $K_x$ of this Hermitian form is called the Levi kernel of the CR structure at the point $x$. The CR structure $(M, D, J)$ is called Levi-nondegenerate if $K_x = 0$ at every point. The local differential geometry of Levi-nondegenerate CR structures is well understood (see [3, 4, 16]). In recent years, interest arose in the uniformly degenerate structures, that is, when $K_x \neq 0$ for every $x$; see [14] for the list of references.

Assume now that the Levi kernel $K_x$ is one-dimensional for every $x$. First, the degenerate Hermitian form $\mathcal{L}_x$ on $E_x$ factors through $E_x/K_x$, defining the nondegenerate Hermitian form $\ell_x$ on $E_x/K_x$, that is, $\ell_x$ is well defined by $\ell_x(\pi v, \pi w) := \mathcal{L}_x(v, w)$ with $\pi : H_x \to H_x/k_x$ denoting the canonical projection projection. Second, for $x \in M$ and $v \in K_x$, we define the antilinear operator $A_x : H_x/K_x \to H_x/K_x$ by choosing $V \in \Gamma(K_x)$ such that $V(x) = v$ and

$$A_x(Y) = \left[ V, \overline{Y} \right]_x \mod K_x \oplus \overline{K}_x \quad \forall Y \in \Gamma(H_x/K_x).$$

The antilinear operator $A_x$ is defined, up to a complex multiple and is $\ell_x$-self-adjoint (see [14]).

It turns out that a pair $(\mathbb{R}\ell_x, \mathbb{C}A_x)$ is a basic invariant of the CR structure under consideration at a point $x$ and our Theorem 2.2 gives the classification of pairs $(\ell_x, A_x)$ and therefore of the these basic invariants. In [14], the structure of an absolute parallelism (i.e., a canonical frame/coframe in a certain bundle over $M$) and maximally symmetric models were found in the particular case where $A_x^3$ is a scalar multiple of $A_x$ at every point. The reason for this somewhat weird condition is that only in this case was it possible to apply a certain version of the machinery of the prolongation of filtered structures for the construction of an absolute parallelism. In many cases where $A_x^3$ is not a scalar multiple of $A_x$ it is not even clear if there exist CR-structures on which its symmetry group acts transitively, and the existing differential geometric methods only offer hope to obtain upper bounds for the dimension of the symmetry group of the most symmetric models with pairs $(\mathbb{R}\ell_x, \mathbb{C}A_x)$ lying in the same prescribed orbit under the natural $GL$-action for every $x$. In any case, we expect that the classification given by our Theorem 2.2 will be useful for finding an upper bound for dimensions of symmetry groups and other questions related to CR structures with one-dimensional Levi kernel for which $A_x^3$ is not a scalar multiple of $A_x$. 
4. Normal Forms for Restrictions to Generalized Eigenspaces

In this section we obtain a canonical form for the restrictions of $\ell$ and $A$ to the spaces $W_\lambda^{(n)}$, and these results can be taken together to obtain the canonical form in Theorem 2.2. The approach we employ varies depending on the eigenvalue $\lambda^2$ of $A^2$, so this section is structured with subsections, each dedicated to a case where $\lambda^2$ belongs to a different family. We repeatedly use the following lemma, which is completely analogous to a standard property of linear self-adjoint operators.

**Lemma 4.1.** If $V \subset \mathbb{C}^n$ is an $A$-invariant subspace on which $\ell$ is nondegenerate then the $\ell$-orthogonal complement $V^{\perp_\ell}$ of $V$ is also $A$-invariant.

**Proof.** Since $V$ is $A$-invariant, for any $w \in V$, we have that $Aw \in V$, which implies that, for $v \in V^{\perp_\ell}$, we have $\ell(Aw, v) = 0$. Therefore, since $A$ is $\ell$-self-adjoint, for $v \in V^{\perp_\ell}$ and $w \in V$, we have $\ell(Aw, v) = \ell(Aw, v) = 0$, which implies that $Av \in V^{\perp_\ell}$. \qed

§4.1. Treating Generalized Eigenspaces with Positive Eigenvalues. Throughout this subsection we assume $\lambda^2 > 0$, and this subsection’s main result is Proposition 4.10.

For this special case with $\lambda^2 > 0$, we define three additional filtrations of $W_\lambda^{(n)}$. Namely,

\begin{equation}
W_\lambda^{(k)} := \{ x \in W : (A \mp \lambda I)(A^2 - \lambda^2 I)^{k-1}x = 0 \},
\end{equation}

and

\begin{equation}
\tilde{W}_\lambda^{(k)} := \{ x \in W : (A - \lambda I)^{k}x = 0 \}.
\end{equation}

The following two lemmas address the relationship between the filtrations $\left\{ W_\lambda^{(k)} \right\}$, $\left\{ W_\lambda^{(k)}\pm \right\}$, and $\left\{ \tilde{W}_\lambda^{(k)} \right\}$, defined by (2.3), (4.1), and (4.2), respectively. Note that, for each $k$, $W_\lambda^{(k)}\pm$ and $\tilde{W}_\lambda^{(k)}$ are vector spaces over $\mathbb{R}$ but not over $\mathbb{C}$. In principle, these lemmas can be deduced from the Hong–Horn canonical forms for antilinear operators from [10, Theorem 3.1] (see also Remark 2.3 above), but we prefer to give an independent geometric proof of these Lemmas, first, in order to make the presentation self-contained (as the source [8], where [10, Theorem 3.1] is proved, is not easily available), second, because our proofs of these Lemmas are the main ingredients in the new geometric proof of Hong and Horn’s result (outlined in section §4.5), and, third, because this proof seems to be interesting by itself.

**Lemma 4.2.** For all positive integers $k$, we have $W_\lambda^{(k)} / W_\lambda^{(k-1)} = W_\lambda^{(k)+} / W_\lambda^{(k-1)} \oplus W_\lambda^{(k)-} / W_\lambda^{(k-1)}$. Moreover, $W_\lambda^{(k)} / W_\lambda^{(k-1)} = \text{span}_\mathbb{C} \left( W_\lambda^{(k)+} / W_\lambda^{(k-1)} \right)$.

**Proof.** If $x \in W_\lambda^{(k)}$ then $\lambda x \pm Ax \in W_\lambda^{(k)}$ because

\[(A \mp \lambda I)(A^2 - \lambda^2 I)^{k-1}(\lambda x \pm Ax) = (A^2 - \lambda^2 I)^{k}x = 0,
\]

which shows

\[\left\{ \lambda x \pm Ax : x \in W_\lambda^{(k)} \right\} / W_\lambda^{(k-1)} \subset W_\lambda^{(k)+} / W_\lambda^{(k-1)}.
\]

Accordingly,

\[W_\lambda^{(k)} / W_\lambda^{(k-1)} = \left\{ \lambda x - Ax : x \in W_\lambda^{(k)} \right\} / W_\lambda^{(k-1)} \oplus \left\{ \lambda x + Ax : x \in W_\lambda^{(k)} \right\} / W_\lambda^{(k-1)} \subset W_\lambda^{(k)+} / W_\lambda^{(k-1)} \oplus W_\lambda^{(k)-} / W_\lambda^{(k-1)}
\]

\[\subset W_\lambda^{(k)} / W_\lambda^{(k-1)},
\]

where ** holds because $A^2 - \lambda^2 I = (A + \lambda I)(A - \lambda I)$ and * holds for the following reason.

Both $\left\{ \lambda x - Ax : x \in W_\lambda^{(k)} \right\} / W_\lambda^{(k-1)}$ and $\left\{ \lambda x + Ax : x \in W_\lambda^{(k)} \right\} / W_\lambda^{(k-1)}$ are disjoint subsets of
$W_{\lambda}^{(k)}/W_{\lambda}^{(k-1)}$ because they belong to the kernel of $A + \lambda I : W_{\lambda}^{(k)}/W_{\lambda}^{(k-1)} \to W_{\lambda}^{(k)}/W_{\lambda}^{(k-1)}$ and $A - \lambda I : W_{\lambda}^{(k)}/W_{\lambda}^{(k-1)} \to W_{\lambda}^{(k)}/W_{\lambda}^{(k-1)}$ respectively, and these kernels are disjoint because if $v$ is in both kernels then $\lambda v \equiv -\lambda v \pmod{W_{\lambda}^{(k-1)}}$. This shows that the direct sum on the right side of $*$ is naturally a subset of $W_{\lambda}^{(k)}/W_{\lambda}^{(k-1)}$. On the other hand, for any $v \in W_{\lambda}^{(k)}/W_{\lambda}^{(k-1)}$, we have

$$v = (\lambda \frac{v}{\lambda}) + (\lambda + \frac{v}{\lambda} + A \frac{v}{\lambda})$$

which shows that $W_{\lambda}^{(k)}/W_{\lambda}^{(k-1)}$ is contained in the direct sum on the right side of $*$. By (4.3), $W_{\lambda}^{(k)}/W_{\lambda}^{(k-1)} = \text{span}_C \left( W_{\lambda}^{(k)+}/W_{\lambda}^{(k-1)} \right)$ because $W_{\lambda}^{(k)+}/W_{\lambda}^{(k-1)} = iW_{\lambda}^{(k)-}/W_{\lambda}^{(k-1)}$. □

**Remark 4.3.** Notice, we have already used the special condition $\lambda^2 > 0$ of §4.1, because Lemma 4.2 relies on the fact that $A^2 - \lambda^2 I = (A + \lambda I)(A - \lambda I)$.

**Lemma 4.4.** Any basis of the real vector space $\widetilde{W}_{\lambda}^{(k)}$ is also a basis of the complex vector space $W_{\lambda}^{(k)}$.

**Proof.** When $k = 1$, the statement follows from Lemma 4.2 because $\widetilde{W}_{\lambda}^{(1)} = W_{\lambda}^{(1)+}$ and $W_{\lambda}^{(0)} = 0$. Proceeding by induction, let us assume any basis of the real vector space $\widetilde{W}_{\lambda}^{(k-1)}$ is also a basis of the complex vector space $W_{\lambda}^{(k-1)}$. Suppose $\dim \widetilde{W}_{\lambda}^{(k-1)} = l$ and $\dim W_{\lambda}^{(k-1)} = m$, and let $\{e_1, \ldots, e_{l+m}\}$ be a basis of $\widetilde{W}_{\lambda}^{(k)}$. Without loss of generality, we can assume $\{e_1, \ldots, e_l\} \subset \widetilde{W}_{\lambda}^{(k-1)}$ because this assumption does not change the real or complex span of $\{e_1, \ldots, e_{l+m}\}$.

First, we show that the vectors $e_{l+1}, \ldots, e_{l+m}$ are linearly independent over $\mathbb{C}$ modulo $W_{\lambda}^{(k-1)}$. For this, consider a vector $v \in \text{span}_C \{e_{l+1}, \ldots, e_{l+m}\}$ with coefficients $\alpha_{l+1}, \ldots, \alpha_{l+m} \in \mathbb{R}$ and $\beta_{l+1}, \ldots, \beta_{l+m} \in \mathbb{R}$ such that

$$v := \sum_{j=l+1}^{l+m} (\alpha_j + i\beta_j)e_j \in W_{\lambda}^{(k-1)}.$$

Set

$$v_+ := \sum_{j=l+1}^{l+m} \alpha_j e_j \quad \text{and} \quad v_- := \sum_{j=l+1}^{l+m} i\beta_j e_j.$$

Since $\widetilde{W}_{\lambda}^k$ is a real vector space, $v_+ \in \widetilde{W}_{\lambda}^k$, and hence, by (4.1),

$$A(A - \lambda I)^{k-1}v_+ = \lambda(A - \lambda I)^{k-1}v_+.$$ 

Therefore

$$A + \lambda I)(A - \lambda I)^{k-1}v_+ = 2\lambda(A - \lambda I)^{k-1}v_+$$

Notice $(A + \lambda I)^{k}v_- = 0$ because, for all $l < j \leq l + m$, $(A + \lambda I)^{k}(i\beta_j e_j) = -i\beta_j (A - \lambda I)^{k}e_j = 0$.

Since $v \in W_{\lambda}^{(k-1)}$ and $(A + \lambda I)^{k}v_- = 0$,

$$0 = (A + \lambda I)(A^2 - \lambda^2 I)^{k-1}v = (A + \lambda I)(A^2 - \lambda^2 I)^{k-1}v_+ + (A - \lambda I)^{k-1}(A + \lambda I)^{k}v_- = (A + \lambda I)(A^2 - \lambda^2 I)^{k-1}v_+,$$

and hence

$$A^2 - \lambda^2 I)^{k-1}v_+ \in \ker(A + \lambda I).$$

Furthermore, $(A - \lambda I)^{k}v_+ = 0$ because, for all $l < j \leq l + m$, $(A - \lambda I)^{k}\alpha_j e_j = 0$, so

$$A^2 - \lambda^2 I)^{k-1}v_+ \in \ker(A - \lambda I).$$
Yet, \( \ker(A - \lambda I) \cap \ker(A + \lambda I) = 0 \), (4.5) and (4.6) imply

\[
(4.7) \quad v_+ \in \ker(A^2 - \lambda^2 I)^{k-1},
\]

and (4.4) implies

\[
(4.8) \quad (A^2 - \lambda^2 I)^{k-1}v_+ = (A + \lambda I)^{k-1}(A - \lambda I)^{k-1}v_+ = (2\lambda)^{k-1}(A - \lambda I)^{k-1}v_+.
\]

Together, (4.7) and (4.8) imply that

\[
v_+ \in \ker(A - \lambda I)^{k-1} = \tilde{W}_{\lambda}^{(k-1)} = \text{span}_\mathbb{R}\{e_1, \ldots, e_l\}
\]

and hence \( v_+ = 0 \) because \( \text{span}_\mathbb{R}\{e_1, \ldots, e_l\} \cap \text{span}_\mathbb{R}\{e_{l+1}, \ldots, e_{l+m}\} = 0 \). Note that \( v_+ = 0 \) implies \( \alpha_{l+1} = \cdots = \alpha_{l+m} = 0 \) because \( e_{l+1}, \ldots, e_{l+m} \) are linearly independent over \( \mathbb{R} \). Repeating the same argument with \( v \) replaced by \( iv \) yields \( v_- = 0 \) and \( \beta_{l+1} = \cdots = \beta_{l+m} = 0 \) as well. Hence \( v = 0 \), which shows that

\[
(4.9) \quad \text{span}_\mathbb{C}\{e_{l+1}, \ldots, e_{l+m}\} \cap W_{\lambda}^{(k-1)} = \text{span}_\mathbb{C}\{e_{l+1}, \ldots, e_{l+m}\} \cap \text{span}_\mathbb{C}\{e_1, \ldots, e_l\} = 0.
\]

Let us now establish the vector space isomorphism \( \tilde{W}_{\lambda}^{(k)} / W_{\lambda}^{(k-1)} \cong W_{\lambda}^{(k)} / W_{\lambda}^{(k-1)} \). The cosets

\[
e_{l+1} + W_{\lambda}^{(k-1)}, \ldots, e_{l+m} + W_{\lambda}^{(k-1)}
\]

are linearly independent vectors (over \( \mathbb{R} \)) in the space \( W_{\lambda}^{(k)} / W_{\lambda}^{(k-1)} \). If we take an arbitrary vector \( w + W_{\lambda}^{(k-1)} \in W_{\lambda}^{(k)} / W_{\lambda}^{(k-1)} \) then \( (A - \lambda I)w \in W_{\lambda}^{(k-1)} \), so

\[
(A + \lambda I)w \equiv 2\lambda w \pmod{W_{\lambda}^{(k-1)}}.
\]

Hence,

\[
(4.10) \quad (2\lambda)^{1-k}(A + \lambda I)^{k-1}w \equiv w \pmod{W_{\lambda}^{(k-1)}}.
\]

Now observe that \( (A + \lambda I)^{k-1}w \in \tilde{W}_{\lambda}^{(k)} \). Indeed, from the definitions (2.3) and (4.1) and the fact that \( w \in W_{\lambda}^{(k)} + \), it follows that

\[
(A - \lambda I)^{k-1}(A + \lambda I)^{k-1}w = (A - \lambda I)(A^2 - \lambda^2 I)^{k-1}w = 0.
\]

Hence, by (4.10), \( w \in \tilde{W}_{\lambda}^{(k)} \). Therefore, there exist real coefficients \( a_{l+1}, \ldots, a_{l+m} \) such that

\[
w \equiv (2\lambda)^{1-k}(A + \lambda I)^{k-1}w \equiv \sum_{i=l+1}^{l+m} a_i e_i \pmod{W_{\lambda}^{(k)}}.
\]

This shows that the cosets

\[
e_{l+1} + W_{\lambda}^{(k-1)}, \ldots, e_{l+m} + W_{\lambda}^{(k-1)}
\]

form a basis of \( W_{\lambda}^{(k)} / W_{\lambda}^{(k-1)} \). On the other hand, the cosets

\[
e_{l+1} + \tilde{W}_{\lambda}^{(k-1)}, \ldots, e_{l+m} + \tilde{W}_{\lambda}^{(k-1)}
\]

form a basis of \( \tilde{W}_{\lambda}^{(k)} / \tilde{W}_{\lambda}^{(k-1)} \), so the real vector spaces \( \tilde{W}_{\lambda}^{(k)} / \tilde{W}_{\lambda}^{(k-1)} \) and \( W_{\lambda}^{(k)} / W_{\lambda}^{(k-1)} \) are isomorphic.

Applying Lemma 4.2, we get

\[
\dim_\mathbb{C} W_{\lambda}^{(k)} / W_{\lambda}^{(k-1)} = \dim_\mathbb{R} W_{\lambda}^{(k)} / W_{\lambda}^{(k-1)} = \dim_\mathbb{R} W_{\lambda}^{(k)} / W_{\lambda}^{(k-1)} = m,
\]

which implies

\[
\dim_\mathbb{C} W_{\lambda}^{(k)} = \dim_\mathbb{R} W_{\lambda}^{(k-1)} = l + m.
\]
We have shown that \( v \in W_{\lambda}^{(k-1)} \) implies \( \alpha_{t+1} = \cdots = \alpha_m = \beta_{t+1} = \cdots = \beta_{t+m} = 0 \). In particular, we have shown that \( v = 0 \) implies \( \alpha_{t+1} = \cdots = \alpha_m = \beta_{t+1} = \cdots = \beta_{t+m} = 0 \). Therefore, \( e_{t+1}, \ldots, e_{t+m} \) are linearly independent over \( \mathbb{C} \), that is,

\[
\dim_{\mathbb{C}} \text{span}_{\mathbb{C}}\{e_{t+1}, \ldots, e_{t+m}\} = m,
\]

so, by the induction hypothesis and (4.9),

\[
\dim_{\mathbb{C}} \text{span}_{\mathbb{C}}\{e_1, \ldots, e_{t+m}\} = l + m.
\]

Therefore, \( W_{\lambda}^{(k)} = \text{span}_{\mathbb{C}}\{e_1, \ldots, e_{t+m}\} \) because \( e_1, \ldots, e_{t+m} \) are \( l + m \) linearly independent vectors (over \( \mathbb{C} \)) in \( W_{\lambda}^{(k)} \).

**Corollary 4.5.** If \( v \in W_{\lambda}^{(k)} \) then there exist unique vectors \( v_+, v_- \in \widetilde{W}_{\lambda}^{(k)} \) such that \( v = v_+ + iv_- \).

Define \( s_1 \) to be the minimal natural number such that \( W_{\lambda}^{(s_1)} = W_{\lambda}^{(n)} \). We would like to find a vector \( v \in \widetilde{W}_{\lambda}^{(s_1)} \) such that the space

\[
V = \text{span}_{\mathbb{C}}\{v, (A - \lambda I)v, \ldots, (A - \lambda I)^{s_1 - 1}v\}
\]

is an \( s_1 \)-dimensional \( A \)-invariant space on which \( \ell \) is nondegenerate because we can then normalize \( A \) and \( \ell \) on the space \( V \) and on the \( \ell \)-orthogonal complement of \( V \) separately. Proceeding throughout this subsection, for \( v \in W_{\lambda}^{(n)} \), we adopt the notation of letting \( v_+, v_- \in \widetilde{W}_{\lambda}^{(n)} \) be the unique vectors such that \( v = v_+ + iv_- \), as given in Corollary 4.5.

**Lemma 4.6.** If \( H \) is a Hermitian \( k \times k \) matrix, \( \lambda > 0 \), and \( HJ_{\lambda,k} \) is symmetric, then \( H \) is a Hankel matrix satisfying

\[
H_{i,j} = 0 \quad \forall i + j \leq k.
\]

**Proof.** Let \( H' \) be the upper left \((k-1) \times (k-1)\) block of \( H \). Symmetry of \( HJ_{\lambda,k} \) implies that \( H'J_{\lambda,k-1} \) is symmetric. Since the Lemma is vacuously true for \( k = 1 \), we can proceed by induction, and assume \( H' \) is a Hankel matrix satisfying

\[
H'_{i,j} = 0 \quad \forall i + j \leq k - 1.
\]

Computing the \((1,k)\) and \((k,1)\) entries of \( HJ_{\lambda,k} \) yields

\[
\left(HJ_{\lambda,k}\right)_{1,k} = \lambda H_{1,k} + H_{1,k-1} \quad \text{and} \quad \left(HJ_{\lambda,k}\right)_{k,1} = \lambda H_{1,k}.
\]

Symmetry of \( HJ_{\lambda,k} \) allows us to equate the terms, so

\[
H_{1,k-1} = \lambda \left( H_{1,k} - H_{1,k} \right) \in \{iz \mid z \in \mathbb{R}\}.
\]

Yet, since \( H' \) is both Hankel and Hermitian, its entries are all real numbers. In particular, \( H_{1,k-1} \in \mathbb{R} \), so

\[
H_{1,k-1} = 0 \quad \text{and} \quad H_{1,k} = H_{k,1} \in \mathbb{R}.
\]

Equating \( \left(HJ_{\lambda,k}\right)_{2,k} \) with \( \left(HJ_{\lambda,k}\right)_{k,2} \) yields

\[
H_{1,k} - H_{2,k-1} = \lambda \left( H_{2,k} - H_{2,k} \right) \in \{iz \mid z \in \mathbb{R}\}
\]

which implies \( H_{k,1} = H_{2,k-1} \) because, by the induction hypothesis, \( H_{2,k-1} \in \mathbb{R} \). Accordingly,

\[
H_{i,j} = H_{1,k} \quad \forall i + j = k + 1
\]

because \( H' \) is Hankel.

We conclude this proof with induction. Supposing, for some \( 1 < m \leq k \), we have \( H_{k,j} = H_{j,k} \) and \( H_{k,j} = H_{j+1,k-1} \) for all \( j < m \), let us establish that \( H_{k,m} = H_{m,k} \) and \( H_{k,m} = H_{m+1,k-1} \), where we interpret \( H_{k,m} = H_{m+1,k-1} \) as vacuously true for \( m = k \) (i.e., since \( H_{i,j} \) is only defined
for \( \max\{i,j\} \leq k \), we can extend the definition of \( H_{i,j} \) for \( \max\{i,j\} > k \) in a way that satisfies the equations \( H_{k,k+j} = H_{k+j,k} \) and \( H_{k,j+1,k} = H_{k+1,k,j} \) for all \( j \) by construction, and, of course, this extension’s definition has no relevance to the normalization of \( H \). Symmetry of \( HJ_{\lambda,k} \) implies

\[
H_{m,k-1} + \lambda H_{m,k} = (HJ_{\lambda,k})_{m,k} = (HJ_{\lambda,k})_{k,m} = H_{k,m-1} + \lambda H_{k,m},
\]

and hence

\[
H_{m,k} - H_{k,m} = \lambda^{-1}(H_{k,m-1} - H_{m,k-1}) = 0
\]

because, by the induction hypothesis, \( H_{k,m-1} = H_{m,k-1} \). If \( m = k \) then there is nothing more to check, that is, \( H_{k,m} = H_{m+1,k-1} \) is vacuously true. Similarly, if \( m = k - 1 \) then we have already shown \( H_{k,m} = H_{m+1,k-1} \). For \( m < k - 1 \), we have

\[
H_{m+1,k-1} + \lambda H_{m+1,k} = (HJ_{\lambda,k})_{m+1,k} = (HJ_{\lambda,k})_{k,m+1} = H_{k,m} + \lambda H_{k,m+1},
\]

and hence

\[
H_{k,m} - H_{m+1,k-1} = \lambda(H_{m+1,k} - H_{m+1,k+1}) = \lambda(H_{m+1,k} - \overline{H_{m+1,k+1}}) \in \{z \mid iz \in \mathbb{R}\},
\]

which implies \( H_{k,m} = H_{m+1,k-1} \) because, since \( H' \) is a real matrix, \( H_{m+1,k-1} \in \mathbb{R} \). This completes the proof by induction.

**Lemma 4.7.** If a nondegenerate Hermitian form \( \ell \) and antilinear operator \( A \) are represented respectively by the \( k \times k \) matrices \( H \) and \( J_{\lambda,k} \), where \( H \) is a Hankel matrix satisfying

\[
H_{i,j} = 0 \quad \forall i + j \leq k,
\]

then there is a basis with respect to which \( \ell \) and \( A \) are represented by \( \pm S_k \) and \( J_{\lambda,k} \) respectively.

**Proof.** Every transformation of the matrices representing \( \ell \) and \( A \) given by the rule (2.2) can be induced by a change of basis, so it will suffice to find \( M \) such that

\[
(4.12) \quad M^*HM = S_k \quad \text{and} \quad M^{-1}J_{\lambda,k}\overline{M} = J_{\lambda,k}.
\]

To satisfy \( M^{-1}J_{\lambda,k}\overline{M} = J_{\lambda,k} \), let us suppose \( M \) is a real upper-triangular Toeplitz matrix, and define \( h_0, \ldots, h_{k-1} \in \mathbb{R} \) and \( \alpha_0, \ldots, \alpha_{k-1} \in \mathbb{R} \) to be the coefficients for which

\[
H = S_k \left( \sum_{i=1}^{k} h_{i-1}T_k^{i-1} \right) \quad \text{and} \quad M = \sum_{i=1}^{k} \alpha_{i-1}T_k^{i-1}.
\]

Note, \( h_1, \ldots, h_k \) must be real because \( H \) is Hermitian and Hankel. For our particular choice of \( M \), we have \( M^* = S_k\overline{M}S_k \), so

\[
M^*HM = S_k \left( \sum_{i=1}^{k} \alpha_{i-1}T_k^{i-1} \right) \left( \sum_{i=1}^{k} h_{i-1}T_k^{i-1} \right) \left( \sum_{i=1}^{k} \alpha_{i-1}T_k^{i-1} \right) = S_k \left( \sum_{i=0}^{k-1} \sum_{r+s+t=i} \alpha_r \alpha_s h_t T_k^{i} \right).
\]

Therefore, we need to solve the equation

\[
(4.13) \quad \sum_{i=0}^{k-1} \sum_{r+s+t=i} \alpha_r \alpha_s h_t T_k^{i} = \pm I_k,
\]

that is, we need to choose \( \alpha_i \) such that (4.13) holds. Comparing entries of the main diagonal in (4.13), we find that \( \alpha_0 = h_0^{-1/2} \), so let us choose \( \alpha_0 = |h_0|^{-1/2} \). Note, \( h_0 \neq 0 \) because \( \ell \) is nondegenerate, and hence this choice of \( \alpha_0 \) is well defined. Having fixed \( \alpha_0 \), comparing entries in the first super-diagonal of (4.13) shows that we can choose \( \alpha_1 \) as the solution to a linear equation with real coefficients so that entries in the first super-diagonal of (4.13) match. Proceeding similarly, for \( 1 < j < k \), after choosing \( \alpha_0, \ldots, \alpha_{j-1} \) so that entries in the main diagonal and the first \( j-1 \) super-diagonals of (4.13) match, comparing entries in the \( j \) super-diagonal of (4.13) shows that we
can choose \( \alpha_j \) as the solution to a linear equation with real coefficients so that entries in the \( j \) super-diagonal of (4.13) match; moreover, the variable \( \alpha_j \) does not appear in the first \( j - 1 \) super-diagonals of (4.13), so, by choosing \( \alpha_j \) in this way, we ensure that entries the first \( j \) super-diagonals of (4.13) match. By choosing \( \alpha_0, \ldots, \alpha_k \) in this way we obtain (4.12) by construction. \( \square \)

**Lemma 4.8.** There exists a vector \( v \in W_{\lambda}^{(s_1)} \) such that the space in (4.11) is an \( s_1 \)-dimensional \( A \)-invariant space on which \( \ell \) is nondegenerate.

**Proof.** It can be seen from the Gohberg–Lancaster–Rodman canonical form for \( \ell \) and \( A^2 \) (given in [6, Theorem 5.1.1] and summarized in Remark 2.1) that there exists a vector \( v' \in W_{\lambda}^{(s_1)} \) for which

\[
(4.14) \quad \ell \left( v', (A^2 - \lambda^2 I)^{(s_1-1)} v' \right) \neq 0.
\]

Using the decomposition of Corollary 4.5, define the coefficients

\[
a_0 := \ell(v'_+, (A - \lambda I)^{(s_1-1)} v'_+) + \ell(v'_-, (A - \lambda I)^{(s_1-1)} v'_-),
\]
\[
a_1 := \ell(v'_+, (A - \lambda I)^{(s_1-1)} v'_+) - \ell(v'_-, (A - \lambda I)^{(s_1-1)} v'_-),
\]
and

\[
b_1 := \ell(v'_-, (A - \lambda I)^{(s_1-1)} v'_+),
\]

By direct computation, we obtain the finite Fourier series

\[
(4.15) \quad 2(2\lambda)^{1-s_1} \ell \left( e^{i\theta} v', (A - \lambda I)^{s_1-1} \left( e^{i\theta} v' \right)_+ \right) = a_0 \cos(2\theta) + b_1 \sin(2\theta)
\]

Also, since \( v'_+, v'_- \in W_{\lambda}^{(s_1)} \), \( A(A - \lambda)^{s_1-1} v'_+ = (A - \lambda I)^{s_1-1} v'_+ \) and \( A(A + \lambda)^{s_1-1} iv'_- = -\lambda(A + \lambda)^{s_1-1} iv'_- \), and hence

\[
(A^2 - \lambda^2 I)^{(s_1-1)} v' = (A + \lambda I)^{(s_1-1)}(A - \lambda I)^{(s_1-1)} v'_+ + (A - \lambda I)^{(s_1-1)}(A + \lambda I)^{(s_1-1)} v'_-
\]
\[
= (2\lambda)^{s_1-1}(A - \lambda I)^{(s_1-1)} v'_+ + i(2\lambda)^{s_1-1}(A - \lambda I)^{(s_1-1)} v'_-.
\]

So, by (4.14),

\[
0 \neq (2\lambda)^{1-s_1} \ell \left( v', (A^2 - \lambda^2 I)^{(s_1-1)} v' \right) = a_0 + ib_1.
\]

If the left side of (4.15) is zero for all \( \theta \in \mathbb{R} \) then \( a_0 = a_1 = b_1 = 0 \), so, by (4.14), there exists \( \theta \in \mathbb{R} \) such that

\[
(4.16) \quad \ell \left( e^{i\theta} v', (A - \lambda I)^{s_1-1} \left( e^{i\theta} v' \right)_+ \right) \neq 0.
\]

Fixing \( \theta \in \mathbb{R} \) so that (4.16) holds, define

\[
(4.17) \quad v := \left( e^{i\theta} v' \right)_+,
\]

so, by (4.16),

\[
(4.18) \quad \ell(v, (A - \lambda I)^{s_1-1} v) \neq 0.
\]

Proceeding, let \( V \) be as in (4.11) with \( v \) as in (4.17). Define basis vectors

\[
e_i = (A - \lambda I)^{i-1} v \quad (i = 1, \ldots, s_1).
\]

The matrix representing the restriction \( A|_V \) of \( A \) to \( V \) with respect to the basis \( \{e_i\}_{1 \leq i \leq s_1} \) is \( J_{\lambda, s_1} \). Let \( H \) be the matrix representing the restriction of \( \ell \) to \( V \) with respect to the basis \( \{e_i\}_{1 \leq i \leq s_1} \). Since \( A \) is \( \ell \)-self-adjoint, \( HJ_{\lambda, s_1} \) is symmetric. Therefore, applying Lemma 4.6, \( H \) is a Hankel matrix satisfying

\[
H_{i,j} = 0 \quad \forall i + j \leq k,
\]
and hence, by (4.18),
\[
\det H = \sqrt{2} \sin \left(\frac{2k\pi + \pi}{4}\right) H_{1,k} = \sqrt{2} \sin \left(\frac{2k\pi + \pi}{4}\right) \left(\ell (v, (A - \lambda I)^{s_1 - 1} v)\right)^k \neq 0.
\]
That is, \(\ell\) is nondegenerate on \(V\), as was needed. \(\square\)

**Corollary 4.9.** There is an \(s_1\)-dimensional \(A\)-invariant space \(V\) on which \(\ell\) is nondegenerate, and there is a basis of \(V\) with respect to which the restrictions \(\ell|_V\) and \(A|_V\) of \(\ell\) and \(A\) to \(V\) are represented by the matrices \(\pm H_{\lambda,s_1}\) and \(C|_{\lambda,s_1}\) respectively.

**Proof.** By Lemma 4.8, there exists an \(s_1\)-dimensional \(A\)-invariant space \(V\) on which \(\ell\) is nondegenerate and there exists a basis of \(V\) with respect to which the restrictions \(\ell|_V\) and \(A|_V\) of \(\ell\) and \(A\) to \(V\) are represented by the matrices \(H_{\lambda,s_1}\) and \(J_{\lambda,s_1}\), where \(H\) is a Hankel matrix satisfying
\[
H_{i,j} = 0 \quad \forall i + j \leq s_1.
\]
Therefore, by Lemma 4.7, there is a basis \(\{e_1, \ldots, e_{s_1}\}\) of \(V\) with respect to which \(\ell|_V\) and \(A|_V\) are represented by \(S_{s_1}\) and \(J_{\lambda,s_1}\) respectively. If \(\lambda > 0\) then this completes the proof because \(J_{\lambda,s_1} = C|_{\lambda,s_1}\). If, on the other hand, \(\lambda < 0\) then we observe \(\ell|_V\) and \(A|_V\) are represented by \(S_{s_1}\) and \(J_{-\lambda,s_1} = C|_{\lambda,s_1}\) with respect to the basis \(\{ie_1, \ldots, ie_{s_1}\}\). So, in either case, we can find a basis with respect to which \(\ell|_V\) and \(A|_V\) are represented by \(H_{\lambda,s_1} = S_{s_1}\) and \(C|_{\lambda,s_1}\). \(\square\)

For the following proposition, let \(r_1, \ldots, r_{n_\lambda}\) and \(s_1, \ldots, s_{n_\lambda}\) be the positive integers satisfying \(s_i > s_{i+1}\) such that the restriction of \(A^2\) to \(W^{(n)}_{\lambda}\) has a Jordan canonical form with \(r_i\) Jordan blocks of size \(s_i \times s_i\). Note, this definition is consistent with the previous definition of \(s_1\), and
\[
W^{(n)}_{\lambda} \cong \mathbb{C}^\mu \quad \text{where} \quad \mu = \sum_{i=1}^{n_\lambda} r_i s_i.
\]

**Proposition 4.10.** There is a basis of \(W^{(n)}_{\lambda}\) with respect to which the restrictions of \(\ell\) and \(A\) to \(W^{(n)}_{\lambda}\) are represented by the matrices
\[
\bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} H_{\lambda,s_i} \right) \quad \text{and} \quad \bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} C|_{\lambda,s_i} \right)
\]
respectively.

**Proof.** By Corollary 4.9, there is a space \(V \subset W^{(n)}_{\lambda}\) that is \(A\)-invariant and \(\ell\)-nondegenerate on which \(\ell\) and \(A\) can be represented by matrices of the desired form. By Lemma 4.1, we can normalize \(\ell\) and \(A\) on \(V\) and the \(\ell\)-orthogonal complement \(V^\perp\) of \(V\) separately, so we can repeat this process, applying Corollary 4.9 to \(V^\perp\) rather than \(W^{(n)}_{\lambda}\). Iterating the process \(\sum_{i=1}^{n_\lambda} r_i\) times completes the normalization. \(\square\)

**§ 4.2. Treating Generalized Eigenspaces with Eigenvalue Zero.** In this subsection we construct a canonical form for the restrictions of \(\ell\) and \(A\) to the space \(W^{(n)}_{0}\). Our approach is the same as in the proof of Theorem 4.5 in [14].

**Proposition 4.11.** The space \(W^{(n)}_{0}\) can be decomposed into \(A\)-invariant, pairwise \(\ell\)-orthogonal subspaces such that there exists a basis with respect to which the restrictions of \(\ell\) and \(A\) to the decomposition’s component subspaces are represented by matrices of the form \(C_{0,k}\) and \(H_{0,k}\).

**Proof.** Let
\[
k = \min \left\{ m \in \mathbb{N} : \left( A|_{W^{(n)}_{0}} \right)^m \equiv 0 \right\}.
\]
Fix a basis, and let $H$ and $C$ be matrices representing $\ell$ and $A$ with respect to this basis. If $k$ is odd, then $A^{k-1}$ is $\ell$-self-adjoint linear, which implies $HC^{k-1}$ is Hermitian, and hence there is a basis with respect to which the mapping

$$v \mapsto H \left( A^{k-1}v \right)$$

is represented by a nonzero diagonal matrix. If, on the other hand, $k$ is even, then $A^{k-1}$ is $\ell$-self-adjoint antilinear, which implies $HC^{k-1}$ is symmetric. By Takagi’s theorem in [15, Theorem 2], for every symmetric matrix $S$, there exists an invertible matrix $U$ such that $USU^{-1}$ is diagonal, and, since the map in (4.19) is antilinear whenever $k$ is even, Takagi’s theorem implies that there is a basis with respect to which the mapping in (4.19) is represented by a nonzero diagonal matrix.

For either parity of $k$, these observations imply that there exists a vector $a_1 \neq 0$ such that

$$H \left( A^{k-1}a_1 \right) = \gamma e^{i\theta} a_1 \quad \text{for some } \theta \in \mathbb{R}, \gamma > 0.$$ 

Furthermore, if $k$ is odd then $\theta \equiv 0 \pmod{\pi}$ because (4.19) is a linear operator represented by a Hermitian matrix. Accordingly, for $z \in \mathbb{C},$

$$\ell(A^{-1}za_1, za_1) = (HA^{-1}za_1, za_1) = \begin{cases} \pm \gamma |z|^2 ||a_1||^2 & \text{if } k \text{ is odd} \\ \gamma e^{i\theta} |z|^2 ||a_1||^2 & \text{if } k \text{ is even.} \end{cases}$$

Therefore

$$\ell \left( A^{-1} \frac{1}{\sqrt{\gamma ||a_1||^2 e^{i\theta/2}}} a_1, \frac{1}{\sqrt{\gamma ||a_1||^2 e^{i\theta/2}}} a_1 \right) = \pm 1.$$ 

Define

$$\tilde{e}_i = A^{-i-1} \frac{1}{\sqrt{\gamma ||a_1||^2 e^{i\theta/2}}} a_1,$$

and define

$$e_1 = \tilde{e}_1 + \alpha_2 \tilde{e}_2 + \ldots + \alpha_k \tilde{e}_k \quad \text{and} \quad e_i = A^{-i-1} e_1,$$

where the coefficients $\alpha_2, \ldots, \alpha_k$ are chosen below. For all $i + j > k + 1$ we have

$$\ell(e_i, e_j) = \ell(A^{-i-1} e_1, A^{-j-1} e_1) = \ell(A^{i+j-2} e_1, e_1) = \ell(0, e_1) = 0.$$ 

Fix the coefficients $\alpha_2, \ldots, \alpha_k$ such that for all $j < k$ we have

$$\ell(e_1, e_j) = 0.$$ 

Since $A$ is $\ell$-self-adjoint, we have

$$\ell(e_i, e_{i+j}) = \ell(e_{i+j}, e_i),$$

so our choices of $\alpha_2, \ldots, \alpha_k$ ensure

$$\ell(e_i, e_j) = 0 \quad \forall i + j < k + 1.$$ 

By construction,

$$\ell(e_1, e_k) = \ell(e_j, e_{k+j-1}) = 1,$$

so the restrictions of $\ell$ and $A$ to the subspace $\text{span}_\mathbb{C}\{e_1, \ldots, e_k\}$ are represented by $S_k$ and $J_{0,k}$ respectively with respect to the basis $\{e_k, \ldots, e_1\}$.

By Lemma 4.1, we can normalize $\ell$ and $A$ on $\text{span}_\mathbb{C}\{e_1, \ldots, e_k\}$ and the orthogonal complement of $\text{span}_\mathbb{C}\{e_1, \ldots, e_k\}$ separately, so this normalization procedure can be repeated on the orthogonal complement of $\text{span}_\mathbb{C}\{e_1, \ldots, e_k\}$ until $W_0^{(n)}$ is exhausted. \qed
4.3. Treating Generalized Eigenspaces with Negative Eigenvalues. Throughout this subsection we assume $\lambda^2 < 0$ and that the restriction of $A^2$ to $W^{(n)}_\lambda$ has a Jordan canonical form with $2r_i$ Jordan blocks of size $s_i \times s_i$, where $r_1, \ldots, r_{n_\lambda}$ and $s_1, \ldots, s_{n_\lambda}$ are positive integers satisfying $s_i > s_{i+1}$.

**Proposition 4.12.** There is a basis of $W^{(n)}_\lambda$ with respect to which the restrictions of $\ell$ and $A$ to $W^{(n)}_\lambda$ are represented by the matrices

$$
\bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} \epsilon_{i,j} H_{\lambda,s_i} \right) \text{ and } \bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} C_{\lambda,s_i} \right)
$$

where $\epsilon_{i,j} = \pm 1$ respectively.

**Proof.** Given the Gohberg–Lancaster–Rodman canonical form for $\ell$ and $A^2$ summarized in Remark 2.1, there exists a vector $a_1 \in W^{(n)}_\lambda$ such that the restrictions of $\ell$ and $A^2$ to the $s_1$-dimensional vector space span$_\mathbb{C} \{a_1, (A^2 - \lambda^2 I)a_1, \ldots, (A^2 - \lambda^2 I)^{s_1-1}a_1\}$ are represented respectively by $S_{s_1}$ and $J_{\lambda^2,s_1}$ with respect to the basis $\{(A^2 - \lambda^2 I)^{s_1-1}a_1, (A^2 - \lambda^2 I)^{s_1-2}a_1, \ldots, a_1\}$.

Defining

$$a_{k+1} = (A^2 - \lambda^2 I)a_k \text{ and } b_k = Aa_k,$$

and letting $e_1 = a_1 + \beta b_1$, we have

$$\ell(e_1, A(A^2 - \lambda^2 I)^{s_1-1}e_1) = \alpha^2 \ell(a_1, b_{s_1}) + \alpha \beta (\ell(a_1, Ab_{s_1}) + \ell(b_1, b_{s_1})) + \beta^2 \ell(b_1, Ab_{s_1})$$

$$= \alpha^2 \ell(a_1, b_{s_1}) + \alpha \beta (\lambda^2 \ell(a_1, a_{s_1}) + \lambda^2 \ell(a_{s_1}, a_{s_1})) + \lambda^2 \beta^2 \ell(b_1, a_{s_1})$$

$$= \alpha^2 \ell(a_1, b_{s_1}) + 2\lambda^2 \alpha \beta + \lambda^2 \beta^2 \ell(b_1, a_{s_1}).$$

Clearly, either $\ell(a_1, b_{s_1}) = 0$ or we can choose $\alpha, \beta \in \mathbb{C}$ such that $(\alpha, \beta) \neq (0, 0)$ and

$$\ell(e_1, A(A^2 - \lambda^2 I)^{s_1-1}e_1) = 0.$$

Accordingly, we can assume, by possibly replacing $a_1$ with $e_1$ as defined above, that $\ell(a_1, b_{s_1}) = 0$.

With this assumption made, we proceed with $e_1$ defined as above, and will determine the coefficients $\alpha$ and $\beta$ later. Note, this assumption implies also that $\ell(b_1, a_{s_1}) = 0$ because $(A^2 - \lambda^2 I)$ is an $\ell$-self-adjoint linear operator, and hence

$$\ell(e_1, A(A^2 - \lambda^2 I)^{s_1-1}e_1) = \pm 2\lambda^2 \alpha \beta.$$

Define

$$e_k := (A^2 - \lambda^2 I)^{k-1}e_1 \text{ and } e_{s_1+k} := Ae_k \quad \forall 1 \leq k \leq s_1,$$

and, on the span of $\{e_i\}$, let $\ell$ and $A$ be represented with respect to the basis $\{e_{s_1}, \ldots, e_1, e_{2s_1}, \ldots, e_{s_1+1}\}$ by the matrices

$$H = \begin{pmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 0 & J_{\lambda^2,s_1} \\ I_{s_1} & 0 \end{pmatrix}$$

where the each $H_{i,j}$ is an $s_1 \times s_1$ matrix. The matrices $H_{i,j}$ are Hankel because $A^2 - \lambda^2 I$ is $H$-self-adjoint. That is, $H_{1,1}$ is Hankel because

$$\ell(e_i, e_j) = \ell((A^2 - \lambda^2)^{i-1}e_1, (A^2 - \lambda^2)^{j-1}e_1) = \ell(e_1, (A^2 - \lambda^2)^{i+j-2}e_1)$$

$$= \ell(e_1, e_{i+j-1}) \quad \forall i + j \leq s_1 + 1$$

and

$$\ell(e_i, e_j) = \ell((A^2 - \lambda^2)^{i-1}e_1, (A^2 - \lambda^2)^{j-1}e_1) = \ell(e_1, (A^2 - \lambda^2)^{i+j-2}e_1)$$

$$= \ell(e_1, 0) \quad \forall i, j \leq s_1 \text{ with } i + j > s_1 + 1.$$
Similarly, using the identity \((A^2 - \lambda^2 I)v, w) = (v, (A^2 - \lambda^2 I)w)\), we can show \(H_{1,2}, H_{2,1}\), and \(H_{2,2}\) are Hankel.

Since \(A\) is \(\ell\)-self-adjoint, \(HC\) is symmetric, which, as in Lemma 4.6, implies that the \((i, j)\) entry of \(H_{2,1}\) is 0 for all \(i + j < s_1 + 1\). On the other hand, if \(s_1 + 1 < i + j\) then still the \((i, j)\) entry of \(H_{2,1}\) is 0 because

\[
\ell((A^2 - \lambda^2)^{i-1}e_1, (A^2 - \lambda^2)^{j-1}Ae_1) = \ell((A^2 - \lambda^2)^{i+j-2}e_1, Ae_1) = \ell(0, Ae_1) = 0.
\]

Therefore,

\[
H_{1,2} = H_{2,1} = \ell(e_1, e_{2s_1})S_{s_1}.
\]

The same analysis shows that the lower left and upper right \(s_1 \times s_1\) blocks of the matrix representing \(\ell\) with respect to the basis \(\{a_{s_1}, \ldots, a_1, b_1, \ldots, b_1\}\) are also multiples of \(S_{s_1}\), that is,

\[
\ell(a_i, b_j) = \ell(a_1, b_{s_1})\delta_{i+j, s_1+1}.
\]

Direct computation also shows that \(H_{1,1}\) is a multiple of \(S_{s_1}\), that is,

\[
H_{1,1} = \ell(e_1, e_{s_1})S_{s_1},
\]

where

\[
\ell(e_1, e_{s_1}) = |\alpha|^2 \ell(a_1, a_{s_1}) + \alpha \beta \ell(a_1, b_{s_1}) + \beta \bar{\alpha} \ell(b_1, a_{s_1}) + |\beta|^2 \ell(b_1, b_{s_1})
\]

\[
= (|\alpha|^2 + \lambda^2 |\beta|^2) \ell(a_1, a_{s_1}) + \alpha \beta \ell(a_1, b_{s_1}) + \beta \bar{\alpha} \ell(b_1, a_{s_1}) + |\beta|^2 \ell(b_1, b_{s_1})
\]

\[
= \pm (|\alpha|^2 + \lambda^2 |\beta|^2) + \alpha \beta \ell(a_1, b_{s_1}) + \beta \bar{\alpha} \ell(a_1, b_{s_1})
\]

Since \(HC\) is symmetric, it follows that

\[
H_{2,2} = (H_{1,1}J_{\lambda^2, s_1})^T = \ell(e_1, e_{s_1})S_{s_1}J_{\lambda^2, s_1}.
\]

Lastly, fixing \(\alpha = 1\) and \(\beta = 0\), the matrices

\[
H = \pm \left(\begin{array}{cc} S_{s_1} & 0 \\ 0 & S_{s_1}J_{\lambda^2, s_1} \end{array}\right) \quad \text{and} \quad C = \left(\begin{array}{cc} 0 & J_{\lambda^2, s_1} \\ I_{s_1} & 0 \end{array}\right)
\]

represent the restrictions of \(\ell\) and \(A\) on \(\text{span}_C \{e_i\}_{1 \leq i \leq 2s_1}\) with respect to a permutation of the basis \(\{e_i\}_{1 \leq i \leq 2s_1}\). Since \(H\) is nonsingular, By Lemma 4.1, we can repeat this construction on the \(\ell\)-orthogonal complement of \(\text{span}_C \{e_1, \ldots, e_{2s_1}\}\), and hence there exists a basis of \(W_{\lambda}^{(n)}\) with respect to which \(\ell\) and \(A\) are represented by the matrices

\[
(4.21) \quad \bigoplus_{i=1}^{n_\lambda} \left(\bigoplus_{j=1}^{r_i} \epsilon_{i,j} \left(\begin{array}{cc} S_{s_1} & 0 \\ 0 & S_{s_1}J_{\lambda^2, s_1} \end{array}\right)\right) \quad \text{and} \quad \bigoplus_{i=1}^{n_\lambda} \left(\bigoplus_{j=1}^{r_i} \left(\begin{array}{cc} 0 & J_{\lambda^2, s_1} \\ I_{s_1} & 0 \end{array}\right)\right)
\]

where \(\epsilon_{i,j} = \pm\). In particular, we have shown that if there is a basis with respect to which \(\ell\) and \(A\) are represented by

\[
\bigoplus_{i=1}^{n_\lambda} \left(\bigoplus_{j=1}^{r_i} \epsilon_{i,j} S_{2s_1} \right) \quad \text{and} \quad \bigoplus_{i=1}^{n_\lambda} \left(\bigoplus_{j=1}^{r_i} \left(\begin{array}{cc} 0 & J_{\lambda^2, s_1} \\ I_{s_1} & 0 \end{array}\right)\right)
\]

then there is a basis with respect to which \(\ell\) and \(A\) are represented by the matrices in (4.21), and hence, noting (2.2), there exist a matrix \(M\) such that

\[
(M^{-1})^* \bigoplus_{i=1}^{n_\lambda} \left(\bigoplus_{j=1}^{r_i} \epsilon_{i,j} \left(\begin{array}{cc} S_{s_1} & 0 \\ 0 & S_{s_1}J_{\lambda^2, s_1} \end{array}\right)\right) M^{-1} = \bigoplus_{i=1}^{n_\lambda} \left(\bigoplus_{j=1}^{r_i} \epsilon_{i,j} H_{\lambda, s_1} \right)
\]
and
\[
M \bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} \left( \begin{array}{cc} 0 & J_{\lambda^2, s_i} \\ I_{s_i} & 0 \end{array} \right) \right) \mathbf{M}^{-1} = \bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} C_{\lambda, s_i} \right),
\]
which completes the proof. \qed

§4.4. **Treating Generalized Eigenspaces with Nonreal Eigenvalues.** Throughout this subsection we assume \( \lambda^2 \not\in \mathbb{R} \) and that the restriction of \( A^2 \) to \( W_\lambda^{(n)} \) has a Jordan canonical form with \( 2r_i \) Jordan blocks of size \( s_i \times s_i \), where \( r_1, \ldots, 2r_{n_\lambda} \) and \( s_1, \ldots, s_{n_\lambda} \) are positive integers satisfying \( s_i > s_{i+1} \).

**Proposition 4.13.** There is a basis of \( W_\lambda^{(n)} \) with respect to which the restrictions of \( \ell \) and \( A \) to \( W_\lambda^{(n)} \) are represented by the matrices
\[
\bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} \epsilon_{i, j} H_{\lambda, s_i} \right) \quad \text{and} \quad \bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} C_{\lambda, s_i} \right) \quad \text{where} \quad \epsilon_{i, j} = \pm 1
\]
respectively.

**Proof.** Given the Gohberg–Lancaster–Rodman canonical form for \( \ell \) and \( A^2 \) summarized in Remark 2.1, there exist vectors \( a_1, a_1' \in W_\lambda^{(n)} \) such that the restrictions of \( \ell \) and \( A^2 \) to the \( 2s_1 \)-dimensional vector space \( \text{span}_\mathbb{C}\{a_1, (A^2 - \lambda^2 I)a_1, \ldots, (A^2 - \lambda^2 I)^{s_1-1}a_1, a_1', \ldots, (A^2 - \lambda^2 I)^{s_1-1}a_1'\} \) are represented respectively by \( S_{2s_1} \) and \( J_{\lambda^2, s_1} \oplus J_{\lambda, s_1} \) with respect to the basis \( \{(A^2 - \lambda^2 I)^{s_1-1}a_1, \ldots, a_1, (A^2 - \lambda^2 I)^{s_1-1}a_1', \ldots, a_1'\} \). Define
\[
a_{k+1} = (A^2 - \lambda^2 I)a_k \quad \text{and} \quad a_{k+1}' = (A^2 - \lambda^2 I)a_k'.
\]
Our goal is to show that there exists a choice of vector \( a_1 \) such that \( \text{span}_\mathbb{C}\{a_1, (A^2 - \lambda^2 I)a_1, \ldots, (A^2 - \lambda^2 I)^{s_1-1}a_1, a_1', \ldots, (A^2 - \lambda^2 I)^{s_1-1}a_1'\} \) is \( A \)-invariant, so let us proceed assuming otherwise and find a new choice for \( a_1 \) that satisfies this property.

Define
\[
b_k := Aa_k \quad \text{and} \quad b_k' := Aa_k'.
\]
For \( 1 \leq i, j \leq s_1 \),
\[
\ell(b_i, b_j) = \ell(\lambda^2 a_j + a_{j+1}' + a_i) = 0 \quad \text{and} \quad \ell(b_i', b_j') = \ell(\lambda^2 a_j + a_{j+1}' + a_i) = 0,
\]
and
\[
\ell(b_i, b_j') = \ell(\lambda^2 a_j + a_{j+1}' + a_i) \quad \text{and} \quad \ell(b_i', b_j) = \ell(\lambda^2 a_j + a_{j+1}' + a_i).
\]
Therefore, the restrictions of \( \ell \) and \( A^2 \) to the \( 2s_1 \)-dimensional vector space \( \text{span}_\mathbb{C}\{b_1, \ldots, b_{s_1}, b_1', \ldots, b_{s_1}'\} \) are represented respectively by
\[
\begin{pmatrix} 0 & J_{\lambda^2, s_1} S_{s_1} \\ J_{\lambda, s_1} S_{s_1} & 0 \end{pmatrix}
\]
with respect to the basis \( \text{span}_\mathbb{C}\{b_1, \ldots, b_{s_1}, b_1', \ldots, b_{s_1}'\} \).

Letting \( e_1 = \alpha a_1 + \beta b_1 \), we have
\[
\ell(e_1, A^2 - \lambda^2 I e_1) = \alpha^2 \ell(a_1, b_{s_1}') + \alpha \beta \ell(a_1, A^2 a_{s_1}') + \alpha \beta \ell(b_1, b_{s_1}') + \beta^2 \ell(a_1, A^2 a_{s_1}')
\]
\[
= \alpha^2 \ell(a_1, b_{s_1}') + \alpha \beta \ell(a_1, a_{s_1}') + \lambda \ell(a_{s_1}', a_{s_1}') + \beta^2 \ell(a_1, A^2 a_{s_1}')
\]
\[
= \alpha^2 \ell(a_1, b_{s_1}') + \alpha \beta \left( \lambda^2 + \lambda^2 \right) + \lambda^2 \beta^2 \ell(a_1, A^2 a_{s_1}')
\]
(4.22)
Define
\[ e_k = (A^2 - \lambda^2)^{k-1}e_1 \quad \text{and} \quad e_{s_1+k} = Ae_k \quad \forall 1 \leq k \leq s_1 \]
and, on the span of \{e_i\}, let \( \ell \) and \( A \) be represented with respect to the basis \{e_i\} by the matrix
\[
H = \begin{pmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 0 & J_{\lambda^2,s_1} \\ I_{s_1} & 0 \end{pmatrix}
\]
where the matrices \( H_{i,j} \) are each \( s_1 \times s_1 \). Direct computation yields \( H_{1,1} = H_{2,2} = 0 \). Furthermore, \( H_{1,2} \) is symmetric because \( HC \) is symmetric, and hence \( H_{2,1} \) is symmetric as well. If \( s_1 + 1 < i + j \) then the \((i,j)\)th entry of \( H_{1,2} \) is zero because
\[
\ell \left( (A^2 - \lambda^2)^{i-1}e_1, (A^2 - \lambda^2)^{j-1}Ae_1 \right) = \ell \left( (A^2 - \lambda^2)^{i+j-2}e_1, Ae_1 \right) = \ell (0, Ae_1) = 0.
\]
Accordingly,
\[
\det(H) = \pm \left( \ell(e_1, e_{2s_1}) \right)^{2s_1},
\]
which, by (4.22), can be made nonzero for an adequate choice of \( \alpha \) and \( \beta \). Since \( HC \) is symmetric, evaluating the lower right \( s_1 \times s_1 \) block of \( HC \) yields, as in Lemma 4.6, that
\[
H_{2,1} = \overline{H_{1,2}} = \ell(e_1, e_{2s_1})S_{1s_1}.
\]
By replacing \( e_1 \) with \( \sqrt{\frac{1}{\ell(e_1, e_{2s_1})}} e_1 \), we can assume \( \ell(e_1, e_{2s_1}) = \pm 1 \), so the restrictions of \( \ell \) and \( A \) to \( \text{span}_C \{e_1, \ldots, e_{2s_1}\} \) are represented by the matrices
\[
\pm S_{2s_1} \quad \text{and} \quad \begin{pmatrix} 0 & J_{\lambda^2,s_1} \\ I_{s_1} & 0 \end{pmatrix}
\]
with respect to the basis \{e_1, \ldots, e_{2s_1}\}. By Lemma 4.1, we can repeat this normalization procedure on the \( \ell \)-orthogonal complement of \( \text{span}_C \{e_1, \ldots, e_{2s_1}\} \), and hence there is a basis of \( W^{(n)}_\lambda \) with respect to which \( \ell \) and \( A \) are represented by the matrices
\[
\bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} e_{i,j} S_{2s_1} \right) \quad \text{and} \quad \bigoplus_{i=1}^{n_\lambda} \left( \bigoplus_{j=1}^{r_i} \begin{pmatrix} 0 & J_{\lambda^2,s_1} \\ I_{s_1} & 0 \end{pmatrix} \right) \quad \text{where} \ e_{i,j} = \pm 1.
\]

\section*{§4.5. A Canonical Form for Antilinear Operators.}

It is worth noting that methods applied above can be used to obtain the canonical form for antilinear operators (without considering Hermitian forms) given in [10, Theorem 3.1], referred to in Remark 2.3, so here we briefly outline how this is done.

On a generalized eigenspace \( W^{(n)}_\lambda \) for \( \lambda \not\in \mathbb{R} \), in subsections §4.3 and §4.4 we normalize the restriction \( A|_V \) of \( A \) to a subspace \( V \), where \( V \) is defined to be the space spanned by some Jordan chain of \( A^2 \) and the image of \( A \) applied to this Jordan chain, and achieve the normalization by first choosing a basis with respect to which \( A^2 \) has the Jordan normal form and then transforming this basis to a new one with respect to which \( A \) has the form in Theorem 2.2, all the while taking care to simultaneously normalize \( \ell \). The very same procedure can be applied to normalize \( A|_V \) without the additional steps needed to normalize \( \ell \), that is, one can normalize \( A|_V \) by reading through the proofs of propositions (4.12) and (4.13) while disregarding all mention of \( \ell \) (e.g., using the Jordan normal form rather than the Gohberg–Lancaster–Rodman form). Next, letting \( U \) denote the \( A \)-invariant space on which we have already normalized \( A \), we repeat this normalization on any \( A \)-invariant subspace of \( W^{(n)}_\lambda \setminus U \) containing a maximal Jordan chain of \( A^2 \) rather than applying Lemma 4.1 to choose a specific \( A \)-invariant complement of \( U \). To find such a subspace, we choose any maximal length Jordan chain of \( A^2 \) in \( W^{(n)}_\lambda \setminus U \) and consider the subspace spanned by this chain and the image of \( A \) applied to this chain.
On the generalized eigenspace $W^{(n)}_0$, we may normalize the restriction $A|_V$ of $A$ to a subspace $V$, where $V$ is a maximal subspace of $W^{(n)}_0$ that has a basis obtained by applying powers of $A$ to a single vector, by using the procedure in the proof of Proposition 4.11, again disregarding all mention of $\ell$, that is, rather than choosing $a_1 \in \{v \mid A^{k-1}v \neq 0\} \cap W^{(n)}_0$ such that (4.20) holds we simply choose $a_1$ to be an arbitrary vector in $\{v \mid A^{k-1}v \neq 0\} \cap W^{(n)}_0$. We repeat this normalization on any maximal $A$-invariant subspace of $W^{(n)}_0 \setminus U$ (where $U$ denotes the $A$-invariant space on which we have already normalized $A$) that has a basis obtained by applying powers of $A$ to a single vector. To find such a subspace, we choose any vector $v \in W^{(n)}_0 \setminus U$ for which the subspace spanned by $\{v, Av, \ldots, A^n v\}$ has maximal dimension.

Lastly, on a generalized eigenspace $W^{(n)}_{\lambda}$ for which $\lambda^2 > 0$, we apply Lemma 4.8 to normalize the restriction $A|_V$ of $A$ to a subspace $V$, where $V$ is the span of a Jordan chain of $A^2$ given by Lemma 4.8. Note, the proof of Lemma 4.8 does not use the assumption that $A$ is $\ell$-self-adjoint for some Hermitian form $\ell$. And as in the previous two cases, we repeat this normalization on any $A$-invariant subspace of $W^{(n)}_{\lambda} \setminus U$ (where, again, $U$ denotes the $A$-invariant space on which we have already normalized $A$) containing a maximal length Jordan chain of $A^2$.

Given that every antilinear operator can be represented by a matrix representing the antilinear operator of a pair in the canonical form of Theorem 2.2, we have the following lemma.

**Lemma 4.14.** Every antilinear operator on $\mathbb{C}^n$ is $\ell$-self-adjoint with respect to some nondegenerate Hermitian form $\ell$.

**5. Alternative Canonical Forms**

We conclude this text with a few remarks regarding an alternative approach to deriving a canonical form for the pair $(\ell, A)$, and we record an alternative canonical form, Theorem 5.1, that naturally arises from this approach. The form in Theorem 2.2 has some advantages. Its matrices have a minimal number of nonzero entries, for example. The form in Theorem 5.1 is, however, better suited for certain applications. Namely, analysis involving antilinear operators often includes consideration of the operators’ squares, making use of the squares’ linearity and well developed theory for linear operators. The alternative canonical forms of Theorems 5.1 and 5.2 below are ideal for studying $A$ and $A^2$ simultaneously because $A^2$ is represented by a Jordan matrix whenever $A$ is represented by the canonical form of Theorem 5.2.

When searching for a canonical form for $(\ell, A)$, after noticing that a linear operator $A^2$ is $\ell$-self-adjoint whenever the antilinear operator $A$ is $\ell$-self-adjoint, it becomes natural to apply the Gohberg–Lancaster–Rodman form to the pair $(\ell, A^2)$. Specifically, one may try to normalize $(\ell, A)$ by bringing $(\ell, A^2)$ to the Gohberg–Lancaster–Rodman form and then changing the basis to normalize $A$ while tracking the changes induced in the matrix representing $\ell$ (ideally, one would like to achieve this without changing the matrix representing $\ell$ at all). Indeed, we use this approach in subsections §4.3 and §4.4, and, from this perspective, noting Lemma 2.4, one must wonder why we do not use this approach in section §4.1 as well. It turns out to be absolutely viable for the normalization carried out in section §4.1, but the method presented in section §4.1 is simply more efficient. Applying this alternative approach to carry out the normalization has its own merit, however, because it naturally leads one to discover the canonical form given in Theorem 5.1 below.

To explore this further, let us consider the special case wherein $A^2 : \mathbb{C}^n \rightarrow \mathbb{C}^n$ has a single eigenvalue $\lambda^2$, its only eigenspace is 1-dimensional, and $\lambda^2 > 0$ (note, applying Lemmas 4.1 and 4.8, one can always reduce to this special case for the normalization carried out in section §4.1). Applying the Gohberg–Lancaster–Rodman form to the pair $(\ell, A^2)$, we can choose a basis of $\mathbb{C}^n$ with respect to which $\ell$ and $A$ are represented by matrices $S_n$ and $C$ respectively such that

\begin{equation}
C\overline{C} = J_{\lambda^2,n}.
\end{equation}
We attempt to normalize $A$ by changing the basis with transformations that preserve the matrix representations of $\ell$ and $A^2$. Hence we consider the transformations represented by matrices in the group

$$G := \{ M \in M_{n \times n}(\mathbb{C}) | M^* S_n M = S_n \text{ and } M C \overline{C} = \overline{C} M \}$$

acting on the subspace

$$\mathcal{C} := \{ C \in M_{n \times n}(\mathbb{C}) | C C^{-1} = J_{\lambda^2,n} \}$$

of $GL_n(\mathbb{C})$, via the action $(M, C) \mapsto M C^{-1}$. It turns out that we can solve (5.1), that is, we can completely describe the general form of a matrix $C$ satisfying (5.1), and $G$ acts transitively on $\mathcal{C}$. Matrices in (5.1) turn out to be upper-triangular and Toeplitz, and, for a matrix $C \in \mathcal{C}$, one can explicitly construct a matrix $M \in G$ such that $M C M^{-1} \in GL_n(\mathbb{R})$ and the eigenvalue of $M C M^{-1}$ is $|\lambda|$. Choosing $M$ to satisfy these conditions, it turns out that $M C M^{-1}$ equals the matrix $M_{\lambda,n}$ defined below, which confirms that $G$ acts transitively on $\mathcal{C}$. Of course, we have omitted details of the calculations summarized here, but the summary provides an outline of how one can apply the aforementioned alternative approach to the normalization carried out in section §4.1. Furthermore, this summary illustrates how, from one perspective, the alternative canonical form given in Theorem 5.1 below arises naturally.

This alternative form features the sequence

$$c_0(\lambda) := \lambda, \quad c_1(\lambda) := \frac{1}{2\lambda}, \quad \text{and} \quad c_i(\lambda) := \frac{-1}{2\lambda} \sum_{j=1}^{i-1} c_j(\lambda)c_{i-j}(\lambda), \quad (5.2)$$

which arises if we try to solve the matrix equation

$$C^2 = J_{\lambda^2,k} \quad \lambda \neq 0 \quad (5.3)$$

by supposing $C$ has the form

$$C = \sum_{i=1}^{k} c_{i-1}(\lambda)T_k^{i-1} \quad (5.4)$$

and comparing coefficients, interpreting each side of the equation as a degree $k - 1$ polynomial in $T_k$. An interesting observation is that the sequence

$$|c_1(1/2)| = 1, |c_2(1/2)| = 1, |c_3(1/2)| = 2, \ldots$$

is known as the Catalan numbers, $|c_i(1/2)| = \frac{1}{i+1}\binom{2i}{i}$, which play an important role in combinatorics. The identity

$$c_i(\lambda) = (-1)^i(2\lambda)^{-2i} |c_i(1/2)| = \frac{(-1)^{i+1}(2\lambda)^{1-2i}}{i+1} \binom{2i}{i}, \quad \text{valid for all positive integers } i,$$

further illuminates the relationship between $\{c_i(\lambda)\}_{i=1}^{\infty}$ and the Catalan numbers.

---

1The space $\mathcal{C}$ turns out to be homeomorphic to the Cartesian product $S^1 \times \mathbb{R}^{n-1}$ of a circle and Euclidean space with the product topology.

2If $C$ satisfies (5.3) then $C$ satisfies (5.4) for some choice of coefficients $c_i(\lambda)$. Nevertheless, proving this fact is not necessary for understanding the provenance of (5.2), so we introduce (5.4) as though it is not a consequence of (5.3).
For $\lambda \in \mathbb{C}$, we define the $k \times k$ or $2k \times 2k$ matrix $M_{\lambda,k}$ by

$$
M_{\lambda,k} := \begin{cases} 
\sum_{i=1}^{k} c_{i-1}(\lambda)T_{k}^{i-1} & \text{if } \lambda \in \mathbb{R} \setminus \{0\} \\
\frac{1}{2} \begin{pmatrix} 
J_{1,k} & -J_{-1,k} \\
J_{-1,k} & -J_{1,k}
\end{pmatrix} & \text{if } \lambda = 0 \text{ and } k \text{ is even} \\
I_{k} & 0 \\
0 & I_{k}
\end{pmatrix} & \text{if } \lambda = 0 \text{ and } k \text{ is odd} \\
\sum_{i=1}^{k} c_{i-1}(\lambda)T_{k}^{i-1} & \text{otherwise,}
\end{cases}
$$

where 0 denotes a matrix of appropriate size with zero in all entries and, for odd $k$, $M_{0,k}$ is a $k \times k$ matrix. We define corresponding matrices $N_{\lambda,k}$ by

$$
N_{\lambda,k} := \begin{cases} 
S_{k} & \text{if } \lambda \in \mathbb{R} \setminus \{0\} \\
S_{\frac{k}{2}} \oplus (-S_{\frac{k}{2}}) & \text{if } \lambda = 0 \text{ and } k \text{ is even} \\
S_{\frac{k}{2}} \oplus S_{\frac{k}{2}} & \text{if } \lambda = 0 \text{ and } k \text{ is odd} \\
S_{k} \oplus (-S_{k}) & \text{if } \lambda^2 < 0 \\
S_{2k} & \text{otherwise,}
\end{cases}
$$

where $[a]$ denotes the smallest integer not less than $a$ and $[a]$ denotes the largest integer not larger than $a$. For the following theorem, we let $\{\lambda_1, \lambda_2, \ldots, \lambda_\gamma\}$ denote the subset of principle square roots of eigenvalues of $A^2$ enumerated in section 2.

**Theorem 5.1.** The domain of an $\ell$-self-adjoint antilinear operator $A$ can be decomposed into $A$-invariant, pairwise $\ell$-orthogonal subspaces such that there exists a basis with respect to which the restrictions of $\ell$ and $A$ to the decomposition's component subspaces are represented by matrices of the form $\pm N_{\lambda,k}$ and $M_{\lambda,k}$ where $\lambda \in \{\lambda_1, \lambda_2, \ldots, \lambda_\gamma\}$ and $k \in \mathbb{N}$. The corresponding block diagonal matrices representing $\ell$ and $A$ are unique up to a permutation of the blocks on the diagonal.

A canonical form for antilinear operators, described in Remark 2.3 and section §4.5, is given by Hong and Horn in [10, Theorem 3.1]. Since, as is noted in Lemma 4.14, every antilinear operator is $\ell$-self-adjoint with respect to some nondegenerate Hermitian form $\ell$, by applying Theorem 5.1 to the pair $\ell$ and $A$ to get another matrix representation for $A$, we obtain the following alternative canonical form for antilinear operators.

**Theorem 5.2.** The domain of an antilinear operator $A$ can be decomposed into $A$-invariant subspaces such that there exists a basis with respect to which the restrictions of $A$ to the decomposition's component subspaces are represented by matrices of the form $M_{\lambda,k}$ where $\lambda \in \{\lambda_1, \lambda_2, \ldots, \lambda_\gamma\}$ and $k \in \mathbb{N}$. The corresponding block diagonal matrix representing $A$ is unique up to a permutation of the blocks on the diagonal.

**Remark 5.3.** In a basis with respect to which $A$ is represented by a matrix with the above canonical form, $A^2$ is represented by a Jordan matrix. Similarly, if $\ell$ and $A$ are represented by matrices in the canonical form of Theorem 5.1 then the pair $(\ell, A^2)$ is represented by matrices in the Gohberg–Lancaster–Rodman form. Noting this connection together with Lemma 2.4, one can readily show that if $A$ is nonsingular then Theorems 2.2 and 5.1 are indeed equivalent. To show that each of these theorems is a consequence of the other in the more general case where $A$ is singular, it is not too difficult to explicitly construct a basis change of the maximal subspace on which $A$ is nilpotent.
transforming the canonical form in Theorem 2.2 to the form in Theorem 5.1 (and vice versa); for example, considering a change of basis transformation represented by

\[ T := \frac{1}{\sqrt{2}} \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
-1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1
\end{bmatrix}, \]

we have \((T^{-1})^*H_{0,6}T^{-1} = N_{0,6}\) and \(TC_{0,6}T^{-1} = M_{0,6}\), that is, this change of basis transforms a certain matrix representation given by Theorem 2.2 to a matrix representation given by Theorem 5.1.

References

[1] Asano, K. and Nakayama, T. (1938). Über halblineare Transformationen. Math. Ann. 115, 87–114.
[2] Benedetti, R. and Cragnolini, P. (1984). On simultaneous diagonalization of one Hermitian and one symmetric form. Linear Algebra and its Applications 57, 215–226.
[3] Cartan, E. (1922). Sur la géométrie pseudo-conforme des hypersurfaces de l’espace de deux variables complexes. Ann. Mat. Pura Appl. 11, 17–90.
[4] Chern, S. and Moser, J. K. (1974). Real hypersurfaces in complex manifolds. Acta Math. 133, 219–271.
[5] Gantmakher, F. R. (1953). The theory of matrices, vol. 131,. American Mathematical Soc.
[6] Gohberg, I., Lancaster, P. and Rodman, L. (2006). Indefinite linear algebra and applications. Springer Science & Business Media.
[7] Haantjes, J. (1935). Klassification der antilinearen Transformationen. Math. Ann. 112, 98–106.
[8] Hong, Y. (1985). Consimilarity: Theory and Applications. Johns Hopkins Univ.
[9] Hong, Y. (1991). A Canonical Form Under \(\varphi\)-Equivalence. Linear Algebra and its Applications 147, 501–549.
[10] Hong, Y. and Horn, R. A. (1988). A canonical form for matrices under consimilarity. Linear Algebra and its Applications 102, 143–168.
[11] Hong, Y. P., Horn, R. A. and Johnson, C. R. (1986). On the reduction of pairs of Hermitian or symmetric matrices to diagonal form by congruence. Linear Algebra and its Applications 73, 213–226.
[12] Jacobowitz, H. (1990). An introduction to CR structures, vol. 38, of Mathematical Surveys and Monographs. American Mathematical Soc.
[13] Jacobson, N. (1943). The Theory of Rings, vol. II, of Mathematical Surveys and Monographs. American Mathematical Soc.
[14] Porter, C. and Zelenko, I. (2017). Absolute parallelism for 2-nondegenerate CR structures via bigraded Tanaka prolongation. arXiv preprint arXiv:1704.03999.
[15] Takagi, T. (1924). On an Algebraic Problem Related to an Analytic Theorem of Carathéodory and Fdjér and on an Allied Theorem of Landau. Proceedings of the Physico-Mathematical Society of Japan. 3rd Series 6, 130–140.
[16] Tanaka, N. (1962). On the pseudo-conformal geometry of hypersurfaces of the space of \(n\) complex variable. J. Math. Soc. Japan 14, 397–429.
[17] Thompson, R. C. (1991). Pencils of Complex and Real Symmetric and Skew Matrices. Linear Algebra and its Applications 147, 323–371.
DAVID SYKES, DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TEXAS, 77843, USA

E-mail address: dgsykes@math.tamu.edu
URL: http://www.math.tamu.edu/~dgsykes

IGOR ZELENKO, DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TEXAS, 77843, USA

E-mail address: zelenko@math.tamu.edu
URL: http://www.math.tamu.edu/~zelenko