CANONICAL FORMS FOR SYMPLECTIC AND HAMILTONIAN MATRICES

ALAN J. LAUB
University of Minnesota, Minn., U.S.A.

and

KENNETH MEYER*
University of Minnesota, Minn., U.S.A.
and
University of Cincinnati, Dept. of Mathematics, Ohio, U.S.A.

(Received 11 December, 1972)

Abstract. This paper gives a constructive method for finding canonical forms for symplectic and Hamiltonian matrices. No restrictions are made on the eigen values or their multiplicity. Real canonical forms are treated in detail.

1. Introduction

This paper exhibits a constructive method for deriving canonical froms for Hamiltonian and symplectic matrices under conjugation by symplectic matrices. Special care has been taken to treat the case of real canonical forms in detail.

The basic result for a real Hamiltonian or symplectic matrix A is that there exists a real symplectic matrix T such that $T^{-1}AT$ has the form

$$\left(\begin{array}{cccc|cccc} A_{i1} & & & O & A_{i2} & & & O \\ & A_{j1} & & & & A_{j2} & & \\ & & \ddots & & & & \ddots & \\ O & & & \cdot & O & & & \cdot \\ \hline A_{i3} & & & O & A_{i4} & & & O \\ & A_{j3} & & & & A_{j4} & & \\ & & \ddots & & & & \ddots & \\ O & & & \cdot & O & & & \cdot \end{array}\right),$$

where the real submatrix

$$\begin{pmatrix} A_{i1} & A_{i2} \\ A_{i3} & A_{i4} \end{pmatrix}$$

* This research partially supported by NSF grant GP-37620.

is a canonical block analogs to the the usual Jordan blocks. In this block, the matrices A_{il}, $l=1, 2, 3, 4$ are all real square matrices and the particular form of the blocks depends on the eigenvalues of A. The canonical blocks for Hamiltonian matrices are given explicitly in Section 3 (see (1), (2), (3), (4), (6), (7), (9)) while those for symplectic matrices are given in Section 4 (see (10), (11), (12), (13). (14), (15), (16)).

Some fundamental work on this topic was presented by John Williamson in three classic papers (Williamson, 1936, 1937, 1939). He gave necessary and sufficient conditions for two Hamiltonian or two symplectic matrices to be conjugate under conjugation by symplectic matrices in terms of elementary divisors and other invariants. Unfortunately, a constructive procedure for computing the canonical forms was not provided. Furthermore, explicit forms were demonstrated only for very low-order cases, some cases for which the extension to higher orders is not obvious. Williamson's work is co-ordinated and extended by Wall (1963). In more recent papers, either explicit canonical forms are not given (or if they are, they are for only specific low order cases), or real transformations are not considered, or the 'difficult' cases (eigenvalues 0 or $\pm i\nu$ in the Hamiltonian case, eigenvalues on the unit circle in the symplectic case) are not satisfactorily discussed or are avoided (Moser, 1958; Robinson, 1971; Roels and Louterman, 1970; Siegel and Moser, 1971). However, explicit canonical forms were simply not the focus of most of these papers. Burgoyne and Cushman (1971) have recently presented a constructive method for the Hamiltonian case of pure imaginary eigenvalues. But the method presented here is not only constructive but also it handles all of the 'difficult' cases. Moreover, both the Hamiltonian and symplectic cases are, to some extent, handled simultaneously. The method is essentially a special case of a far more general algebraic result of Springer and Steinberg (1970). (See also Springer, 1951.)

Section 2 contains some preliminary lemmas and definitions which will be needed for the construction of the canonical forms. Many of these lemmas are either elementary or well known and so many proofs are omitted. The survey (Robinson, 1971) has many of these results. The main result of this section is Lemma 5 which reduces the problem to finding canonical forms for the linear transformation restricted to generalized eigenspaces.

Section 3 deals with the canonical forms for Hamiltonian matrices on these generalized eigenspaces. Certain cases are, in fact, trivial. The non-trivial cases where the matrix has zero or pure imaginary eigenvalues are treated in detail by using an extension of the symplectic form. This extension of the symplectic form, Ω, was introduced by Springer (1951) and Springer and Steinberg (1970) and greatly simplifies the discussion of these difficult cases.

Section 4 deals with the canonical forms for symplectic matrices on the generalized eigenspaces.

2. Preliminary Lemmas

A *symplectic space* is a pair $\langle V, \omega \rangle$ where V is a $2n$-dimensional vector space over a field F and ω is a nondegenerate alternating bilinear 2-form on V. Since ω is nondegenerate the mapping $\omega^*: V \to V^*: v \to \omega(v, \cdot)$ is an isomorphism. The standard example is $V = F^{2n}$ and $\omega(x, y) = x^T J y$ where J is the $2n \times 2n$ shew symmetric matrix

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

and I_n is the $n \times n$ identity matrix. A *symplectic basis* for V is a basis $v_1, v_2, \cdots, v_{2n}$ such that $\omega(v_i, v_j) = J_{ij}$, where $J = (J_{ij})$. A symplectic space always admits a symplectic basis and so all symplectic spaces are isomorphic to the standard example.

A *symplectic subspace* U of a symplectic space is a subspace such that $\omega \mid U \times U$ is nondegenerate. If U is a symplectic subspace of V then there exists a unique symplectic subspace W such that $V = U \oplus W$ and $\omega(U, W) = 0$; in fact, $W = \{v \in V : \omega(u, v) = 0$ for all $u \in U\}$. Conversely, if $\omega(U, W) = 0$ and $V = U \oplus W$ then both U and W are symplectic subspaces. Such a decomposition will be called a *symplectic decomposition.*

U is a *Lagrange subspace* of V if it is a maximal subspace of V such that $\omega(\mu_1, \mu_2) = 0$ for all $u_1, u_2 \in U$. A Lagrange subspace always has dimension n and there is a second Lagrange subspace Y such that $V = U \oplus Y$ but this decomposition is not unique. Such a decomposition will be called a *Lagrange splitting* of V.

LEMMA 1. *Let $V = U \oplus Y$ be a Lagrange splitting of V. Then if $u_1, \ldots, u_n$ is any basis for U, there exists a unique basis $y_1, \ldots, y_n$ for Y such that $u_1, \cdots, u_n, y_1, \ldots, y_n$* is symplectic basis for V.

Proof. Let $\hat{\omega} = \omega \mid U \times Y$ so that $\hat{\omega}: U \times Y \to F$.

Define $\varphi_i: Y \to F$ (i.e., $\varphi_i \in Y^* =$ dual space of Y) as follows:

$$\varphi_i(\cdot) = \hat{\omega}(u_1, \cdot), \qquad i = 1, \cdots, n.$$

We show that $\varphi_1, \ldots, \varphi_n$ are a basis for Y^*.

Since $\dim Y^* = n$ it suffices to show that $\varphi_1, \ldots, \varphi_n$ are linearly independent (over F).

Suppose

$$\sum_{i=1}^{n} \alpha_i \varphi_i = 0, \qquad \alpha_i \in F.$$

By definition of the φ_i,

$$\sum_{i=1}^{n} \hat{\omega}(a_i u_i, y) = 0, \qquad \forall y \in Y.$$

Now, since U is a Lagrange subspace we have

$$0 = \sum_{i=1}^{n} \hat{\omega}(\alpha_i u_i, y) = \sum_{i=1}^{n} \omega(\alpha_i u_i, u + y), \qquad \forall u \in U, \forall y \in Y,$$

whence

$$0 = \sum_{i=1}^{n} \omega(\alpha_i u_i, v) = \omega\left(\sum_{i=1}^{n} \alpha_i u_i, v\right), \qquad \forall v \in V$$

since $V = U \oplus Y$.

Thus $\sum_{i=1}^{n} \alpha_i u_i = 0$ by the nondegeneracy of ω. Hence $\alpha_1 = \cdots = \alpha_n = 0$ since $u_1, \ldots, u_n$ are a basis for U and so $\varphi_1, \ldots, \varphi_n$ are linearly independent and thus a basis for Y^*. Now let $y_1, \ldots, y_n \in Y$ be the unique dual basis of $\varphi_1, \ldots, \varphi_n$. (Actually, the dual basis of $\varphi_1, \ldots, \varphi_n$ is a basis for Y^{**} but we make the usual identification between elements of Y^{**} and Y since for $\dim Y < \infty$, $Y \approx Y^{**}$ by the ismorphism $\tau: Y \to Y^{**}$ defined by $\tau(y) = T_y$, $\forall y \in Y$ where $T_y \in Y^{**}$, $Ty: Y^* \to F$ is defined by $T_y(\varphi) = \varphi(y)$, $\forall \varphi \in Y^*$).

Then $\varphi_i(y_j) = \delta_{ij}$ or $\hat{\omega}(u_i, y_j) = \delta_{ij} = \omega(u_i, y_j)$ for $i, j = 1, \ldots, n$.

Thus $u_1, \ldots, u_n, y_1, \ldots, y_n$ is a symplectic basis for V.

DEFINITION 1. Let V be a $2n$-dimensional symplectic space and let $A: V \to V$ be a real linear transformation (matrix). Define A, the adjoint of A, by the relation $\omega(Ax, y) = \omega(x, \tilde{A}y)$. Since this is equivalent to the statement that $x^T A^T J y = x^T J \tilde{A} y$, an equivalent definition is $\tilde{A} = -JA^T J$.

Now, the usual definition that A be Hamiltonian is that

$$\omega(Ax, y) + \omega(x, Ay) = 0, \qquad \forall x, y \in V.$$

This is equivalent to

$$\begin{aligned} 0 &= \omega(x, \tilde{A}y) + \omega(x, Ay) \\ &= \omega(x, (\tilde{A} + A)\, y) \end{aligned}$$

whence $\tilde{A} = -A$ by the nondegeneracy of ω.

The usual definition that A be symplectic is that

$$\omega(Ax, Ay) - \omega(x, y) = 0, \qquad \forall x, y \in V.$$

Since A is symplectic, it is nonsingular so let $Ay = v$, or $y = A^{-1}v$.
Then

$$\begin{aligned} 0 &= \omega(Ax, v) - \omega(x, A^{-1}v) \\ &= \omega(x, \tilde{A} - A^{-1})\, v) \end{aligned}$$

whence $\tilde{A} = A^{-1}$ by the nondegeneracy of ω.

DEFINITION 2.

$$\sigma(A) = \begin{cases} -A & \text{if } \ A \text{ is Hamiltonian} \\ A^{-1} & \text{if } \ A \text{ is symplectic}. \end{cases}$$

DEFINITION 3. A is said to be σ-symplectic if $\tilde{A} = \sigma(A)$, i.e., if A is Hamiltonian or symplectic.

Given a symplectic space $\langle V, \omega \rangle$, $\{A : \omega(Ax, Ay) = \omega(x, y),\ \forall x, y \in V\}$ is a Lie group often denoted by $\mathrm{Sp}(V)$, $\mathrm{sp}(V)$, the Lie algebra associated with $\mathrm{Sp}(V)$, can be shown to be

$$\{A : \omega(Ax, y) = -\omega(x, Ay),\ \forall x, y \in V\}.$$

LEMMA 2. *Let A be σ-symplectic. If λ is an eigenvalue of A then $\sigma(\lambda)$ is also an eigenvalue of A.*

Proof. See Robinson (1971).

COROLLARY 2.1. *If λ is an eigenvalue of the σ-symplectic matrix A $(V=\mathbb{R}^{2n})$, then $\sigma(\lambda)$, $\bar{\lambda}$, and $\sigma(\bar{\lambda})$ are also eigenvalues of A with the same multiplicity as λ. In particular, 1 and -1 have even multiplicity if they are eigenvalues of $A\in \mathrm{Sp}(\mathbb{R}^{2n})$ and 0 has even multiplicity if it is an eigenvalue of $A\in \mathrm{sp}(\mathbb{R}^{2n})$.*

Suppose now, unless otherwise stated, that $A: V\to V$ is *σ-symplectic. Let* $\eta_k(\lambda)=\ker(A-\lambda I)^k$, $\eta(\lambda)=\bigcup_k \eta_k(\lambda)$ *(finite union)*

$$\tilde{\eta}_k(\lambda)=\ker(\tilde{A}-\lambda I)^k, \qquad \tilde{\eta}(\lambda)=\bigcup_k \tilde{\eta}_k(\lambda).$$

LEMMA 3. *If $\mu\neq\tilde{\lambda}=\sigma(\lambda)$, then $\omega(\eta_1(\lambda),\eta_1(\mu))=0$.*

Proof. Let $x\in\eta_1(\lambda)$, i.e., $Ax=\lambda x$

and $y\in\eta_1(\mu)$, i.e., $Ay=\mu y$.

As in the proof of Lemma 2, we also have $\sigma(A)\,x=\sigma(\lambda)\,x=\tilde{\lambda}x$.
Now,

$$\begin{aligned}\mu\omega(x,y)&=\omega(x,\mu y)=\omega(x,Ay)=\omega(\sigma(A)\,x,y)\\&=\omega(\lambda x,y)=\lambda\omega(x,y).\end{aligned}$$

Thus $(\mu-\tilde{\lambda})\,\omega(x,y)=0$ so $\mu\neq\tilde{\lambda}$ implies $\omega(x,y)=0$.

COROLLARY 3.1. *Let A have distinct eigenvalues $\lambda_1,\dots,\lambda_n$, $\sigma(\lambda_1),\dots,\sigma(\lambda_n)\in F$ (dim $V=2n$ implies multiplicity of each eigenvalue is 1) with eigenvectors $q_1,\dots,q_n$, $y_1,\dots,y_n$ respectively. There exists a symplectic basis for which A takes the form*

$$\begin{pmatrix}\lambda_1 & & & & & \\ & \ddots & & & \bigcirc & \\ & & \lambda_n & & & \\ & & & \sigma(\lambda_1) & & \\ & \bigcirc & & & \ddots & \\ & & & & & \sigma(\lambda_n)\end{pmatrix}.$$

LEMMA 4. $\eta_k(\mu)=\tilde{\eta}_k(\tilde{\eta})$.

Proof. The proof is by induction on k.

Let $x\in\eta_1(\mu)$, i.e., $Ax=\mu x$.

Then $\sigma(A)\,x=\sigma(\mu)\,x$ or $\tilde{A}x=\tilde{\mu}x$ whence $x\in\tilde{\eta}_1(\tilde{\mu})$.

Thus $\eta_1(\mu) \subseteq \tilde{\eta}_1(\tilde{\mu})$. By symmetry (i.e., using the fact that $\sigma^2 = \text{identity}$) we get $\tilde{\eta}_1(\tilde{\mu}) \subseteq \eta_1(\mu)$ so $\eta_1(\mu) = \tilde{\eta}_1(\tilde{\eta})$.

Induction hypothesis: $\eta_k(\mu) = \tilde{\eta}_k(\tilde{\eta})$.

Let $x \in \eta_{k+1}(\mu)$, i.e., $(A - \mu I)^{k+1} x = 0$. Rewriting this as $(A - \mu I)^k (A - \mu I) x = 0$ we have $(A - \mu I) x \in \eta_k(\mu)$ and so $(A - \mu I) x \in \tilde{\eta}_k(\tilde{\mu})$ by induction hypothesis. Thus

$$0 = (\tilde{A} - \tilde{\mu} I)^k (A - \mu I)\, x = (A - \mu I)\,(\tilde{A} - \tilde{\mu} I)^k\, x$$

since for σ analytic, $\sigma(A)\, A = A \sigma(A)$. (In particular, $AA^{-1} = A^{-1}A$ and $A(-A) = (-A)A$).

Hence $(\tilde{A} - \tilde{\mu} I)^k\, x \in \eta_1(\mu) = \tilde{\eta}_1(\tilde{\mu})$ so

$$0 = (\tilde{A} - \tilde{\mu} I)\,(\tilde{A} - \tilde{\mu} I)^k\, x = (\tilde{A} - \tilde{\mu} I)^{k+1}\, x$$

whence $x \in \tilde{\eta}_{k+1}(\tilde{\mu})$ and we have $\eta_{k+1}(\mu) \subseteq \tilde{\eta}_{k+1}(\tilde{\mu})$.

A symmetrical argument shows the reverse inclusion so $\eta_{k+1}(\mu) = \tilde{\eta}_{k+1}(\tilde{\mu})$ and the lemma is proved.

LEMMA 5. *If* $\mu \neq \tilde{\lambda} = \sigma(\lambda)$, *then* $\omega(\eta(\lambda), \eta(\mu)) = 0$.

Proof. Since $\eta(\lambda) = \bigcup_k \eta_k(\lambda)$ and $\eta(\mu) = \bigcup_k \eta_k(\mu)$, where the unions are finite and $\eta_j \subseteq \eta_{j+1}$, the proof is by induction on k. For $k = 1$, $\eta(\lambda) = \eta_1(\lambda)$ and $\eta(\mu) = \eta_1(\mu)$ so the result follows by Lemma 3.

Induction hypothesis: $\omega(\eta_k(\lambda), \eta_k(\mu)) = 0$.

We show this implies $\omega(\eta_{k+1}(\lambda), \eta_k(\mu)) = 0$ and by reversing the roles of the arguments we show $\omega(\eta_{k+1}(\lambda), \eta_{k+1}(\mu)) = 0$.

Suppose $u \in \eta_{k+1}(\lambda)$, i.e., $(A - \lambda I)^{k+1} u = 0$; $v \in \eta_k(\mu)$, i.e., $(A - \mu I)^k\, v = 0$.

Then

$$\begin{aligned} 0 &= \omega\big(u, (A - \mu I)^k\, v\big) \\ &= \omega\big(u, (A - \tilde{\lambda} I + [\tilde{\lambda} - \mu]\, I)^k\, v\big) \\ &= \sum_{i=0}^{k} \binom{k}{i} (\tilde{\lambda} - \mu)^{k-i}\, \omega\big(u, (A - \tilde{\lambda} I)^i\, v\big) \\ &= \sum_{i=0}^{k} \binom{k}{i} (\tilde{\lambda} - \mu)^{k-i}\, \omega\big((\tilde{A} - \tilde{\lambda} I)^i\, u, v\big). \end{aligned}$$

Now, by Lemma 4, $u \in \eta_{k+1}(\lambda)$ implies $u \in \tilde{\eta}_{k+1}(\tilde{\lambda})$ and thus

$$0 = (\tilde{A} - \tilde{\lambda} I)^{k+1}\, u = (\tilde{A} - \tilde{\lambda} I)^{k+1-i} (\tilde{A} - \tilde{\lambda} I)^i\, u$$

whence $(\tilde{A} - \tilde{\lambda} I)^i\, u \in \tilde{\eta}_{k+1-i}(\tilde{\lambda})$.

Then by the induction hypothesis $\omega((\tilde{A} - \tilde{\lambda} I)^i\, u, v) = 0$, $i = 1, \ldots, k$ since $(\tilde{A} - \tilde{\lambda} I)^i\, u \in \tilde{\eta}_{k+1-i}(\tilde{\lambda}) = \eta_{k+1-i}(\lambda)$, $i = 1, \ldots, k$ and $v \in \eta_k(\mu)$.

Hence from the equations above, we have $0 = (\tilde{\lambda} - \mu)^k\, \omega(u, v)$ which implies, since $\mu \neq \tilde{\lambda}$, that $\omega(u, v) = 0$ and hence that $\omega(\eta_{k+1}(\lambda), \eta_k(\mu)) = 0$.

Similarly, we show $\omega(\eta_{k+1}(\lambda), \eta_{k+1}(\mu)) = 0$.

Therefore, by the induction process, $\omega(\eta(\lambda), \eta(\mu)) = 0$.

Suppose $\langle V, \omega\rangle$ is a $2n$-dimensional symplectic space and that $A: V\to V$ is σ-symplectic. Let A have the distinct eigenvalues $\lambda_1,\ldots, \lambda_k$, $\sigma(\lambda_1),\ldots. \sigma(\lambda_k)$, ϱ_1, ϱ_r where $\sigma(\varrho_i)=\varrho_i$. If A is Hamiltonian, $r=1$ and $\varrho_1=0$ while if A is symplectic, $r=1$ or $r=2$ and $\varrho_1=\pm 1, \varrho_2=\mp 1$.

Now, by the Jordan Canonical Form Theorem,

$$V=\left(\bigoplus_i \eta(\lambda_i)\right)\oplus\left(\bigoplus_i \eta(\sigma(\lambda_i))\right)\oplus\left(\bigoplus_i \eta(\varrho_i)\right),$$

where the direct sums indexed by i are, of course, finite. With this decomposition of V we have the following lemmas:

COROLLARY 5.1. *$\eta(\varrho_j)$ is a symplectic subspace of V.*

COROLLARY 5.2. *If A is symplectic then* $\det A=1$.

Proof. By Corollary 5.1, the eigenvalues $+1$ or -1 of A must have even multiplicity, while for eigenvalues $\lambda\neq\pm 1$, $1/\lambda$ is also an eigenvalue so $\det A=$product of its eigenvalues$=1$.

COROLLARY 5.3. $W=(\oplus_i\eta(\lambda_i))\oplus(\oplus_i\eta(\sigma(\lambda_i)))$ *is a symplectic subspace of V.*

Note. By Lemma 5, $(\oplus_i\eta(\lambda_i))$ and $(\oplus_i\eta(\sigma(\lambda_i)))$ are a Lagrange splitting of W.

3. Canonical Forms for Hamiltonian Matrices

Suppose $\langle V, \omega\rangle$ is a $2n$-dimensional symplectic space, $F=\mathbb{R}$, and $A: V\to V$ is Hamiltonian. Consider the following decomposition of V:

$$\begin{aligned}V=&\left(\bigoplus_j[\eta(\mu_j)\oplus\eta(-\mu_j)]\right)\oplus\\ &\oplus\left(\bigoplus_j[[\eta(\mu_j+iv_j)\oplus\eta(\mu_j-iv_j)]\oplus[\eta(-\mu_j-iv_j)\oplus\eta(-\mu_j+iv_j)]]\right)\\ &\oplus\left(\bigoplus_j[\eta(iv_j)\oplus\eta(-iv_j)]\right)\oplus(\eta(0)),\end{aligned}$$

where μ_j, $v_\alpha>0$ for all j (finite number). Dropping the subscripts, we shall derive canonical forms for each of these four classes of eigenvalues. Note that the third subspace cannot be treated simply as a special case of the second, for while $\eta(iv)$, $\eta(-iv)$ are a Lagrange splitting of the symplectic space $\eta(iv)\oplus\eta(-iv)$, the space is not real. However, $\eta(\mu+iv)\oplus\eta(\mu-iv)$, $\eta(-\mu-iv\oplus\eta(-\mu+iv)$ are a Lagrange splitting of $[\eta(u+iv)\oplus\eta(\mu-iv)]\oplus[\eta(-\mu_-iv)\oplus\eta(-\mu+_iv)]$ and both Lagrange subspaces are simply complexifications of real spaces.

Case 1. Consider $A\mid\eta(\mu)\oplus\eta(-\mu)$

A is invariant on both $\eta(\mu)$ and $\eta(-\mu)$. Choose a basis $q_1,\ldots, q_k$ of $\eta(\mu)$ such that w.r.t. this basis the matrix of $A \mid \eta(\mu)$ is in Jordan canonical form

$$A_1 = \begin{pmatrix} \mu & & & & O \\ \delta_2 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ O & & & \delta_k & \mu \end{pmatrix},$$

where $\delta_i = 0$ or 1.

Then by Lemma 5, $\omega(q_i, q_j) = 0$, $i, j = 1, \ldots, k$. By Lemma 1, we can complete $q_1, \ldots, q_k$ to a symplectic basis $q_1, \ldots, q_k, p_1, \ldots, p_k$ where $p_i \in \eta(-\mu)$.

With respect to this basis, the matrix of A is

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

Since we must have $-JA^T J = -A$, a short computation shows that we must have $A_2 = -A_1^T$. Thus the canonical form is:

$$\left(\begin{array}{ccccc|ccccc} \mu & & & & O & & & & & \\ \delta_2 & \ddots & & & & & & & & \\ & \ddots & \ddots & & & & & O & & \\ & & \ddots & \ddots & & & & & & \\ O & & & \delta_k & \mu & & & & & \\ \hline & & & & & -\mu & -\delta_2 & & & O \\ & & & & & & \ddots & \ddots & & \\ & & O & & & & & \ddots & \ddots & \\ & & & & & & & & \ddots & -\delta_k \\ & & & & & O & & & & -\mu \end{array}\right), \tag{1}$$

Case 2. Consider $A \mid [\eta(\mu+i\nu) \oplus \eta(\mu-i\nu)] \oplus [\eta(-\mu-i\nu) \oplus \eta(-\mu+i\nu)]$

By the same reasoning as in Case 1, we may choose $A \mid [\eta(\mu+i\nu) \oplus \eta(\mu-i\nu)]$ to be in real Jordan canonical form. The desired canonical form is then seen to be:

$$\left(\begin{array}{ccccc|ccccc} B & & & & O & & & & & \\ \Delta_2 & \ddots & & & & & & & & \\ & \ddots & \ddots & & & & & O & & \\ & & \ddots & \ddots & & & & & & \\ O & & & \Delta_k & B & & & & & \\ \hline & & & & & -B & -\Delta_2 & & & O \\ & & & & & & \ddots & \ddots & & \\ & & O & & & & & \ddots & \ddots & \\ & & & & & & & & \ddots & -\Delta_k \\ & & & & & O & & & & -B \end{array}\right), \tag{2}$$

where

$$B = \begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix} \quad \text{and} \quad \Delta_i = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

We shall now consider $A \mid \eta(0)$ before $A \mid \eta(iv) \oplus \eta(-iv)$ in order to better illustrate the constructive procedure of Springer and Steinberg.

Case 3. Consider $A \mid \eta(0)$

For simplicity, let A denote $A \mid \eta(0)$ and let $V = \eta(0)$. Suppose $\langle V, \omega \rangle$ is a symplectic space of dimension $2n$ and that $A: V \to V$ is a linear Hamiltonian transformation which is nilpotent of index $k+1 \leqslant 2n$, i.e., $A^k \neq 0$, $A^{k+1} = 0$ ($A^{2n} = 0$ by Cayley-Hamilton Theorem). Let $F = \mathbb{R}$. By the Jordan Canonical Form Theorem, V has a basis of the form:

$$\begin{aligned} &v_1, Av_1, \ldots, A^{s_1} v_1, \\ &v_2, Av_2, \ldots, A^{s_2} v_2, \\ &\quad\vdots \\ &v_r, Av_r, \ldots, A^{s_r} v_r, \end{aligned}$$

where $A^{s_i+1} v_i = 0$, $i = 1, \ldots, r$.

This, however, is not a symplectic basis.

Let $\mathfrak{A}$ be the commutative algebra generated by A, i.e., $\Phi \in \mathfrak{A}$ iff $\Phi = \alpha_0 I + \alpha_1 A + \cdots + \alpha_k A^k$, $\alpha_i \in F$. ($\Phi = 0$ iff $\alpha_0 = \alpha_1 = \cdots = \alpha_k = 0$). Then $\mathfrak{A} \approx F[x]/(x^{k+1})$. In general, if the minimal polynomial of A is $p^{k+1}(x)$ with $p(x)$ irreducible over F, we consider $\mathfrak{A} \approx \approx F[x]/(p^{k+1}(x))$.

Let $\mathfrak{B}$ denote V considered as a module over $\mathfrak{A}$. We then notice the pleasant fact that $v_1, v_2, \ldots, v_r$ are a basis for $\mathfrak{B}$. This result, along with the *-sesquilinear form $\Omega: V \times V \to \mathfrak{A}$ to be introduced below, will be exploited to constructively derive canonical forms for Hamiltonian matrices all of whose eigenvalues are zero. Let $\Lambda: \mathfrak{A} \to F$ be linear and nonzero on $p^k(x) F[x] \bmod p^{k+1}(x) F[x]$. For example, if $p(x) = x$, $\Phi = \alpha_0 I + \alpha_1 A + \cdots + \alpha_k A^k \in \mathfrak{A}$, $\alpha_i \in F$, then $\Lambda(\Phi) = \alpha_k$.

We now define $\Omega: V \times V \to \mathfrak{A}$ so as to satisfy the equation

$$\omega(\Phi x, y) = \Lambda(\Phi \Omega(x, y)), \qquad \forall x, y \in V, \qquad \forall \Phi \in \mathfrak{A}.$$

An easy computation gives:

$$\Omega(x, y) = \omega(A^k x, y) I + \omega(A^{k-1} x, y) A + \cdots + \omega(x, y) A^k.$$

DEFINITION 4. Let $\Phi = \alpha_0 I + \alpha_1 A + \ldots + \alpha_k A^k \in \mathfrak{A}$. Define $\Phi^* \in \mathfrak{A}$ by

$$\Phi^* = \alpha_0 I - \alpha_1 A + \cdots + (-1)^k \alpha_k A^k.$$

DEFINITION 5. Ω is said to be a *-sesquilinear form if for all $\alpha_1, \alpha_2 \in F$; $\Phi_1, \Phi_2 \in \mathfrak{A}$; and $x_1, x_2, y_1, y_2 \in V$ we have

$$\Omega(\alpha_1 \Phi_1 x_1 + \alpha_2 \Phi_2 x_2, y_1) = \alpha_1 \Phi_1 \Omega(x_1, y_1) + \alpha_2 \Phi_2 \Omega(x_2, y_1)$$

and

$$\Omega(x_1, \alpha_1\Phi_1 y_1 + \alpha_2\Phi_2 y_2) = \alpha_1\Phi_1^*\Omega(x_1, y_1) + \alpha_2\Phi_2^*\Omega(x_1, y_2).$$

LEMMA 6. *Ω is nondegenerate *-sesquilinear form. Furthermore,* $\Omega(x, y) = (-1)^{k+1}\,\Omega(y, x)^*$.

LEMMA 7. $\Phi = \alpha_0 I + \alpha_1 A + \cdots + \alpha_k A^k \in \mathfrak{A}$ *is nonsingular iff* $\alpha_0 \neq 0$.

LEMMA 8. *Suppose* $\Phi = \alpha_0 I + \alpha_1 A + \cdots + \alpha_k A^k \in \mathfrak{A}$ *is nonsingular, i. e.,* $\alpha_0 \neq 0$. *Then there exists a square root Ψ of Φ such that* $\Psi^2 = \operatorname{sgn}(\alpha_0)\,\Phi$.

Furthermore, Ψ is also nonsingular.

Proof. Suppose $\alpha_0 > 0$. Let $\Psi = \beta_0 I + \beta_1 A + \cdots + \beta_k A^k$ and compare coefficients in the equation $\Psi^2 = \Phi$.

We find

$$\alpha_j = \sum_{h=0}^{j} \beta_h \beta_{j-h}, \qquad j = 0, 1, \ldots, k.$$

Clearly, if $\alpha_j \in \mathbb{R}$ and $\alpha_0 > 0$, we can solve recursively for $\beta_j \in \mathbb{R}, j = 0, 1, \ldots, k$. In particular, $\beta_0 = \pm\sqrt{\alpha_0} \neq 0$ so Ψ is also nonsingular.

If $\alpha_0 < 0$, find Ψ such that $\Psi^2 = -\Phi$ as above.

THEOREM 9. *Let V, ω, A be as in the introduction to case 3. Then V has a symplectic decomposition* $V = U_1 \oplus \cdots \oplus U_\alpha \oplus Y_1 \oplus \cdots \oplus Y_\beta$ *into A-invariant subspaces. Further, U_i has a basis* $e_i, Ae_i, \ldots, A^{k_i} e_i$ ($A \mid U_i$ *is nilpotent of index* $k_i + 1 \leqslant k + 1$) *with*

$$\omega(A^s e_i, e_i) = \begin{cases} \pm 1 & \text{if} \quad s = k_i \\ 0 & \textit{otherwise} \end{cases},$$

and Y_j has a basis $f_j, Af_j, \ldots, A^{m_j} f_j, g_j, Ag_j, \ldots, A^{m_j} g_j$ ($A \mid Y$ *is nilpotent of index* $m_j + 1 \leqslant k + 1$) *with*

$$\omega(A^s f_j, g_j) = \begin{cases} 1 & \text{if} \quad s = m_j \\ 0 & \textit{otherwise} \end{cases}$$

and

$$\omega(A^s f_j, f_j) = \omega(A^s g_j, g_j) = 0, \qquad \forall s.$$

LEMMA 10. *Let W be a subspace of V. Let* $A_W = A \mid W$. *Suppose* $A_W : W \to W$ *is a linear Hamiltonian transformation which is nilpotent of index* $k_W + 1$ ($\leqslant k + 1$) *and let* $\mathfrak{A}_W$ *be the commutative algebra generated by* A_W. *Let $\mathfrak{W}$ denote W considered as a vector space over* $\mathfrak{A}_W$. *Define* $\Omega_W : W \times W \to \mathfrak{A}_W$ *by* $\Omega_W = \Omega \mid (W \times W)$. *Let* $\xi_1, \ldots, \xi_\gamma$ *be a basis for $\mathfrak{W}$. Suppose, by relabeling if necessary, that* $\Omega_W(\xi_1, \xi_1)$ *is nonsingular. Then there exists a basis* $e_W, \xi_2', \ldots, \xi_\gamma'$ *for $\mathfrak{W}$ such that* $\Omega_W(e_W, e_W) = \pm I$, $\Omega_W(e_W, \xi_i') = 0$ *and* $\mathfrak{W} = \mathfrak{L}\{e_W\} \oplus \mathfrak{L}\{\xi_2', \ldots, \xi_\gamma'\}$ *where* $\mathfrak{L}\{\ \}$ *means linear span in the module sense and* $\mathfrak{L}\{e_W\}$ *has a basis, as a vector space over* $F = \mathbb{R}$, $e_W, A_W e_W, \ldots, A_W^{k_W} e_W$.

LEMMA 11. *W. $\mathfrak{W}$, A_W, Ω_W as in Lemma* 10 (*except now A_W is nilpotent of index $m_W+1(\leqslant k+1)$*). *Suppose again that $\xi_1,\ldots,\xi_\gamma$ is a basis for $\mathfrak{W}$ but that now $\Omega_W(\xi_i,\xi_i)$ is singular for all $i=1,\ldots,\gamma$ and, by relabeling if necessary, that $\Omega_W(\xi_1,\xi_2)$ is nonsingular. Then there exists a basis $f_W, g_W, \xi_3',\ldots,\xi_\gamma'$ for $\mathfrak{W}$ such that $\Omega_W(f_W,g_W)=I$, $\Omega_W(f_W,f_W)=\Omega_W(g_W,g_W)=0$, $\Omega_W(f_W,\xi_i')=\Omega_W(g_W,\xi_i')=0$ and $\mathfrak{W}=\mathfrak{L}\{f_W,g_W\}\oplus\mathfrak{L}\{\xi_3',\ldots,\xi_\gamma'\}$ where $\mathfrak{L}\{f_W,g_W\}$ has a basis, as a vector space over $F=\mathbb{R}$, $f_W, A_Wf_W,\ldots,A_W^{m_W}f_W, g_W, A_Wg_W,\ldots,A_W^{m_W}g_W$.*

Proof of Theorem 9. Let $\xi_1,\ldots,\xi_\gamma$ be a basis for $\mathfrak{V}$. We first show that $\Omega(\xi_i,\xi_j)$ must be nonsingular for some i and j (possibly equal). Suppose $\Omega(\xi_i,\xi_j)$ were singular for all i and j. By Lemma 7, the coefficient of I must be 0, i.e., $\omega(A^k\xi_i,\xi_j)=0$. Furthermore, $\omega(A^{k+l_1+l}\xi_i,\xi_j)=0$ for all nonnegative integers l_1, l_2 by the nilpotency of A. Now fix l_1. Then $\omega(A^k(A^{l_1}\xi_i),A^{l_2}\xi_j)=0$ $\forall l_2\geqslant 0$. Since $\{A^{l_2}\xi_j: l_2\geqslant 0\}$ forms a basis for V over F, the nondegeneracy of ω implies $A^k(A^{l_1}\xi_i)=0$. But this holds for all $l_1\geqslant 0$ and $\{A^{l_1}\xi_i: l_1\geqslant 0\}$ forms a basis for V over F so we conclude that $A^k=0$ which contradicts the hypothesis of $k+1$ as the index of nilpotency of A.

Having established the existence of a nonsingular $\Omega(\xi_i,\xi_j)$ suppose, by relabeling if necessary, that $\Omega(\xi_1,\xi_1)$ is nonsingular. (If $\Omega(\xi_i,\xi_i)$ is singular for all i, proceed to the use of Lemma 11 below.) Apply Lemma 10 with $W=V$ to get $V=U_1\oplus W_{11}$, with e_1, $Ae_1,\ldots,A^ke_1$ being a basis for U_1 (the numerical subscripts on the basis corresponding to the first application of Lemma 10 and A being written for $A\mid U_1$). If $W_{11}\neq\{0\}$, consider next W_{11}, $A\mid W_{11}$, $\Omega\mid(W_{11}\times W_{11})$, etc. instead of V, A, Ω, etc. The remarks at the beginning of this proof may again be applied and it may again be possible to apply Lemma 10 with $W=W_{11}$ to get $W_{11}=U_2\oplus W_{12}$ and hence $V=U_1\oplus U_2\oplus W_{12}$, with $e_2, Ae_2,\ldots,A^{k_2}e_2$ being a basis for U_2 $(A=A\mid U_2)$. Then consider W_{12}, $A\mid W_{12}$, etc. and so on until finally $V=U_1\oplus\ldots\oplus U_\alpha\oplus W_{1\alpha}$, where, if $\xi_1,\ldots,\xi_\delta$ is a basis for $W_{1\alpha}$, $\Omega(\xi_i,\xi_i)$ is singular for $i=1,\ldots,\delta$ (where Ω means $\Omega\mid(W_{1\alpha}\times W_{1\alpha})$).

If $W_{1\alpha}\neq\{0\}$, the remarks at the beginning of this proof still hold so that we may apply Lemma 11 to $W_{1\alpha}$ to get $V=U_1\oplus\cdots\oplus U_\alpha\oplus Y_1\oplus W_{21}$ with $f_1, Af_1,\ldots,A^{m_1}f_1, g_1, Ag_1,\ldots,A^{m_2}g_1$ being a basis for Y_1 (again numerical subscripts on the basis correspond to the first application of Lemma 11 and A stands for $A\mid Y_1$). Then consider W_{21} to generate a Y_2 and so on. Since $\dim V<\infty$, the process must terminate with V finally being decomposed into $V=U_1\oplus\cdots\oplus U_\alpha\oplus Y_1\oplus\cdots\oplus Y_\beta$. That the decomposition is a symplectic decomposition follows immediately from Lemmas 10 and 11.

Proof of Lemma 10. For notational convenience in the proof we shall drop the 'W' subscripts. Note first that k must be odd for from Lemma 6 we have $\Omega(\xi_1,\xi_1)=$ $=(-1)^{k+1}\Omega(\xi_1,\xi_1)^*$. The coefficient, α_0, of I for $\Omega(\xi_1,\xi_1)$ must be nonzero by Lemma 7 and is invariant under the *-operation so $\alpha_0=(-1)^{k+1}\alpha_0$ whence k is odd. Thus $\Phi=\Omega(\xi_1,\xi_1)=\Omega(\xi_1,\xi_1)^*$ so that Φ must clearly be of the form $\alpha_0I+\alpha_2A^2+$ $+\alpha_4A^4+\cdots+\alpha_{k-1}A^{k-1}$.

Since $\Phi=\Phi^*$ is nonsingular, by Lemma 8 there exists a square root Ψ of Φ such that $\Psi^2=\operatorname{sgn}(\alpha_0)\,\Phi$. Moreover, by examining the proof of Lemma 8, we see that

$\Psi=\Psi^*$ also. Suppose $\operatorname{sgn}(\alpha_0)=+1$. (A precisely analogous proof works when $\operatorname{sgn}(\alpha_0)=-1$). Let e be defined by the equation $\xi_1=\Psi e$ (or $e=\Psi^{-1}\xi_1$). Then

$$\Psi\Psi^* = \Psi^2 = \Phi = \Omega(\xi_1, \xi_1) = \Omega(\Psi e, \Psi e) = \Psi\Psi^*\Omega(e, e).$$

Thus $\Omega(e, e)=I$ (or $\Omega(e, e)=-I$ when $\alpha_0<0$).

Now, for $i=2,\ldots,\gamma$, let $\xi_i'=\xi_i-\Omega(e, \xi_i)^* e$. Then

$$\begin{aligned}\Omega(e, \xi_i') &= \Omega(e, \xi_1) - \Omega(e, \Omega(e, \xi_i)^*e)\\ &= \Omega(e, \xi_i) - \Omega(e, \xi_i\,\Omega(e, e))\\ &= 0.\end{aligned}$$

Since the transformation from the basis $\xi_1,\cdots,\xi_\gamma$ to $e, \xi_2',\ldots,\xi_\gamma'$ is invertible, the latter set of vectors is also a basis for $\mathfrak{W}$ and the lemma is proved.

Proof of Lemma 11. Again, as in the proof of Lemma 10, we shall drop the 'W' subscripts for notational convenience. Also, we may as well assume $\Omega(\xi_1, \xi_2)=I$; if not, simply make a nonsingular change of variables. Now, there are two cases to consider.

Case A: m is even. By Lemma 6, $\Omega(\xi_i, \xi_i)=(-1)^{m+1}\,\Omega(\xi_i, \xi_i)^*=-\Omega(\xi_i, \xi_i)^*$, $\forall i$.

Thus $\Omega(\xi_i, \xi_i)$ must clearly be of the form $\alpha_1 A+\alpha_3 A^3+\alpha_5 A^5+\cdots+\alpha_{m-1}A^{m-1}$ which we shall call type I. Also by Lemma 6, $\Omega(\xi_2, \xi_1)=-I$. Let $f=\xi_1+\Psi\xi_2$. We determine Ψ so that $\Omega(f,f)=0$ and Ψ is of type I, i.e., $\Psi=-\Psi^*$.

$$\begin{aligned}0 = \Omega(f,f) &= \Omega(\xi_1, \xi_1) + \Omega((\xi_1, \Psi\xi_2) + \Omega(\Psi\xi_2, \xi_1) + \Omega(\Psi\xi_2, \Psi\xi_2)\\ &= \Omega(\xi_1, \xi_1) + \Psi^* - \Psi + \Psi\Psi^*\Omega(\xi_2, \xi_2).\end{aligned}$$

If $\Psi=-\Psi^*$, we wish to solve

$$\Psi = \tfrac{1}{2}[\Omega(\xi_1, \xi_1) + \Psi^2\Omega(\xi_2, \xi_2)].$$

Notice that the product of three type I terms is again of type I so that the right hand side of the equation is of type I. Clearly, we may solve recursively for the coefficients of Ψ starting with the coefficient for A.

Case B: m is odd. By Lemma 6, $\Omega(\xi_i, \xi_i)=(-1)^{m+1}\,\Omega(\xi_i, \xi_i)^*=\Omega(\xi_i, \xi_i)^*$, $\forall i$. Thus $\Omega(\xi_i, \xi_i)$ must clearly be of the form $\alpha_2 A^2+\alpha_4 A^4+\cdots+\alpha_{m-1}A^{m-1}$ (recall: $\Omega(\xi_i, \xi_i)$ singular implies $\alpha_0=0$) which we shall call type II. Also by Lemma 6, $\Omega(\xi_2, \xi_1)=I$. Let $f=\xi_1+\Psi\xi_2$. We determine Ψ so that $\Omega(f,f)=0$ and Ψ is of type II, i.e., $\Psi=\Psi^*$.

$$0 = \Omega(f,f) = \Omega(\xi_1, \xi_1) + \Psi^* + \Psi + \Psi\Psi^*\Omega(\xi_2, \xi_2).$$

If $\Psi=\Psi^*$, we wish to solve.

$$\Psi = -\tfrac{1}{2}[\Omega(\xi_1, \xi_1) + \Psi^2\Omega(\xi_2, \xi_2)].$$

Notice that the right hand side of the equation is of type II and we can clearly solve recursively for the coefficients of Ψ starting with the coefficient for A^2.

In either case, $f, \xi_2, \dots, \xi_\gamma$ is a basis and, without loss of generality, we may as well assume that $\Omega(f, \xi_2)=I$ (and $\Omega(\xi_2, f)=\pm I$ according as m is odd or even). Let $h=\xi_2-\tfrac{1}{2}\Omega(\xi_2, \xi_2)f$. Then one can check that $\Omega(h, h)=0$ whether m is even or odd. Let g be defined by the equation $h=\Omega(f, h)^* g$. Then $\Omega(g, g)=0$ and $\Omega(f, h)=$ $=\Omega(f, h)\,\Omega(f, g)$ so $\Omega(f, g)=I$. Finally, noting that

$$\Omega(g, f) = \begin{cases} I & \text{if} \quad m \text{ is odd} \\ -I & \text{if} \quad m \text{ is even}, \end{cases}$$

let $\xi_i'=\xi_i\mp\Omega(g, \xi_i)^* f-\Omega(f, \xi_i)^* g$ (the minus sign being chosen when m is odd, the plus when even). Then one may check that $\Omega(f, \xi_i')=\Omega(g, \xi_i')=0$, $i=2, \dots, \gamma$, and the lemma is proved.

Canonical Blocks for $A \mid U_i$

Note. In this case and in canonical blocks for $A \mid Y_j$ we continue to exclude the proper subspace subscripts on certain variables. Assume $\Omega(e, e)=+I$ and recall k must be odd in this case. Let $l=(k+1)/2$. Since $\Omega(e, e)=\omega(A^k e, e)\,I+\omega(A^{k-1}e, e)\,A+\cdots+$ $+\omega(e, e)\,A^k=I$, we have

$$\omega(A^s e, e) = \delta_{sk} = \begin{cases} 1 & \text{if} \quad s = k \\ 0 & \text{otherwise}. \end{cases}$$

Let $q_j=A^{j-1}e, p_j=(-1)^{k+1-j}A^{k+1-j}e; j=1, \dots, l$. Then

$$\begin{aligned} \omega(q_i, q_j) &= \omega(A^{i-1}e, A^{j-1}e)\,; \qquad i, j=1, \dots, l \\ &= (-1)^{j-1}\,\omega(A^{i+j-2}e, e) \\ &= 0 \quad \text{since} \quad i+j-2 \leqslant k-1 \\ \omega(p_i, p_j) &= \omega((-1)^{k+1-i}A^{k+1-i}e, (-1)^{k+1-j}A^{k+1-j}e)\,; \quad i, j = 1, \dots, l \\ &= 0 \quad \text{since} \quad 2(k+1)-i-j \geqslant k+1 \end{aligned}$$

and

$$\begin{aligned} \omega(q_i, p_j) &= \omega(A^{i-1}e, (-1)^{k+1-j}A^{k+1-j}e); \qquad i, j = 1, \dots, l \\ &= \omega(A^{k+(i-j)}e, e) \\ &= \begin{cases} +1 & \text{if} \quad i=j \\ 0 & \text{if} \quad i \neq j. \end{cases} \end{aligned}$$

Thus $q_1, \dots, q_l, p_1, \dots, p_l$ form a symplectic basis for U_i. With respect to this basis, $A \mid U_i$ has the following canonical form: $((k+1)\times(k+1)$ matrix)

$$\left(\begin{array}{cccc|cccc}
0 & & & & & & & \\
1 & \ddots & & \text{O} & & & & \\
 & \ddots & \ddots & & & \text{O} & & \\
\text{O} & & 1 & 0 & & & & \\
\hline
 & & & 0 & 0 & -1 & & \text{O} \\
 & \text{O} & & \vdots & & \ddots & \ddots & \\
 & & & \vdots & & & \ddots & -1 \\
0 & \cdots & \cdots & (-1)^l & \text{O} & & & 0
\end{array}\right). \tag{3}$$

Note that $Aq_l = A^l e = (-1)^l p_l$.

In case $\Omega(e, e) = -I$, let $q_j = (-1)^{j-1} A^{j-1} e$, $p_j = A^{k+1-j} e$ and check results anaogous to above to find that the canonical form is precisely the negative of the one shown above.

Canonical Blocks for $A \mid Y_j$

Since $\Omega(f, f) = \Omega(g, g) = 0$ and $\Omega(f, g) = I$, by the definition of Ω we have $\omega(A^s f, f) = = \omega(A^s g, g) = 0$, $\forall s$; $\omega(A^s f, g) = \delta_{sm}$.

Let $q_j = A^{j-1} f$, $p_j = (-1)^{m+1-j} A^{m+1-j} g$; $j = 1, \ldots, m+1$.

By the properties above we find

$$\begin{aligned}
\omega(q_i, q_j) &= 0, \qquad i, j = 1, \ldots, m+1 \\
\omega(p_i, p_j) &= 0, \qquad i, j = 1, \ldots, m+1
\end{aligned}$$

and

$$\begin{aligned}
\omega(q_i, p_j) &= \omega(A^{i-1} f, (-1)^{m+1-j} A^{m+1-j} g) \\
&= \omega(A^{m+(i-j)} f, g) \\
&= \delta_{ij}.
\end{aligned}$$

Thus $q_1, \ldots, q_{m+1}, p_1, \ldots, p_{m+1}$ are a symplectic basis for Y_j. (They are clearly independent and span Y_j since $Y_j = \mathfrak{L}\{f, g\}$ implies a typical element is of the form $(\alpha_0 I + \cdots + \alpha_m A^m) f + (\beta_0 I + \ldots + \beta_m A^m) g$ and this is clearly a linear combination of q_j's and p_j's.) With respect to this basis, $A \mid Y_j$ has the following canonical form: $(2(m+1) \times 2(m+1)$ matrix)

$$\left(\begin{array}{cccc|cccc}
0 & & & & & & & \\
1 & \ddots & & \text{O} & & & & \\
 & \ddots & \ddots & & & \text{O} & & \\
\text{O} & & 1 & 0 & & & & \\
\hline
 & & & & 0 & -1 & & \text{O} \\
 & \text{O} & & & & \ddots & \ddots & \\
 & & & & & & \ddots & -1 \\
 & & & & \text{O} & & & 0
\end{array}\right). \tag{4}$$

Case 4. Consider $A \mid \eta(iv) \oplus \eta(-iv)$

Much of this case is analogous to case 3 so only the essential differences will be noted here. Let $V = \eta(iv) \oplus \eta(-iv)$, $F = \mathbb{C}$. Suppose $\langle V, \omega \rangle$ is a symplectic space of dimension $2n$ and that $A: V \to V$ is a linear Hamiltonian transformation. Let $B = (A \mid \eta(iv)) - ivI$, $\bar{B} = (A \mid \eta(-iv)) + ivI$ and suppose B is nilpotent of index $k+1 \leqslant 2n$. (and hence also $\bar{B}$). Let $\mathfrak{A}$ be the commutative algebra generated by B. $\mathfrak{A} \approx F[x]/((x-iv)^{k+1})$.
$\Phi \in \mathfrak{A}$ iff $\Phi = \alpha_0 I + \alpha_1 B + \cdots + \alpha_k B^k$, $\alpha_i \in F = \mathbb{C}$. This time we define $\Omega: \eta(iv) \times \eta(-iv) \to \mathfrak{A}$ as follows:

$$\Omega(x, \bar{y}) = \omega(B^k x, \bar{y})\, I + \omega(B^{k-1} x, \bar{y})\, B + \cdots + \omega(x, \bar{y})\, B^k$$

for all $x \in \eta(iv)$, $\bar{y} \in \eta(-iv)$.

DEFINITION 6. Let $\Phi = \alpha_0 I + \alpha_1 B + \cdots + \alpha_k B^k \in \mathfrak{A}$. Define $\Phi^* \in \mathfrak{A}$ by $\Phi^* = \bar{\alpha}_0 I - \bar{\alpha}_1 B + \cdots + (-1)^k \bar{\alpha}_k B^k$.

LEMMA 12. *Ω is a nondegenerate *-sesquilinear form.*

LEMMA 13. $\Phi = \alpha_0 I + \alpha_1 B + \cdots + \alpha_k B^k \in \mathfrak{A}$ *is nonsingular iff* $\alpha_0 \neq 0$.

THEOREM 14. *Let V, ω, A, and B be as in the introduction to case* 4. *Then V has a symplectic decomposition*

$$V = (U_1 \oplus \bar{U}_1) \oplus \ldots \oplus (U_\alpha \oplus \bar{U}_\alpha) \oplus (Y_1 \oplus \bar{Y}_1) \oplus \cdots \oplus (Y_\beta \oplus \bar{Y}_\beta),$$

where the $U_i \oplus \bar{U}_i$ and $Y_j \oplus \bar{Y}_j$ are A-invariant subspaces (U_i and Y_j are subspaces of $\eta(iv)$). Furthermore, U_i has a basis $e_i, Be_i, \ldots, B^{k_i} e_i$ where *$B \mid U_i$ is nilpotent of index $k_i + 1 \leqslant k+1$, ($\bar{U}_i$ has a basis $\bar{e}_i, \bar{B}\bar{e}_i, \ldots, \bar{B}^{k_i}\bar{e}_i$), and*

$$\omega(B^s e_i, \bar{e}_i) = \begin{cases} \pm 1 & \text{if } \; s = k_i, \quad k_i \text{ odd} \\ \pm i & \text{if } \; s = k_i, \quad k_i \text{ even}, \\ 0 & \text{otherwise} \end{cases}$$

Y_j has a basis $f_j, Bf_j, \ldots, B^{m_j} f_j, g_j, Bg_j, \ldots, B^{m_j} g_j$ where $B \mid Y_j$ is nilpotent of index $m_j + 1 \leqslant k+1$ ($\bar{Y}_j$ has a basis $\bar{f}_j, \ldots, \bar{B}^{m_j}\bar{g}_j$) and $\omega(B^s f_j, \bar{g}_j) = \delta_{sm_j}$, $\omega(B^s f_j, \bar{f}_j) = \omega(B^s g_j, \bar{g}_j) = 0$, for all s.

Remark 1. In terms of certain reality conditions, to be given in detail later, $U_i \oplus \bar{U}_i$ will be shown to have a real symplectic basis in terms of the real and imaginary parts of a basis for U_i. In terms of this real basis, real $2(k_i+1) \times 2(k_i+1)$ canonical blocks for $A \mid U_i \oplus \bar{U}_i$ will be derived. Similarly, real $4(m_j+1) \times 4(m_j+1)$ canonical blocks for $A \mid Y_j \oplus \bar{Y}_j$ will be derived. Thus we shall derive real canonical forms for $A \mid \eta(iv) \oplus \eta(-iv)$. Note that

$$\sum_{i=1}^{\alpha} 2(k_i + 1) + \sum_{j=1}^{\beta} 4(m_j + 1) = 2n$$

with $\max_i k_i$ and/or $\max_j m_j$ equal to k.

Remark 2. In fact $e_1,\dots,e_\alpha, f_1,\dots,f_\beta, g_1,\dots,g_\beta, \bar{e}_1,\dots,\bar{e}_\alpha, \bar{f}_1,\dots,\bar{f}_\beta, \bar{g}_1,\dots,\bar{g}_\beta$ form a basis for $\mathfrak{V}$ (recall $\mathfrak{V}$ denotes V considered as a module over $\mathfrak{A}$) and $U_i\oplus\bar{U}_i = \mathfrak{L}\{e_i, \bar{e}_i\}$, $Y_j\oplus\bar{Y}_j=\mathfrak{L}\{f_j, g_j, \bar{f}_j, \bar{g}_j\}$.

Remark 3. The proof of Theorem 14 will follow essentially the same pattern as that of Theorem 9 and follows from Lemmas 15 and 16. The proofs of these lemmas follow the pattern of Lemmas 10 and 11.

LEMMA 15. *Let $W\oplus\bar{W}$ be a subspace of V (W is a subspace of $\eta(iv)$). Let $A_W = A \mid W\oplus\bar{W}$. Suppose $A_W: W\oplus\bar{W}\to W\oplus\bar{W}$ is a linear Hamiltonian transformation and that $B_W=(A_W \mid W)-ivI$ is nilpotent of index k_W+1 $(\leqslant k+1)$. Let $\mathfrak{A}_W$ be the commutative algebra generated by B_W and let $\mathfrak{W}$ denote $W\oplus\bar{W}$ considered as a module over $\mathfrak{A}_W$. Define $\Omega_W: W\times W\to\mathfrak{A}_W$ by $\Omega_W=\Omega \mid (W\times\bar{W})$. Let $\xi_1,\dots,\xi_\gamma, \bar{\xi}_1,\cdots,\bar{\xi}_\gamma$ be a basis for $\mathfrak{W}$. Suppose, by relabeling if necessary, that $\Omega_W(\xi_1,\bar{\xi}_1)$ is nonsingular. Then there exists a basis $e_W, \xi_2',\dots,\xi_\gamma', \bar{e}_W, \bar{\xi},\dots,\bar{\xi}_{2\gamma}'$ for $\mathfrak{W}$ such that*

$$\Omega_W(e_W,\bar{e}_W)=\begin{cases}\pm I & \text{if } k_W \text{ is odd}\\ \pm iI & \text{if } k_W \text{ is even}\end{cases},\qquad \Omega_W(e_W,\bar{\xi}_i')=0,$$

and $\mathfrak{W}=\mathfrak{L}\{e_W,\bar{e}_W\}\oplus\mathfrak{L}\{\xi_2',\dots,\xi_\gamma',\bar{\xi}_2,\dots,\bar{\xi}_\gamma'\}$ where $\mathfrak{L}\{e_W,\bar{e}_W\}$ has a basis (as a vector space over $F=\mathbb{C}$)

$$e_W, B_We_W,\dots,B_W^{k_W}e_W, \bar{e}_W, \bar{B}_W\bar{e}_W,\dots,\bar{B}_W^{k_W}\bar{e}_W$$

$(\bar{B}_W=(A_W \mid \bar{W})+ivI)$.

LEMMA 16. *Let W, $\mathfrak{W}$, A_W, B_W, Ω_W be as in Lemma 15 (except now B_W is nilpotent of index m_W+1 $(\leqslant k+1)$). Suppose again that $\xi_1,\dots,\xi_\gamma,\bar{\xi}_1,\dots,\bar{\xi}_\gamma$ is a basis for $\mathfrak{W}$ such that $\Omega_W(\xi_i,\bar{\xi}_i)$ is singular for $i=1,\dots,\gamma$ and, by relabeling if necessary, that $\Omega_W(\xi_1,\bar{\xi}_2)$ is nonsingular. Then there exists a basis $f_W, g_W, \xi_3',\dots,\xi_\gamma', \bar{f}_W, \bar{g}_W, \bar{\xi}_3,\dots,\bar{\xi}_\gamma$ for $\mathfrak{W}$ such that $\Omega_W(f_W,\bar{g}_W)=I$, $\Omega_W(f_W,\bar{f}_W)=\Omega_W(g_W,\bar{g}_W)=0$, $\Omega_W(f_W,\bar{\xi}_i')=\Omega_W(g_W,\bar{\xi}_i')=0$ and $\mathfrak{W}=\mathfrak{L}\{f_W,g_W,\bar{f}_W,\bar{g}_W\}\oplus\mathfrak{L}\{\xi_3',\dots,\xi_\gamma',\bar{\xi}_3',\dots,\bar{\xi}_\gamma'\}$ where $\mathfrak{L}\{f_W,g_W,\bar{f}_W,\bar{g}_W\}$ has a basis (as a vector space over F)*

$$f_W, B_Wf_W,\dots,B_W^{m_W}f_W, g_W, B_Wg_W,\dots,B_W^{m_W}g_W,$$
$$\bar{f}_W,\dots,\bar{B}_W^{m_W}\bar{f}_W, \bar{g}_W,\dots,\bar{B}_W^{m_W}\bar{g}_W.$$

Canonical Blocks for $A \mid U_i\oplus\bar{U}_i$

Note. In this case and in canonical blocks for $A \mid Y_j\oplus\bar{Y}_j$ we continue to exclude the proper subspace subscripts on certain variables in the interests of convenience.

Case A: k even. We have demonstrated above the existence of a vector e such that $\Omega(e,\bar{e})=iI$. (The case of $-iI$ is analogous.) By definition of Ω we have $\omega(B^se,\bar{e})=i\delta_{sk}$.

Let $u_j=B^{j-1}e$, $v_j=(-1)^{k-j+1}\,i\bar{B}^{k-j+1}\bar{e}$; $j=1,\dots,k+1$. Then for $i,j=1,\dots,k+1$, $k+1$, $\omega(u_i,u_j)=\omega(v_i,v_j)=0$ since U_i and $\bar{U}_i$ are a Lagrange splitting, while $\omega(u_i, v_j)=0$, $i\neq j$ and $\omega(u_j,v_j)=-i\omega(B^ke,\bar{e})=1$, $j=1,\dots,k+1$. With respect to this basis for $U_i\oplus\bar{U}_i$, one may quickly check that the matrix of $A \mid U_i\oplus\bar{U}_i$ is given by the $2(k+1)\times 2(k+1)$ complex canonical form:

$$\left(\begin{array}{cccc|cccc}
iv & & & & & & & \\
1 & \ddots & & \bigcirc & & & \bigcirc & \\
 & \ddots & \ddots & & & & & \\
\bigcirc & & 1 & iv & & & & \\
\hline
 & & & & -iv & -1 & & \bigcirc \\
 & \bigcirc & & & & \ddots & \ddots & \\
 & & & & & & \ddots & -1 \\
 & & & & \bigcirc & & & -iv
\end{array}\right). \tag{5}$$

With the reality conditions $u_j=(-1)^{k-j+2}\, i\bar{v}_{k-j+2}$, $j=1,\dots,k+1$, consider the following real basis for $U_i \oplus \bar{U}_i$:

$$\left.\begin{aligned}
q_j &= \sqrt{2}\,\mathrm{Re}\,u_j = \frac{1}{\sqrt{2}}\left(u_j + (-1)^{j+1}\, iv_{k-j+2}\right) \\
p_j &= (-1)^j \sqrt{2}\,\mathrm{Im}\,u_j = \frac{1}{\sqrt{2}\,i}\left((-1)^j\, u_{k-j+2} + iv_j\right)
\end{aligned}\right\} j = 1,\dots,k+1\,.$$

It is clear, almost by inspection, that $\omega(q_i, q_j)=\omega(p_i, p_j)=0$, $i,j=1,\dots,k+1$ and $\omega(q_i, p_j)=0$, $i\neq j$ while

$$\begin{aligned}
\omega(q_j, p_j) &= \frac{1}{2i}\left[i\omega(u_j, v_j) + i(-1)^{j+1}(-1)^j\, \omega(v_{k-j+2}, u_{k-j+2})\right] \\
&= 1\,, \qquad j = 1,\dots,k+1\,.
\end{aligned}$$

Thus $q_1,\dots,q_{k+1}, p_1,\dots,p_{k+1}$ are a symplectic basis and one may check that the real canonical form for $A \mid U_i \oplus \bar{U}_i$ is the $(2k+2)\times(2k+2)$ matrix:

$$\left(\begin{array}{cccc|cccc}
0 & & & & & & & -v \\
1 & \ddots & & \bigcirc & \bigcirc & & v & \\
 & \ddots & \ddots & & & \cdot^{\cdot^{\cdot}} & & \\
 & & \ddots & \ddots & -v & & & \\
\bigcirc & & & & v & & & \bigcirc \\
 & & 1 & 0 & -v & & & \\
\hline
 & & & v & 0 & -1 & & \\
\bigcirc & & -v & & & \ddots & \ddots & \bigcirc \\
 & \cdot^{\cdot^{\cdot}} & & & & & \ddots & \ddots \\
 & v & & & & & & \\
-v & & & \bigcirc & \bigcirc & & & -1 \\
v & & & & & & & 0
\end{array}\right). \tag{6}$$

Case B: k is odd. We have demonstrated above the existence of a vector e such that $\Omega(e, \bar{e})=I$. (The case of $-I$ is analogous.) By definition of Ω, we have $\omega(B^s e, \bar{e})=\delta_{sk}$.

Let $u_j=B^{j-1}e$, $v_j=(-1)^j \bar{B}^{k-j+1}\bar{e}$; $j=1,\ldots,k+1$. Then $\omega(u_i, u_j)=\omega(v_i, v_j)=0$ for $i.j=1,\ldots,k+1$ since U_i and $\bar{U}_i$ are a Lagrange splitting, while $\omega(u_i, v_j)=0$, $i\neq j$, and

$$\begin{aligned}\omega(u_j, v_j) &= (-1)^j\,\omega(B^{j-1}e, \bar{B}^{k-j+1}\bar{e})\\ &= (-1)^{k+1}\,\omega(B^k e, \bar{e})=1\,.\end{aligned}$$

With respect to this basis for $U_i\oplus\bar{U}_i$, A again has the complex canonical form (5).

With the reality conditions $u_j=(-1)^{k-j+2}\bar{v}_{k-j+2}$, $j=1,\ldots,k+1$, consider the following real basis for $U_i\oplus\bar{U}_i$:

$$q_j=\begin{Bmatrix}\sqrt{2}\,\mathrm{Re}\,u_j\\ \sqrt{2}\,\mathrm{Im}\,u_j\end{Bmatrix}=\begin{cases}\dfrac{1}{\sqrt{2}}(u_j+v_{k-j+2}); & j \text{ odd}\\[2ex] \dfrac{1}{\sqrt{2}\,i}(u_j+v_{k-j+2}); & j \text{ even}\end{cases}$$

$$p_j=\begin{Bmatrix}-\sqrt{2}\,\mathrm{Re}\,u_{k-j+2}\\ \sqrt{2}\,\mathrm{Im}\,u_{k-j+2}\end{Bmatrix}=\begin{cases}\dfrac{1}{\sqrt{2}}(-u_{k-j+2}+v_j); & j \text{ odd}\\[2ex] \dfrac{1}{\sqrt{2}i}(u_{k-j+2}-v_j); & j \text{ even}\end{cases}$$

It is clear that $\omega(q_i, q_j)=\omega(p_i, p_j)=0$, $i, j=1,\ldots,k+1$ and $\omega(q_i, p_j)=0$, $i\neq j$ while $j=1,\ldots,k+1$.

$$\begin{aligned}\omega(q_j, p_j) &= \begin{cases}\tfrac{1}{2}\omega(u_j+v_{k-j+2}, -u_{k-j+2}+v_j); & j \text{ odd}\\ -\tfrac{1}{2}\omega(u_j+v_{k-j+2}, u_{k-j+2}-v_j); & j \text{ even}\end{cases}\\ &=1\,, \qquad j=1,\ldots,k+1\,.\end{aligned}$$

Thus $q_1,\ldots,q_{k+1}, p_1,\ldots,p_{k+1}$ are a symplectic basis and one may check that the real canonical form for $A \mid U_i\oplus\bar{U}_i$ is the $(2k+2)\times(2k+2)$ matrix:

$$\left(\begin{array}{c|c} \bigcirc & \begin{matrix} \bigcirc & & & \nu \\ & & \nu+1 & -1 \\ & \cdots & \cdots & \\ \nu-1 & +1 & & \\ \nu+1 & & & \bigcirc \end{matrix} \\ \hline \begin{matrix} \bigcirc & & & -1-\nu \\ & & & +1-\nu \\ & \cdots & \cdots & \\ & -1 & & \\ & +1 & & \\ -1-\nu & & & \\ -\nu & & \bigcirc & \end{matrix} & \bigcirc \end{array}\right). \tag{7}$$

Canonical Blocks for $A \mid Y_j \oplus \bar{Y}_j$

Since $\Omega(f, \bar{f}) = \Omega(g, \bar{g}) = 0$ and $\Omega(f, \bar{g}) = I$, by the definition of Ω we have $\omega(B^s f, \bar{f}) = \omega(B^s g, \bar{g}) = 0$, $\forall s$ and $\omega(B^s f, \bar{g}) = \delta_{sm}$.

Let

$$u_j = \begin{cases} B^{j-1} f; & j = 1, \ldots, m+1 \\ (-1)^{j-m-1} B^{j-m-2} g; & j = m+2, \ldots, 2m+2 \end{cases}$$

and

$$v_j = \begin{cases} (-1)^{m-j+1} \bar{B}^{m-j+1} \bar{g}; & j = 1, \ldots, m+1 \\ \bar{B}^{2m-j+2} \bar{f}; & j = m+2, \ldots, 2m+2. \end{cases}$$

Now, $\omega(u_i, u_j) = \omega(v_i, v_j) = 0$, $i, j = 1, \ldots, 2m+2$ since Y_j, $\bar{Y}_j$ are a Lagrange splitting while $\omega(u_i, v_j) = 0$, $i \neq j$ and for $j = 1, \ldots, m+1$;

$$\begin{aligned} \omega(u_j, v_j) &= \omega(B^{j-1} f, (-1)^{m-j+1} \bar{B}^{m-j+1} \bar{g}) \\ &= \omega(B^m f, \bar{g}) = 1 \end{aligned}$$

and for $j = m+2, \ldots, 2m+2$,

$$\begin{aligned} \omega(u_j, v_j) &= \omega((-1)^{j-m-1} B^{j-m-2} g, \bar{B}^{2m-j+2} \bar{f}) \\ &= (-1)^{j-m} \omega(\bar{B}^{2m-j+2} \bar{f}, B^{j-m-2} g) \\ &= (-1)^{j-m} (-1)^{j-m-2} \omega(\bar{B}^m \bar{f}, g) = 1, \end{aligned}$$

where, in the last equality, we have used the fact that

$$1 = \bar{1} = \overline{\omega(B^m f, \bar{g})} = \omega(\bar{B}^m \bar{f}, g).$$

With respect to this basis for $Y_j \oplus \bar{Y}_j$, one may quickly check that the matrix of $A \mid Y_j \oplus \bar{Y}_j$ is given by the $4(m+1) \times 4(m+1)$ complex canonical form:

$$\left(\begin{array}{cc|cc} \begin{matrix} iv & & O \\ 1 & \ddots & \\ O & 1 & iv \end{matrix} & O & & \\ \hline O & \begin{matrix} iv & & O \\ -1 & \ddots & \\ O & -1 & iv \end{matrix} & \multicolumn{2}{c}{O} \\ \hline & & \begin{matrix} -iv & -1 & O \\ & \ddots & -1 \\ O & & -iv \end{matrix} & O \\ \multicolumn{2}{c|}{O} & O & \begin{matrix} -iv & 1 & O \\ & \ddots & 1 \\ O & & -iv \end{matrix} \end{array}\right). \tag{8}$$

With the reality conditions

$$u_j = \begin{cases} \bar{v}_{2m-j+3}; & j = 1, \ldots, m+1 \\ -\bar{v}_{2m-j+3}; & j = m+2, \ldots, 2m+2 \end{cases}$$

consider the following real basis for $Y_j \oplus \bar{Y}_j$:

$$q_j = \begin{Bmatrix} \sqrt{2}\,\mathrm{Re}\, u_j \\ -\sqrt{2}\,\mathrm{Im}\, u_{2m-j+3} \end{Bmatrix} = \begin{cases} \dfrac{1}{\sqrt{2}}(u_j + v_{2m-j+3}); & j = 1, \ldots, m+1 \\ \dfrac{1}{\sqrt{2}\, i}(v_j - u_{2m-j+3}); & j = m+2, \ldots, 2m+2 \end{cases}$$

$$p_j = \begin{Bmatrix} \pm\sqrt{2}\,\mathrm{Re}\, u_{2m-j+3} \\ \pm\sqrt{2}\,\mathrm{Im}\, u_j \end{Bmatrix} = \begin{cases} \dfrac{1}{\sqrt{2}}(-u_{2m-j+3} + v_j); & j = 1, \ldots, m+1 \\ \dfrac{1}{\sqrt{2}\, i}(u_j + v_{2m-j+3}); & j = m+2, \ldots, 2m+2. \end{cases}$$

Clearly $\omega(q_i, q_j) = \omega(p_i, p_j) = 0$, $i, j = 1, \ldots, 2m+2$, while $\omega(q_i, p_j) = 0$, $i \neq j$ and $\omega(q_j, p_j) = 1$, $j = 1, \ldots, 2m+2$. Thus $q_1, \ldots, q2m+2$, $p_1, \ldots, p_{2m+2}$ are a symplectic basis and one may check that the real canonical form for $A \mid Y_j \oplus \bar{Y}_j$ is the $(4m+4) \times (4m+4)$ matrix:

$$\left(\begin{array}{c|c}
\begin{array}{c|c}
\begin{matrix} 0 & & & O \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ O & & 1 & 0 \end{matrix} &
\begin{matrix} O & & & -v \\ & & ⋰ & \\ & -v & & \\ -v & & & O \end{matrix} \\ \hline
\begin{matrix} O & & & v \\ & & ⋰ & \\ & v & & \\ v & & & O \end{matrix} &
\begin{matrix} 0 & 1 & & O \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ O & & & 0 \end{matrix}
\end{array} & \huge O \\ \hline
\huge O &
\begin{array}{c|c}
\begin{matrix} 0 & -1 & & O \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ O & & & 0 \end{matrix} &
\begin{matrix} O & & & -v \\ & & ⋰ & \\ & -v & & \\ -v & & & O \end{matrix} \\ \hline
\begin{matrix} O & & & v \\ & & ⋰ & \\ & v & & \\ v & & & O \end{matrix} &
\begin{matrix} 0 & & & O \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ O & & -1 & 0 \end{matrix}
\end{array}
\end{array}\right). \tag{9}$$

4. Canonical Forms for Symplectic Matrices

Suppose $\langle V, \omega\rangle$ is a $2n$-dimensional sympletic space, $F=\mathbb{R}$, and $A: V\to V$ is symplectic. Consider the following decomposition of V:

$$V=\Big(\bigoplus_j [\eta(\mu_j)\oplus\eta(\mu_j^{-1})]\Big)\oplus$$

$$\oplus\Big(\bigoplus_j [[\eta(\mu_j+iv_j)\oplus\eta(\mu_j-iv_j)]\oplus[\eta((\mu_j+iv_j)^{-1})\oplus\eta((\mu_j-iv_j)^{-1})]]\Big)\oplus$$

$$\oplus\Big(\bigoplus_j [\eta(\mu_j+iv_j)\oplus\eta(\mu_j-iv_j)]\Big)\oplus(\eta(1))\oplus(\eta(-1)),$$

where $\mu_j\neq 0$ (except $\mu_j\neq 0$ in the first class and μ_j could be zero in the third class), $v_j\neq 0$ for all j and where $\mu_j^2+v_j^2=1$ in the third class of eigenvalues. Dropping the subscripts, we shall derive canonical forms for each of these four classes of eigenvalues ($+1$ and -1 are treated in exactly the same say so we do not distinguish these classes).

Case 1. Consider $A \mid \eta(\mu)\oplus\eta(\mu^{-1})$

A is invariant on both $\eta(\mu)$ and $\eta(\mu^{-1})$. Choose a basis $q_1,\ldots, q_k$ of $\eta(\mu)$ such that w.r.t. this basis the matrix of $A \mid \eta(\mu)$ is in Jordan canonical form

$$A=\begin{pmatrix} \mu & & & & \text{O} \\ \delta_2 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ \text{O} & & & \delta_k & \mu \end{pmatrix},$$

where $\delta_i=0$ or 1.

Then by Lemma 5, $\omega(q_i, q_j)=0$, $i, j,=1,\ldots, k$. By Lemma 1 we can complete $q_1,\ldots, q_k$ to a symplectic basis $q_1,\ldots, q_k, p_1,\ldots, p_k$ where $p_i\in\eta(\mu^{-1})$. With respect to this basis, the matrix of A is

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

where a short computation shows that $A_2=(A_1^{-1})^T$ (since $-JA^TJ=A^{-1}$). Thus the canonical form is composed of canonical blocks (combined as in Section 1) of the form:

$$\left(\begin{array}{cccc|cccc} \mu & & & & & & & \\ 1 & \ddots & & O & & & O & \\ & \ddots & \ddots & & & & & \\ O & & 1 & \mu & & & & \\ \hline & & & & \mu^{-1} & -\mu^{-2} & \cdots & (-1)^{k+1}\mu^{-k} \\ & O & & & & \ddots & \ddots & \vdots \\ & & & & & & \ddots & -\mu^{-2} \\ & & & & & O & & \mu^{-1} \end{array}\right) \tag{10}$$

Case 2. Consider $A \mid [\eta(\mu+i\nu)\oplus\eta(\mu-i\nu)]\oplus[\eta((\mu+i\nu)^{-1})\oplus\eta((\mu-i\nu)^{-1})]$

By the same reasoning as in case 1, we may choose $A \mid [\eta(\mu+i\nu)\oplus\eta(\mu-i\nu)]$ to be in real Jordan canonical form.

$$A=\begin{pmatrix} B & & & O \\ \Delta_2 & \ddots & & \\ & \ddots & \ddots & \\ O & & \Delta_k & B \end{pmatrix} \quad \text{where} \quad B=\begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix} \quad \text{and} \quad \Delta_i=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then the desired canonical form will be composed of canonical blocks of the form:

$$\left(\begin{array}{cccc|cccc} B & & & & & & & \\ I & \ddots & & O & & & & \\ & \ddots & \ddots & & & & O & \\ O & & I & B & & & & \\ \hline & & & & B^{-1} & -B^{-2} & \cdots & (-1)^{k+1}B^{-k} \\ & & & & & \ddots & \ddots & \vdots \\ & O & & & & & \ddots & -B^{-2} \\ & & & & O & & & B^{-1} \end{array}\right), \tag{11}$$

where

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and B^{-j} means $(B^{-1})^j$.

Remark 4. We now depart from the techniques of Section 3 and, in the interests of brevity, use (3), (4), (6), (7), and (9) to develop canonical forms for $A \mid \eta(\mu+iv) \oplus \oplus \eta(\mu-iv)$, $\mu^2+v^2=1$, and for $A \mid \eta(1)$. (The case of $A \mid \eta(-1)$ is completely analogous so will not be discussed). We shall first consider:

Case 3. $A \mid \eta(1)$. (Let A denote $A \mid \eta(1)$ below)

A is a symplectic matrix all of whose eigenvalues are 1.

Consider the following transformation:

$$A \rightarrow (A-I)(A+I)^{-1} = H$$

or the inverse transformation:

$$H \rightarrow (I+H)(I-H)^{-1} = A$$

LEMMA 17. *A is symplectic iff H is Hamiltonian.*

Proof. Suppose A is symplectic, i.e., $-JA^TJ=A^{-1}$.
Then

$$\begin{aligned}
-JH^TJ &= J^{-1}(A^T+I)^{-1}JJ^{-1}(A^T-I)J \\
&= (J^{-1}A^TJ+I)^{-1}(J^{-1}A^TJ-I) \\
&= (A^{-1}+I)^{-1}(A^{-1}-I) \\
&= (A^{-1}(I+A))^{-1}A^{-1}(I-A) \\
&= (I-A)(A+I)^{-1} \\
&= -H \qquad \text{so } H \text{ is Hamiltonian.}
\end{aligned}$$

Conversely, suppose H is Hamiltonian, i.e., $JH^TJ=H$. Then

$$\begin{aligned}
-JA^TJ &= (I-J^{-1}H^TJ)^{-1}(I+J^{-1}H^TJ) \\
&= (I-H(I+H)^{-1} \\
&= A^{-1} \qquad \text{so } A \text{ is symplectic.}
\end{aligned}$$

Remark 5. If all eigenvalues of A are 1, then all eigenvalues of H are 0 and conversely. This fact together with Remark 6 will enable us to derive canonical forms for $A \mid \eta(1)$.

Remark. 6. If the transformation P puts H in canonical form, then P also puts A in some canonical form (and conversely). This follows since

$$P^{-1}AP=(I+P^{-1}HP)(I-P^{-1}HP)^{-1}$$

and

$$P^{-1}HP=(P^{-1}AP-I)(P^{-1}AP+I)^{-1}.$$

So suppose A is a symplectic matrix all of whose eigenvalues are 1. Compute $H=(A-I)(A+I)^{-1}$ (If A is symplectic with all eigenvalues -1, compute $(A+I)(A-I)^{-1}$). Then H is a Hamiltonian matrix all of whose eigenvalues are zero so by the results of Section 3, there exists a real transformation T such that $T^{-1}HT$ takes one of the canonical forms (3) or (4), say $T^{-1}HT=G$.

Then by Remark 6, T also puts A in some canonical form, namely $(I+G)(I-G)^{-1}$ since

$$\begin{aligned}
T^{-1}AT &= T^{-1}(I+TGT^{-1})(I-TGT^{-1})^{-1}T \\
&= (I+G)(I-G)^{-1}.
\end{aligned}$$

Thus, corresponding to the real canonical form (3) for Hamiltonian matrices, we have the real canonical form: $((k+1)\times(k+1)$ matrix$)$

$$\left(\begin{array}{cccc|ccccc}
1 & & & O & & & & & \\
2 & \ddots & & & & & O & & \\
\vdots & \ddots & \ddots & & & & & & \\
2 & \cdots & 2 & 1 & & & & & \\ \hline
-2 & -2 & \cdots & -2 & 1 & -2 & +2 & \cdots & (-1)^{l+1}2 \\
2 & 2 & \cdots & 2 & & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & \vdots & & & \ddots & \ddots & +2 \\
\vdots & & & \vdots & & O & & \ddots & -2 \\
(-1)^{l}2 & \cdots & & (-1)^{l}2 & & & & & 1
\end{array}\right), \tag{12}$$

where $l=(k+1)/2$ (recall k must be odd in this case), while corresponding to (4) we have the following real canonical form: ($2(m+1)\times 2(m+1)$ matrix).

$$\left(\begin{array}{cccc|ccccc}
1 & & & O & & & & & \\
2 & \ddots & & & & & O & & \\
\vdots & \ddots & \ddots & & & & & & \\
2 & \cdots & 2 & 1 & & & & & \\ \hline
 & & & & 1 & -2 & +2 & \cdots & (-1)^{m}2 \\
 & & & & & \ddots & \ddots & \ddots & \vdots \\
 & O & & & & & \ddots & \ddots & +2 \\
 & & & & & O & & \ddots & -2 \\
 & & & & & & & & 1
\end{array}\right). \tag{13}$$

Case 4. Consider $A \mid \eta(\mu+iv) \oplus \eta(\mu-iv)$ where $\mu^2+v^2=1$, $\mu\neq 0$, $v>0$

By precisely the same reasoning as in Case 3, one can compute $(I+G)(I-G)^{-1}$, where G is one of (6), (7), or (9), to get the real canonical forms (14), (15), and (16) but since they are somewhat complicated, they are not presented here. For example, if

$$G = \left(\begin{array}{cc|cc} 0 & 0 & 0 & v \\ 0 & 0 & v & 1 \\ \hline -1 & -v & 0 & 0 \\ -v & 0 & 0 & 0 \end{array}\right),$$

then

$$(I+G)(I-G)^{-1} = \begin{pmatrix} \varphi(v) & 0 & 0 & \psi(v) \\ \varphi'(v) & \varphi(v) & \psi(v) & \psi'(v) \\ -\psi'(v) & -\psi(v) & \varphi(v) & \varphi'(v) \\ -\psi(v) & 0 & 0 & \varphi(v) \end{pmatrix} = S,$$

where $\varphi(v)=(1-v^2)/(1+v^2)$ and $\Psi(v)=2v/(1+v^2)$.

Note that S has eigenvalues $\varphi(v)\pm i\Psi(v)$ of multiplicity 2 and that $\varphi^2(v)+\Psi^2(v)= =1$.

References

Burgoyne, N. and Cushman, R.: 1971, Normal Forms for Real Linear Hamiltonian Systems with Purely Imaginary Eigenvalues', Dept. of Mathematics, University of California at Santa Cruz.

Moser, J.: 1958, *Comm. Pure Appl. Math.* **11**, 81–114.

Robinson, R. C.: 1971, 'Lectures on Hamiltonian Systems', Instituto de Matematica Pura e Aplicada, Guanabara.

Roels, J. and Louterman, G.: 1970: *Celes. Mech.* **3**, 129–140.

Siegel, C. L. and Moser, J.: 1971, 'Lectures on Celestial Mechanics,' in *Die Grundlehren der Mathematische Wissenschaften in Einzeldarstellungen*, Band 187, Springer-Verlag, New York, p. 97.

Springer, T. A.: 1951, Thesis, University of Leiden.

Springer, T. A. and Steinberg, R.: 1970, 'Conjugacy Classes, Part E. of Seminar on Algebraic Groups and Related Finite Groups', in *Lecture Notes in Mathematics*, **131**, Springer-Verlag, Berlin-Heidelberg-New York, 1970.

Wall, G. E.: 1963, *J. Austr. Math. Soc.* **3**, 1–62.

Williamson, J.: 1936, *Am. J. Math.* **58**, 141–163.

Williams, J.: 1937, *Am. J. Math.* **59**, 599–617.

Williams, J.: 1939, *Am. J. Math.* **61**, 897–911.