Normal forms for linear Hamiltonian vector fields commuting with the action of a compact Lie group *

Ian Melbourne†
Department of Mathematics, University of Houston
Houston, Texas 77204-3476, USA

Michael Dellnitz‡
Institut für Angewandte Mathematik
Universität Hamburg
D-2000 Hamburg 13, Germany

Abstract

We obtain normal forms for infinitesimally symplectic matrices (or linear Hamiltonian vector fields) that commute with the symplectic action of a compact Lie group of symmetries. In doing so we extend Williamson's theorem on normal forms when there is no symmetry present.

Using standard representation-theoretic results the symmetry can be factored out and we reduce to finding normal forms over a real division ring. There are three real division rings consisting of the real, complex and quaternionic numbers. Of these, only the real case is covered in Williamson's original work.

*Appeared: *Math. Proc. Camb. Phil. Soc.* **114** (1993) 235-268

†Supported in part by NSF Grant DMS-9101836 and by the Texas Advanced Research Program (003652037)

‡Supported in part by the Deutsche Forschungsgemeinschaft and by NSF Grant DMS-9101836

1 Introduction

In this paper, we obtain a list of normal forms for linear Hamiltonian vector fields that commute with the action of a compact Lie group of symmetries. This is an equivariant generalization of work of Williamson [17] who gave a list of normal forms in the absence of symmetry. Linear Hamiltonian vector fields can be viewed as infinitesimally symplectic matrices, and the Williamson normal form theorem fills the role for infinitesimally symplectic matrices that the Jordan normal form theorem fills for general square matrices.

Galin's work [4] on versal deformations of infinitesimally symplectic matrices depends heavily on Williamson's Theorem. Similarly, an equivariant version of Williamson's Theorem is a necessary first step (indeed the main step) in an equivariant generalization of Galin's results. Such a generalization is currently in progress [11] and renders many of the ad hoc methods in [3] unnecessary.

We note that there is a vast literature on Williamson's and Galin's Theorems in the absence of symmetry, see [1], [2], [8], [9] and [13]. Also see [14] and [15] when there is a time-reversal symmetry.

Let us describe in more detail the problem addressed in [17]. Suppose that $\mathbb{R}^{2n}$ is equipped with the standard inner product $<\,,\,>$ and define the nonsingular skew-symmetric matrix $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$. Then J induces the *canonical* symplectic form ω on $\mathbb{R}^{2n}$ by $\omega(x,y) =< x, Jy >$ for $x, y \in \mathbb{R}^{2n}$. The Lie group of *symplectic* transformations $\mathbf{Sp}_{2n}$ consists of those linear mappings $P : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ that preserve the symplectic structure: $\omega(Px, Py) = \omega(x, y)$. Equivalently $PJP^T = J$.

The Lie algebra of $\mathbf{Sp}_{2n}$ is denoted by $\mathbf{sp}_{2n}$ and consists of *infinitesimally symplectic* transformations. These satisfy the condition $\omega(Ax, y) + \omega(x, Ay) = 0$, equivalently $AJ + JA^T = 0$. The group $\mathbf{Sp}_{2n}$ acts on $\mathbf{sp}_{2n}$ by similarity transformations. (In more sophisticated language, this is the adjoint action of the Lie group $\mathbf{Sp}_{2n}$ on its Lie algebra $\mathbf{sp}_{2n}$.) The list of normal forms in Williamson [17] has the property that each matrix in $A \in \mathbf{sp}_{2n}$ can be transformed by a symplectic change of coordinates to precisely one of these normal forms.

An elementary result in linear algebra states that any nonsingular skew-symmetric matrix R can be transformed into J by an orthogonal change of coordinates. It turns out to be convenient to enlarge the problem as follows. Let $\mathbf{sk}_{2n}$ denote the set of nonsingular skew-symmetric $2n \times 2n$ matrices. If

$R \in \mathbf{sk}_{2n}$, define $\mathbf{sp}_{2n}(R)$ to be the set of $2n \times 2n$ matrices M satisfying $MR + RM^T = 0$. If P is nonsingular then $PRP^T \in \mathbf{sk}_{2n}$ and $PMP^{-1} \in \mathbf{sp}_{2n}(PRP^T)$ so that we obtain an equivalence relation on *symplectic pairs* (M, R) that restricts to the original equivalence relation on $\mathbf{sp}_{2n}$.

Williamson begins by computing normal forms for symplectic pairs. Then if (M, R) is in normal form, R can be transformed into J by an orthogonal change of coordinates. Simultaneously transforming M we obtain a matrix $A \in \mathbf{sp}_{2n}$. This provides a method for converting normal forms (M, R) for symplectic pairs into normal forms A for infinitesimally symplectic matrices. For many theoretical purposes, for example in the work of Galin [4] and the equivariant analog, it is actually more convenient to work with (M, R) rather than A. This is because the coordinates are nicer, R having been chosen to complement the structure of M.

The setup described above can be generalized to the equivariant context. Let $\Gamma \subset \mathbf{O}(2n)$ be a compact Lie group acting orthogonally on $\mathbb{R}^{2n}$. We assume that the action is symplectic, that is $\Gamma \subset \mathbf{Sp}_{2n}$. It follows that J is Γ-equivariant (that is J commutes with the action of Γ) and that ω is invariant under the action of Γ. Define $\mathbf{Sp}_\Gamma$ and $\mathbf{sp}_\Gamma$ to consist of the equivariant matrices in $\mathbf{Sp}_{2n}$ and $\mathbf{sp}_{2n}$. Then $\mathbf{Sp}_\Gamma$ acts on $\mathbf{sp}_\Gamma$ via similarity transformations.

Again it is convenient to enlarge the problem. Begin with an orthogonal action of Γ on a finite-dimensional vector space V and let $\mathbf{sk}_\Gamma$ denote the set of Γ-equivariant nonsingular skew-symmetric matrices. If $R \in \mathbf{sk}_\Gamma$ define $\mathbf{sp}_\Gamma(R)$ to be the space of equivariant matrices M satisfying $MR + RM^T = 0$. Then a Γ-symplectic pair is a pair of matrices (M, R) such that $R \in \mathbf{sk}_\Gamma$ and $M \in \mathbf{sp}_\Gamma(R)$.

The task of finding normal forms for Γ-symplectic pairs can be reduced by applying standard results from representation theory. The action of Γ is factored out and it suffices to obtain (nonequivariant) results over each of the real division rings $\mathcal{D} = \mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ (the reals, complexes and quaternions). Of course, the case $\mathcal{D} = \mathbb{R}$ is solved in [17]. Now Williamson's results were formulated over an arbitrary commutative ring, so it may seem that they should incorporate the case $\mathcal{D} = \mathbb{C}$. In fact they do not and this is related to the fact that a complex matrix can be regarded as a real matrix with twice as many rows and columns. This distinction is not important in Jordan normal form theory but is fundamental here since the changes of coordinates involve the transposition of matrices. The real transpose of a complex matrix is the conjugate transpose (or adjoint). However in [17] the transpose without conjugation is taken regardless of the underlying ring.

The heart of the paper deals with normal forms for Γ-symplectic pairs. The main theorem, Theorem 2.4, is stated in Section 2 and proved in Sections 3, 4 and 5. Section 3 contains the main ideas of the proof, all of which were present in Williamson's original paper (the first of these, which we have already described, is to enlarge the problem). Although the proof relies only on elementary methods in linear algebra, the argument is quite lengthy and is not easy to follow. We have attempted to present a simpler exposition of the proof which divides naturally into three steps. The first two steps taken alone are trivial, but judicious choices lead to a convenient framework in which to carry out the third step. We present an abstract version of Williamson's Theorem within an axiomatic framework (Section 4) and then fit the choices made in the first two steps into this framework (Section 5). It turns out that there is only one case in which this cannot be done. The exceptional case occurs when there are zero eigenvalues and $\mathcal{D} = \mathbb{R}$.

We began with the intention of finding normal forms for matrices in $\mathbf{sp}_\Gamma$ that are infinitesimally symplectic with respect to the canonical symplectic structure induced by J on $\mathbb{R}^{2n}$. We return to this in Section 8 but there are some issues that must be addressed prior to this.

In enlarging the problem from one concerning infinitesimally Γ-symplectic matrices to one concerning Γ-symplectic pairs, we have shifted the emphasis considerably. We began with the canonical symplectic form on $\mathbb{R}^{2n}$ (induced by J) and considered symmetry groups $\Gamma \subset \mathbf{Sp}_{2n}$. Now we have a group Γ acting on V and we take into consideration all Γ-invariant symplectic forms (induced by the matrices in $\mathbf{sk}_\Gamma$). If Γ acts trivially on V, then $\mathbf{sk}_\Gamma$ is nonempty if and only if V is even-dimensional in which case all matrices in $\mathbf{sk}_\Gamma$ can be transformed into the matrix J. The situation is more complicated in general: usually $\mathbf{sk}_\Gamma$ will not contain J even if the set is nonempty. Moreover the matrices in $\mathbf{sk}_\Gamma$ need not be related to each other under equivariant changes of coordinates.

In Section 6 we describe for any representation of Γ a set of *canonical* Γ-symplectic forms having the property that any Γ- invariant symplectic form can be transformed into precisely one of these canonical forms. This result, originally stated in [12], can be recovered as an immediate consequence of Theorem 2.4. The nonuniqueness of the canonical Γ-symplectic forms referred to above stems from the complex representations (where $\mathcal{D} = \mathbb{C}$). It seems that this fact was first appreciated by Montaldi, Roberts and Stewart [12]. In particular, the equivariant version of Darboux's Theorem as cited in [7] is incorrect.

It might be argued that we should consider only those groups $G \subset \mathbf{Sp}_{2n}$, namely those that preserve the symplectic form induced by J. In applications, it is usually the case that we start with such a representation. However, the 'nonstandard' representations often crop up in the analysis of such problems, see for example [12]. Some of these representations have the property of *cyclospectrality* which has important implications for the linear stability of relative equilibria. The cyclospectral representations were defined and classified in [12]. In Section 7 we recover these results, again as an easy consequence of Theorem 2.4. In addition, we define and classify the *weakly cyclospectral* representations which have implications for the existence of Liapunov centers. The information obtained from the cyclospectrality (weak or otherwise) is independent of the precise structure of the Hamiltonian and depends only on its invariance under Γ.

Finally, in Section 8, we return to the original problem of obtaining normal forms for infinitesimally Γ-symplectic matrices. (An infinitesimally Γ-symplectic matrix is an equivariant matrix A that satisfies $AK + KA^T = 0$ where K induces on V a canonical Γ-symplectic form as defined in Section 6.)

The remainder of this section is divided into two subsections. Subsection 1.1 is mainly notational, and in Subsection 1.2 we recall the Jordan normal form theorem over a real division ring. (This is of course standard except over the quaternions.)

1.1 Notation

In this section we review the notation that will be used in this paper. The notation in (a) and (b) below is not standard but is used heavily in the statement and proof of our theorems. In contrast, the notation in (c) is standard but is used only in the proof of the theorems.

(a) Definition of I_k, N_k, T_k and X_k

We define four $k \times k$ real matrices. Let I_k denote the identity matrix, and N_k the standard nilpotent matrix of order k,

$$N_k = \begin{pmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}.$$

The matrices T_k and X_k are given by

$$T_k = \begin{pmatrix} & & & 1 \\ & 0 & 1 & \\ & \cdot^{\cdot^{\cdot}} & & 0 \\ 1 & & & \end{pmatrix}, X_k = \begin{pmatrix} & 0 & & -1 \\ & & 1 & \\ & \cdot^{\cdot^{\cdot}} & & 0 \\ (-1)^k & & & \end{pmatrix}.$$

Proposition 1.1

(i) *The matrices I_k, N_k, $N_k^2, \dots, N_k^{k-1}$ form a linearly independent set.*

(ii) $T_k^T = T_k^{-1} = T_k$.

(iii) $X_k^T = X_k^{-1} = (-1)^{k-1} X_k$.

(iv) $N_k T_k = T_k N_k^T, \quad N_k X_k = -X_k N_k^T$.

(v) *If $k \geq \ell$ and A is an $\ell \times \ell$ matrix, then*

$$T_k \begin{pmatrix} 0 \\ A \end{pmatrix} = \begin{pmatrix} T_\ell A \\ 0 \end{pmatrix}, \quad X_k \begin{pmatrix} 0 \\ A \end{pmatrix} = \begin{pmatrix} X_\ell A \\ 0 \end{pmatrix}.$$

(b) Matrices over a real division ring

Throughout this paper we shall work with matrices over a real division ring. Let $\mathcal{D}$ denote one of the real division rings $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. Let $d = \dim_{\mathbb{R}} \mathcal{D} = 1, 2$ or 4. Suppose that α is a $p \times p$ matrix over $\mathcal{D}$. Then there are standard homomorphisms that identify α with a $dp \times dp$ real matrix A. Any matrix A that can be obtained in this way is said to be real, real-complex or real-quaternionic respectively.

Often we shall use this identification in reverse, so that a real matrix which just happens to have the structure of a real-complex or real-quaternionic matrix can be written as a matrix over $\mathbb{C}$ or $\mathbb{H}$. The number of entries is thus reduced by a factor of four or sixteen.

Of course, if $\mathcal{D}=\mathbb{R}$, then we just let $A=\alpha$. Suppose that $\mathcal{D}=\mathbb{C}$. Then $\alpha=\{\alpha_{rs}\}_{1\le r,s\le p}$ where $\alpha_{rs}=a_{rs}+ib_{rs}\in\mathbb{C}$. Let $A_{rs}=\begin{pmatrix} a_{rs} & -b_{rs} \\ b_{rs} & a_{rs}\end{pmatrix}$. Then define $\alpha_{\mathbb{C}}=A$ where A is the $2p\times 2p$ matrix obtained from α replacing α_{rs} by A_{rs}.

Finally suppose that $\mathcal{D}=\mathbb{H}$. Then $\alpha=\{\alpha_{rs}\}_{1\le r,s\le p}$ where $\alpha_{rs}=a_{rs}+ib_{rs}+jc_{rs}+kd_{rs}\in\mathbb{H}$. Let $A_{rs}=\begin{pmatrix} a_{rs} & -b_{rs} & c_{rs} & -d_{rs} \\ b_{rs} & a_{rs} & d_{rs} & c_{rs} \\ -c_{rs} & -d_{rs} & a_{rs} & b_{rs} \\ d_{rs} & -c_{rs} & -b_{rs} & a_{rs}\end{pmatrix}$. Then define $\alpha_{\mathbb{H}}=A$ where A is the $4p\times 4p$ matrix obtained from α replacing α_{rs} by A_{rs}.

For example,

$$(a+ib+jc+kd)_{\mathbb{H}}=\begin{pmatrix} a+ib & c+id \\ -c+id & a-ib\end{pmatrix}_{\mathbb{C}}=\begin{pmatrix} a & -b & c & -d \\ b & a & d & c \\ -c & -d & a & b \\ d & -c & -b & a\end{pmatrix}.$$

Notice that the matrix $1_{\mathcal{D}}$ corresponds to three different matrices depending on $\mathcal{D}$. That is

$$1_{\mathbb{R}}=1,\quad 1_{\mathbb{C}}=I_2,\quad 1_{\mathbb{H}}=(I_2)_{\mathbb{C}}=I_4.$$

On the other hand, The matrix $i_{\mathcal{D}}$ is undefined if $\mathcal{D}=\mathbb{R}$ and

$$i_{\mathbb{C}}=\begin{pmatrix} 0 & -1 \\ 1 & 0\end{pmatrix},\quad i_{\mathbb{H}}=\begin{pmatrix} i & 0 \\ 0 & -i\end{pmatrix}_{\mathbb{C}}=\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0\end{pmatrix}.$$

The *conjugate* of an element of $\mathcal{D}$ is defined in the usual way: if $a\in\mathbb{R}$, $\bar{a}=a$, if $\alpha=a+ib\in\mathbb{C}$, $\bar{\alpha}=a-ib$, and if $\alpha=a+ib+jc+kd\in\mathbb{H}$, $\bar{\alpha}=a-ib-jc-kd$. Then if $\alpha=\{\alpha_{rs}\}_{1\le r,s\le p}$, $\alpha_{rs}\in\mathcal{D}$ we define the transpose matrix α^T,

$$\alpha^T=\{\bar{\alpha}_{rs}\}_{1\le r,s\le p}.$$

Notice that if $\alpha_{\mathcal{D}}=A$ then $\alpha^T_{\mathcal{D}}=A^T$.

(c) Tensor products of matrices

Suppose that A is a $k \times k$ matrix and α is a $p \times p$ matrix. The tensor product of the matrices is defined to be the $kp \times kp$ matrix $A \otimes \alpha$ formed by replacing each of the entries of A by the product of that entry with α. Thus if $A = \{a_{ij}\}_{1 \le i,j \le k}$ then $A \otimes \alpha = \{a_{ij}\alpha\}_{1 \le i,j \le k}$.

Proposition 1.2 *Suppose that A, B are $k \times k$ matrices and α, β are $p \times p$ matrices. Then*

(a) $(A \otimes \alpha)(B \otimes \beta) = (AB) \otimes (\alpha\beta)$.

(b) $(A \otimes \alpha)^T = A^T \otimes \alpha^T$.

(c) If A and α are nonsingular, then $A \otimes \alpha$ is nonsingular and $(A \otimes \alpha)^{-1} = A^{-1} \otimes \alpha^{-1}$.

(d) $(A \otimes \alpha) \oplus (B \otimes \alpha) = (A \oplus B) \otimes \alpha$.

(e) $(A \otimes \alpha) \otimes \beta = A \otimes (\alpha \otimes \beta)$.

(f) $\alpha \otimes A$ is similar to $A \otimes \alpha$ by an orthogonal change of coordinates.

The orthogonal transformation in part (f) consists of rearrangement of rows and columns in the order

$$\begin{array}{l} 1, p+1, 2p+1, \dots, (k-1)p+1 \\ 2, p+2, 2p+2, \dots, (k-1)p+2 \\ \quad\vdots \\ p, 2p, 3p, \dots, kp \end{array}$$

1.2 Jordan normal forms over real division rings

We recall the Jordan normal form theorems over $\mathbb{R}, \mathbb{C}$ and $\mathbb{H}$. The real and (real)-complex cases are standard. A reference for the real-quaternionic case is Wiegmann [16].

Theorem 1.3 *(a) Any real matrix is similar to a direct sum of matrices of the form*

$$\mu I_k + N_k,\ \mu \in \mathbb{R},$$

and

$$(\mu I_k + N_k)_{\mathbb{C}},\ \mu = \alpha + i\beta,\ \alpha \in \mathbb{R},\ \beta > 0.$$

(b) Any real-complex matrix is similar to a direct sum of matrices of the form

$$(\mu I_k + N_k)_{\mathbb{C}},\ \mu = \alpha + i\beta,\ \alpha, \beta \in \mathbb{R}.$$

(c) Any real-quaternionic matrix is similar to a direct sum of matrices of the form

$$(\mu I_k + N_k)_{\mathbb{H}},\ \mu = \alpha + i\beta,\ \alpha \in \mathbb{R},\ \beta \geq 0.$$

In each case, the direct sum is unique up to permutation of summands.

Remark 1.4 In the real and real-quaternionic cases, it is possible to assume that in each summand μ has nonnegative imaginary part. This is not true in the real-complex case and has important implications for the results in [11]. It turns out that this is the reason why the complex irreducible case is different from the others in [3].

2 The Equivariant Williamson Theorem

Let Γ be a compact Lie group acting on V. We define $\mathrm{Hom}_\Gamma(V)$ to be the vector space of Γ-equivariant real matrices

$$\mathrm{Hom}_\Gamma(V) = \{L : V \to V \text{ linear};\ L\gamma = \gamma L \text{ for all } \gamma \in \Gamma\}.$$

Let $\mathbf{sk}_\Gamma$ be the set of nonsingular skew-symmetric Γ-equivariant matrices. Then if $R \in \mathbf{sk}_\Gamma$, define $\mathbf{sp}_\Gamma(R)$ to consist of those Γ-equivariant matrices M satisfying $MR + RM^T = 0$. We call such pairs of matrices (M, R) Γ-symplectic pairs.

We define an equivalence relation on Γ-symplectic pairs based on the following observation. Suppose that (M, R) is a Γ-symplectic pair, and $P \in \mathrm{Hom}_\Gamma(V)$ is nonsingular. Then it is easily verified that (PMP^{-1}, PRP^T) is a Γ-symplectic pair. Hence we define two pairs to be equivalent, $(M, R) \sim (M', R')$, if there exists a nonsingular matrix $P \in \mathrm{Hom}_\Gamma(V)$ such that

$$PMP^{-1} = M', \quad PRP^T = R'.$$

The equivariant Williamson theorem gives a list of normal forms for Γ-symplectic pairs under this equivalence relation.

We divide this section into two parts. In Subsection 2.1 we show that real matrices that commute with a compact Lie group action can be decomposed into the direct sum of matrices whose entries lie in a real division ring, thus reducing the problem. We make use of basic results from real representation theory, see for example [6]. Then in Subsection 2.2 we state Williamson's theorem over each real division ring. (There are only three nonisomorphic real division rings: the reals, complexes and quaternions.)

2.1 Equivariant matrices

A subspace U is said to be Γ-*irreducible* if it is invariant under Γ and has no proper invariant subspaces. It is known that if U is an irreducible subspace, then $\mathrm{Hom}_\Gamma(U)$ is a real division ring and hence is isomorphic to $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$.

The space V may be written as a direct sum of irreducible subspaces

$$V = U_1 \oplus \cdots \oplus U_k.$$

Group together those U_i on which Γ acts isomorphically to obtain the *isotypic decomposition*

$$V = W_1 \oplus \cdots \oplus W_\ell,$$

where each *isotypic component* W_j is the sum of isomorphic irreducible subspaces. The isotypic decomposition is unique, and moreover each isotypic component is left invariant by matrices in $\mathrm{Hom}_\Gamma(V)$. It follows that

$$\mathrm{Hom}_\Gamma(V) = \mathrm{Hom}_\Gamma(W_1) \oplus \cdots \oplus \mathrm{Hom}_\Gamma(W_\ell). \tag{2.1}$$

Next suppose that W is an isotypic component. We may write $W = U \oplus \cdots \oplus U = \bigoplus_{i=1}^m U$ where U is irreducible and $\mathrm{Hom}_\Gamma(U) \cong \mathcal{D} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. Let $A \in \mathrm{Hom}_\Gamma(W)$. Then $A = \{A_{jk}\}_{1 \le j,k \le m}$ where $A_{jk} : U \to U$. It is easy to check that $A_{jk} \in \mathrm{Hom}_\Gamma(U)$. Since $\mathrm{Hom}_\Gamma(U) \cong \mathcal{D}$ we have shown that

$$\mathrm{Hom}_\Gamma(W) \cong \mathrm{Hom}(\mathcal{D}^m), \tag{2.2}$$

where $\mathrm{Hom}(\mathcal{D}^m)$ denote the space of $m \times m$ matrices with entries in $\mathcal{D}$. Often it will be convenient to denote the isotypic component W by $\mathcal{D}^m$. We say that an isotypic component $\mathcal{D}^m$ is *real*, *complex* or *quaternionic* depending on $\mathcal{D}$. Also we define the *dimension* of the isotypic component $\mathcal{D}^m$ to be the integer m. Note that the dimension of the corresponding (real) subspace W is a multiple of m but is in general not equal to m.

Proposition 2.1 *The space of* Γ*-equivariant matrices has the direct sum decomposition*

$$\mathrm{Hom}_\Gamma(V) \cong \mathrm{Hom}(\mathcal{D}_1^{m_1}) \oplus \cdots \oplus \mathrm{Hom}(\mathcal{D}_\ell^{m_\ell}),$$

where for each $j = 1, \ldots, \ell$, $\mathcal{D}_j = \mathbb{R}, \mathbb{C}$ *or* $\mathbb{H}$.

Proof This follows immediately from equations (2.1) and (2.2). ■

Corollary 2.2 *Suppose that* (M, R) *is a* Γ*-symplectic pair. Then* (M, R) *is a direct sum with summands* (M_j, R_j) *where* $M_j, R_j \in \mathrm{Hom}(\mathcal{D}_j^{m_j})$*, with* R_j *nonsingular, and*

$$R_j = -R_j^T, \quad M_j R_j + R_j M_j^T = 0.$$

Moreover $(M, R) \sim (M', R')$ *if and only if for each* j *there exists a nonsingular matrix* $P_j \in \mathrm{Hom}(\mathcal{D}_j^{m_j})$ *satisfying*

$$P_j M_j P_j^{-1} = M_j', \quad P_j R_j P_j^T = R_j'.$$

Note that in this corollary, transposition is interpreted as applied to the corresponding real matrix (as in Subsection 1.1(b)). Alternatively, take the conjugate transpose of the matrix with entries in $\mathcal{D}$.

In light of the corollary we make some more definitions. Let $\mathbf{sk}_{\mathcal{D}^m}$ consist of nonsingular skew-symmetric $m \times m$ matrices with entries in $\mathcal{D}$. If $R \in \mathbf{sk}_{\mathcal{D}^m}$, we define $\mathbf{sp}_{\mathcal{D}^m}(R)$ to consist of those $m \times m$ matrices M with entries in $\mathcal{D}$ satisfying $MR + RM^T = 0$. A symplectic pair over $\mathcal{D}$ is a pair (M, R) where $R \in \mathbf{sk}_{\mathcal{D}^m}$ and $M \in \mathbf{sp}_{\mathcal{D}^m}(R)$. Two such pairs (M, R) and (M', R') are equivalent, $(M, R) \sim (M', R')$, if there is a nonsingular matrix $P \in \mathrm{Hom}(\mathcal{D}^m)$ such that $PMP^{-1} = M'$, $PRP^T = R'$. Then the corollary can be rephrased as the following.

Corollary 2.3 *Suppose that* (M, R) *is a* Γ*-symplectic pair. Then* (M, R) *is a direct sum with summands* (M_j, R_j) *where* (M_j, R_j) *is a symplectic pair over* $\mathcal{D}_j$*. Moreover* $(M, R) \sim (M', R')$ *if and only if* $(M_j, R_j) \sim (M_j', R_j')$ *for each* j*.*

Hence, the problem of finding normal forms for real Γ-symplectic pairs reduces to one of finding normal forms for symplectic pairs over a real division ring.

2.2 Statement of the theorem

In this subsection we shall regard $\mathcal{D}$ and m as fixed and write $\mathbf{sk}$ instead of $\mathbf{sk}_{\mathcal{D}^m}$. If $R \in \mathbf{sk}$, we define $\mathbf{sp}(R)$ to consist of those $m \times m$ matrices M with entries in $\mathcal{D}$ satisfying $MR + RM^T = 0$. We call such a pair of matrices (M, R) a symplectic pair.

Theorem 2.4 (Equivariant Williamson Theorem) *Suppose that (M, R) is a symplectic pair. Then (M, R) is equivalent to a direct sum of summands where the summands are taken from Table 1 if $\mathcal{D} \cong \mathbb{R}$, from Table 2 if $\mathcal{D} \cong \mathbb{C}$, and from Table 3 if $\mathcal{D} \cong \mathbb{H}$. The direct sum is unique up to order of summands.*

We refer to the summands as *normal form summands* and direct sums of normal form summands as *normal forms*. In the following discussion we shall make more explicit the uniqueness part of the theorem.

First observe that if (M, R) and (M', R') are equivalent, then M and M' are similar by a change of coordinates with entries in $\mathcal{D}$. Now the eigenvalues of each normal form summand occur in a quadruplet $\{\pm\mu, \pm\bar{\mu}\}$ and it follows that the eigenvalues (and Jordan blocks) of $M \in \mathbf{sp}(R)$ occur in quadruplets. In fact this is easy to verify directly, see Lemma 5.1. A less trivial consequence (and one for which we do not know a direct proof) is that if $\mathcal{D} = \mathbb{R}$, odd-dimensional Jordan blocks with zero eigenvalues occur with even multiplicity.

	M	R	Size	μ	ρ
1	$\begin{pmatrix} \mu I_k + N_k & 0 \\ 0 & -\bar{\mu} I_k - N_k \end{pmatrix}_{\mathbb{C}}$	$\begin{pmatrix} 0 & -T_k \\ T_k & 0 \end{pmatrix}_{\mathbb{C}}$	$k \in \mathbb{N}$	$\alpha + i\beta$	1
2	$\begin{pmatrix} \alpha I_k + N_k & 0 \\ 0 & -\alpha I_k - N_k \end{pmatrix}$	$\begin{pmatrix} 0 & -T_k \\ T_k & 0 \end{pmatrix}$	$k \in \mathbb{N}$	α	1
3	$(i\beta I_k + N_k)_{\mathbb{C}}$	$\rho(X_k)_{\mathbb{C}}$	k even	$i\beta$	± 1
4	$(i\beta I_k + N_k)_{\mathbb{C}}$	$\rho(iX_k)_{\mathbb{C}}$	k odd	$i\beta$	± 1
5	N_k	ρX_k	k even	0	± 1
6	$(N_k)_{\mathbb{C}}$	$(iX_k)_{\mathbb{C}}$	k odd	0	1

Table 1: Normal form summands over $\mathbb{R}$ of size k, modulus μ and index ρ; $\alpha, \beta > 0$

	M	R	Size	μ	ρ
1	$\begin{pmatrix} \mu I_k + N_k & 0 \\ 0 & -\bar{\mu} I_k - N_k \end{pmatrix}_{\mathbb{C}}$	$\begin{pmatrix} 0 & -T_k \\ T_k & 0 \end{pmatrix}_{\mathbb{C}}$	$k \in \mathbb{N}$	$\alpha + i\beta$	1
2	$(i\beta I_k + N_k)_{\mathbb{C}}$	$\rho(X_k)_{\mathbb{C}}$	k even	$i\beta$	± 1
3	$(i\beta I_k + N_k)_{\mathbb{C}}$	$\rho(iX_k)_{\mathbb{C}}$	k odd	$i\beta$	± 1

Table 2: Normal form summands over $\mathbb{C}$ of size k, modulus μ and index ρ; $\alpha > 0, \beta \in \mathbb{R}$

	M	R	Size	μ	ρ
1	$\begin{pmatrix} \mu I_k + N_k & 0 \\ 0 & -\bar{\mu} I_k - N_k \end{pmatrix}_{\mathbb{H}}$	$\begin{pmatrix} 0 & -T_k \\ T_k & 0 \end{pmatrix}_{\mathbb{H}}$	$k \in \mathbb{N}$	$\alpha + i\beta$	1
2	$(i\beta I_k + N_k)_{\mathbb{H}}$	$\rho(X_k)_{\mathbb{H}}$	k even	$i\beta$	± 1
3	$(i\beta I_k + N_k)_{\mathbb{H}}$	$\rho(iX_k)_{\mathbb{H}}$	k odd	$i\beta$	$\pm 1; \quad \beta > 0$ $1; \quad \beta = 0$

Table 3: Normal form summands over $\mathbb{H}$ of size k, modulus μ and index ρ; $\alpha > 0$, $\beta \geq 0$

Given these restrictions on the eigenvalues of M, it follows from Jordan normal form theory, Theorem 1.3, that M is equivalent to a direct sum of the matrices listed in the 'M' column of the tables. Moreover the direct sum is unique up to the order of the summands.

The summands of the normal form of M are uniquely determined by the *size* $k \in \mathbb{N}$ and the *modulus* $\mu \in \mathbb{C}$ as listed in the tables. Note that we do not call μ the eigenvalue even though it is one of the quadruplet of eigenvalues. This is because we wish to stress that when $\mathcal{D} = \mathbb{C}$ summands with modulus μ are not similar to those with modulus $\bar{\mu}$. In particular the corresponding normal form summands are not equivalent.

In addition we associate to each normal form summand the number $\rho = \pm 1$. Suppose that (M, R) is in normal form and that $\mu \in \mathbb{C}$, $k \in \mathbb{N}$. Define the *index* $\mathrm{ind}_{\mu,k}(M, R)$ to be the sum of the indices of all summands of size k and modulus μ (if there is no summand of size k with modulus μ, set $\mathrm{ind}_{\mu,k}(M, R) = 0$). More generally, we can define the indices of a pair (M, R) to be equal to the indices of the corresponding normal form. Implicit in Theorem 2.4 is the statement that the indices form a set of invariants for the equivalence relation. (Again, we should stress that the indices $\mathrm{ind}_{\mu,k}(M, R)$ and $\mathrm{ind}_{\bar{\mu},k}(M, R)$ are treated independently when $\mathcal{D} = \mathbb{C}$.)

To sum up we have the following result.

Corollary 2.5 *Two symplectic pairs (M, R) and (M', R') are equivalent if and only if*

(a) M is similar to M' by a change of coordinates with entries in $\mathcal{D}$*, and*

(b) $\mathrm{ind}_{\mu,k}(M, R) = \mathrm{ind}_{\mu,k}(M', R')$ *for all $\mu \in \mathbb{C}$, $k \in \mathbb{N}$.*

Remark 2.6 The invariance of the indices may seem like a nuisance factor and possibly this is true when $\mathcal{D} = \mathbb{R}$ or $\mathbb{H}$. However as shown in Section 6, the indices are responsible for the existence of nonisomorphic symplectic structures when $\mathcal{D} = \mathbb{C}$. We observe here that the issue of indices does not arise when the modulus μ has nonzero real part and also when $\mu = 0$ for k odd and $\mathcal{D} \neq \mathbb{C}$.

3 The main ideas behind Williamson's theorem

In this section, we describe some of the main ideas behind Williamson's original proof. This leads to a reformulation of the problem. Also we describe some 'coordinate-free' results that are fundamental in the solution of the reformulated problem.

Suppose that (M, R) is a symplectic pair over the real division ring $\mathcal{D}$. The method for reducing (M, R) to normal form divides into three main steps.

Step 1 Reduce M to a simpler matrix L so that $(M, R) \sim (L, S)$ for some $S \in \mathbf{sk}$.

It is easy to prove that eigenvalues of L occur in quadruplets $\pm\mu$, $\pm\bar{\mu}$. Let V_μ denote the sum of the corresponding quadruplet of generalized eigenspaces.

Step 2 Write down a nonsingular matrix Q such that

(i) $LQ + QL^T = 0$.

(ii) $Q(V_\mu) \subset V_\mu$ for any eigenvalue μ of L.

We call Q the matrix *associated* to L. Write $S = GQ$ so that $(M, R) \sim (L, GQ)$.

Step 3 Reduce G to a simpler matrix H: $(L, GQ) \sim (L, HQ)$.

Step 1 relies on Jordan normal form theory, but apart from this Steps 1 and 2 taken alone are trivial. The important thing is to choose the matrices L and Q with Step 3 in mind. We discuss this more fully at the end of the section. For the time being we concentrate on those aspects of Step 3 that do not rely on a judicious choice of L and Q.

Define the centralizer $Z(L)$ to be the set of matrices with entries in $\mathcal{D}$ that commute with L. A computation, see Proposition 3.1 below, shows that the matrix G lies in $Z(L)$. In addition, $(L, GQ) \sim (L, HQ)$ if and only if there is a nonsingular matrix P such that

$$PLP^{-1} = L, \quad PGQP^T = HQ.$$

In particular, the matrices P, G, H and L all lie in $Z(L)$. Now matrices that commute with L leave invariant the generalized eigenspaces of L and hence preserve the subspaces V_μ. In addition, Q is defined so as to have this property. It follows that all of the matrices concerned, L, Q, G, H, P, have a block-diagonal structure corresponding to the decomposition of $\mathbb{R}^n$ into subspaces V_μ. Hence, in Step 3 we may assume without loss of generality that the matrix L has a single quadruplet of eigenvalues.

Suppose then that L has a single quadruplet of eigenvalues. Usually we take Q to be skew-symmetric, so assume that this is the case. By Proposition 3.1 below, $G \in Z(L, Q)$ where

$$Z(L, Q) = \{G \in Z(L);\, GQ = QG^T\}.$$

Define an equivalence relation on $Z(L, Q)$: $G \sim G'$ if there is a nonsingular matrix $P \in Z(L)$ such that $PGQP^T = G'Q$. Then Step 3 amounts to finding a normal form H for the matrix $G \in Z(L, Q)$. The strategy in the reduction of G is to first transform G into a block-diagonal matrix. Lemma 3.2 below is the main tool in performing this block-diagonalization.

In the sequel (L, Q) will denote a pair of $m \times m$ matrices where Q is nonsingular and $LQ + QL^T = 0$. At this point we drop the assumption that Q is skew-symmetric so that (L, Q) need not be a symplectic pair.. The resulting gain in generality is required in the case of zero eigenvalues and $\mathcal{D} = \mathbb{R}$.

Proposition 3.1 *Suppose that (L, S) is a symplectic pair, and that Q is a nonsingular matrix satisfying $LQ + QL^T = 0$. Then $S = GQ$ where G is a nonsingular matrix satisfying*

$$GL = LG, \quad GQ = -Q^T G^T.$$

In particular, if Q is skew-symmetric, then $G \in Z(L, Q)$.

Proof Set $G = SQ^{-1}$. Then G is nonsingular. We compute that

$$LG = LSQ^{-1} = -SL^TQ^{-1} = SQ^{-1}L = GL,$$

and that

$$GQ = S = -S^T = -Q^TG^T,$$

as required. ∎

Lemma 3.2 *Suppose that the matrices $L = L_1 \oplus L_2 \oplus L_3$, $Q = Q_1 \oplus Q_2 \oplus Q_3$, and* $G = \begin{pmatrix} G_{11} & 0 & 0 \\ 0 & G_{22} & G_{23} \\ 0 & G_{32} & G_{33} \end{pmatrix}$ *satisfy the following conditions.*

(a) Q is nonsingular and $Q_2^T = \pm Q_2$.

(b) $GL = LG, \quad GQ = -Q^T G^T$.

(c) G_{22} is nonsingular.

Then there is a nonsingular matrix $P \in Z(L)$ such that $PGQP^T = HQ$ where H has the form

$$H = \begin{pmatrix} G_{11} & 0 & 0 \\ 0 & G_{22} & 0 \\ 0 & 0 & H_{33} \end{pmatrix}.$$

*In particular, if $Q \in$ **sk** then $G, H \in Z(L, Q)$ and $G \sim H$.*

Proof The conditions in (b) imply in particular that

$$G_{j2}L_2 = L_j G_{j2}, \quad G_{2j}Q_j = \mp Q_2 G_{j2}^T.$$

Let P be the matrix

$$P = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & -G_{32}G_{22}^{-1} & I \end{pmatrix}.$$

Then P is nonsingular and it is easily checked that P commutes with L and that $H = PGQP^TQ^{-1}$ has the required form. ■

4 An abstraction of Williamson's Theorem

In this section we state and prove a result, Theorem 4.6, that may be thought of as an abstract version of Williamson's Theorem. It corresponds to Step 3 described in Section 3.

Let $\mathcal{D}$ be one of the real division rings $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ and let $\mathcal{M}$ denote the space of $p \times p$ matrices with entries in $\mathcal{D}$. We denote by $1_{\mathcal{M}}$ the identity matrix in $\mathcal{M}$. The properties of $\mathcal{M}$ that we use are:

1. If $a \in \mathcal{M}$, then $a^T \in \mathcal{M}$.
2. If $a \in \mathcal{M}$ is nonsingular, then $a^{-1} \in \mathcal{M}$.
3. Suppose that $\mathcal{M} \neq \mathbb{R}$, and $g \in \mathcal{M}$, $g \neq 0$. Then there exists a matrix $a \in \mathcal{M}$ such that ag is neither symmetric nor skew-symmetric.

We defined the real $k \times k$ matrices I_k and N_k in Subsection 1.1(a). Recall also the definitions of $Z(L)$ and $Z(L,Q)$ in Section 3.

Definition 4.1 A pair of matrices (L,Q) is a *W-summand (of size k)* if $L = I_k \otimes \pi + N_k \otimes \phi$ and $Q = Y_k \otimes \tau$, where $\pi, \phi, \tau \in \mathcal{M}$ and Y_k is a real nonsingular $k \times k$ matrix such that the following hypotheses are satisfied for some choice of $\sigma_1, \sigma_2 = \pm 1$:

(H1) ϕ, τ, Y_k are nonsingular, and either π is nonsingular or $Z(\pi) = \mathcal{M}$.

(H2) π is semisimple, $\quad \tau^T = \tau^{-1} = \sigma_1 \tau, \quad Y_k^T = Y_k^{-1} = -\sigma_1 Y_k$.

(H3) $\pi\tau = -\tau\pi^T, \quad \phi\tau = -\sigma_2 \tau \phi^T, \quad N_k Y_k = \sigma_2 Y_k N_k^T$.

(H4) $Z(\pi) \subset Z(\phi)$.

(H5) $Z(\pi, \tau)$ is contained in a real division ring. In addition, if $p \in Z(\pi,\tau)$ then $p^T \in Z(\pi, \tau)$.

Definition 4.2 Suppose that (L_i, Q_i) is a W-summand of size k_i for $i = 1, \dots, r$, and that $k_1 \geq k_2 \geq \cdots \geq k_r$. Let $L = L_1 \oplus \cdots \oplus L_r$ and $Q = Q_1 \oplus \cdots \oplus Q_r$. Then (L,Q) is a *W-sum* if

(a) $L_i = I_{k_i} \otimes \pi + N_{k_i} \otimes \phi$ where π and ϕ are independent of i,

(b) If $k_i = k_j$ then $L_i = L_j$ and $Q_i = Q_j$,

(c) If $k_i > k_j$ and A is a $k_j \times k_j$ matrix, then $Y_{k_i} \begin{pmatrix} 0 \\ A \end{pmatrix} = \begin{pmatrix} Y_{k_j} A \\ 0 \end{pmatrix}$.

Remark 4.3 (a) The matrices π and ϕ that appear in the tensor product form of L_i do not depend on i. However the matrices τ_i may depend on k_i. (b) $Z(\pi)$ is a subring of $\mathcal{M}$ and if $s \in Z(\pi)$ is nonsingular, then $s^{-1} \in Z(\pi)$. In general $Z(\pi, \tau)$ is not closed under multiplication (unless $Z(\pi)$ is commutative) and so is not a subring of $\mathcal{M}$.

(c) By hypothesis (H5), $Z(\pi, \tau)$ is contained in a real division ring and may be viewed as a subset of the reals, complexes or quaternions ($\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$). With this identification, we may define the norm of an element $p \in Z(\pi, \tau)$ to be $\|p\| = \sqrt{pp^T} = \sqrt{p\bar{p}}$. Let $Z_1(\pi, \tau) = \{p \in Z(\pi, \tau); \|p\| = 1\}$. We may view this set as being contained in the space of unit reals, complexes or quaternions.
(d) In practice, we choose $Y_{k_i} = T_{k_i}$ for each i or $Y_{k_i} = X_{k_i}$ for each i. Then part (c) of Definition 4.2 is satisfied by Proposition 1.1(v).

Proposition 4.4 *Any W-sum (L, Q) is a symplectic pair (over $\mathcal{D}$).*

Proof We must verify that Q is nonsingular and skew-symmetric, and that $LQ + QL^T = 0$. It is sufficient to consider the case where (L, Q) is a W-summand. Recall that the matrices Y_k and τ are assumed to be nonsingular. Also, by hypothesis (H2) one of these matrices is symmetric and the other skew-symmetric. Hence $Q = Y_k \otimes \tau$ is nonsingular and skew-symmetric. Using hypothesis (H3) we compute that

$$\begin{aligned}
LQ &= (I_k \otimes \pi + N_k \otimes \phi)(Y_k \otimes \tau) \\
&= Y_k \otimes (\pi\tau) + (N_k Y_k) \otimes (\phi\tau) \\
&= -Y_k \otimes (\tau\pi^T) - (Y_k N_k^T) \otimes (\tau\phi^T) \\
&= -(Y_k \otimes \tau)(I_k \otimes \pi + N_k \otimes \phi)^T \\
&= -QL^T
\end{aligned}$$

as required. ■

Corollary 4.5 *Suppose that (L, Q) is a W-sum and that $L \in \mathbf{sp}(S)$ for some $S \in \mathbf{sk}$. Then $S = GQ$ where $G \in Z(L, Q)$ is nonsingular.*

Proof This is immediate from Proposition 4.4 and Proposition 3.1. ■

We shall say that a W-sum is of *type I* if there is a division ring $\mathcal{R}$ (not necessarily the same as $\mathcal{D}$) and an isomorphism (denoted by $\cong$) from the space of 2×2 matrices over $\mathcal{R}$ into $\mathcal{M}$ such that

(i) $\tau_i \cong \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, for $i = 1, \ldots, r$.

(ii) $Z(\pi)$ contains all matrices of the form $p \cong \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ where $a, b \in \mathcal{R}$.

(iii) $Z(\pi, \tau_i) \cong \left\{ \begin{pmatrix} a & 0 \\ 0 & \bar{a} \end{pmatrix} ; a \in \mathcal{R} \right\}$.

A W-sum is of *type II* if $Z(\pi) \subset Z(\tau_i)$ for $i = 1, \dots, r$.

If (L, Q) is a W-sum, define $Z(L, Q)^0$ to consist of those matrices $H \in Z(L, Q)$ that have the form $H = H_1 \oplus \cdots \oplus H_r$ where $H_i = \rho_i I_{k_i} \otimes 1_{\mathcal{M}}$ and $\rho_i = \pm 1$ if (L_i, Q_i) is of type II, $\rho_i = 1$ otherwise. If $H \in Z(L, Q)^0$ define $\mathrm{ind}_k(H) = \sum_{k_i = k} \rho_i$.

Theorem 4.6 (Abstract Williamson Theorem) *Suppose that (L, Q) is a W-sum with summands (L_i, Q_i) of size k_i, $i = 1, \dots, r$. Let $G \in Z(L, Q)$ be a nonsingular matrix. Then*

(a) $G \sim H$ where $H = H_1 \oplus \cdots \oplus H_r$ and $H_i = I_{k_i} \otimes h_i$, $h_i \in Z_1(\pi, \tau_i)$.

(b) If (L, Q) is of type I or of type II, then $G \sim H$ where $H \in Z(L, Q)^0$.

(c) Suppose that $G, H \in Z(L, Q)^0$, $G \sim H$. Then $\mathrm{ind}_k(G) = \mathrm{ind}_k(H)$ *for all k.*

The remainder of this subsection is devoted to proving the theorem. We break the proof into four stages. First we obtain some results on the structure of matrices in $Z(L)$ and $Z(L, Q)$. Then there is a block-diagonalization step where we reduce G to a matrix $H = H_1 \oplus \cdots \oplus H_r$. Next we treat each summand H_i separately completing parts (a) and (b) of the theorem. Finally we prove the uniqueness statement in part (c).

4.1 Structure of $Z(L)$ and $Z(L, Q)$

Suppose that (L, Q) is a W-sum with W-summands (L_i, Q_i) of size k_i. We can partition a matrix $P \in Z(L)$ into blocks P_{ij}, $1 \le i, j \le r$ where P_{ij} is a $k_i \times k_j$ matrix with entries in $\mathcal{M}$. If $P \in Z(L, Q)$ the conditions $PL = LP$, $PQ = QP^T$ become

$$P_{ij} L_j = L_i P_{ij}, \quad P_{ij} Q_j = Q_i P_{ji}^T.$$

Let $k_{ij} = \min(k_i, k_j)$.

Proposition 4.7 *A matrix $P = \{P_{ij}\}$ lies in $Z(L)$ if and only if the following is true.*

(a) $P_{ij} = \begin{pmatrix} 0 & F_{ij} \end{pmatrix}$ *if* $k_i \le k_j$ *or* $P_{ij} = \begin{pmatrix} F_{ij} \\ 0 \end{pmatrix}$ *if* $k_i \ge k_j$, *where* F_{ij} *is a* $k_{ij} \times k_{ij}$ *matrix with entries in* $\mathcal{M}$.

(b) $F_{ij} = \displaystyle\sum_{s=0}^{k_{ij}-1} N_{k_{ij}}^s \otimes f_{ij,s}$ *where* $f_{ij,s} \in Z(\pi)$.

Proof See the appendix. ∎

We shall refer to the blocks P_{ii} as diagonal blocks and P_{ij}, $i \ne j$, as off-diagonal blocks. Note that $F_{ii} = P_{ii}$ for each i. The element $f_{ij,0}$ appearing in the expansion of F_{ij} is called the *leading coefficient*. Define

$$\tilde{F}_{ij} = \sum_{s=0}^{k_{ij}-1} \sigma_2^s (N_{k_{ij}}^s)^T \otimes f_{ij,s},$$

and let $\tilde{P}_{ij} = \begin{pmatrix} \tilde{F}_{ij} & 0 \end{pmatrix}$ if $k_i \le k_j$ and $\tilde{P}_{ij} = \begin{pmatrix} 0 \\ \tilde{F}_{ij} \end{pmatrix}$ if $k_i \ge k_j$.

Proposition 4.8 *Suppose that $P = \{P_{ij}\} \in Z(L)$. If $i \le j$, then*

$$P_{ij}(Y_{k_j} \otimes 1_{\mathcal{M}}) = (Y_{k_i} \otimes 1_{\mathcal{M}})\tilde{P}_{ij}.$$

Proof Since $i \le j$, $k_i \ge k_j$ and we compute that

$$\begin{aligned} P_{ij}(Y_{k_j} \otimes 1_{\mathcal{M}}) &= \sum_{s=0}^{k_j-1} \begin{pmatrix} N_{k_j}^s \\ 0 \end{pmatrix} Y_{k_j} \otimes f_{ij,s} \\ &= \sum_{s=0}^{k_j-1} \begin{pmatrix} N_{k_j}^s Y_{k_j} \\ 0 \end{pmatrix} \otimes f_{ij,s} \\ &= \sum_{s=0}^{k_j-1} \sigma_2^s \begin{pmatrix} Y_{k_j}(N_{k_j}^s)^T \\ 0 \end{pmatrix} \otimes f_{ij,s} \\ &= \sum_{s=0}^{k_j-1} \sigma_2^s Y_{k_i} \begin{pmatrix} 0 \\ (N_{k_j}^s)^T \end{pmatrix} \otimes f_{ij,s} \\ &= (Y_{k_i} \otimes 1_{\mathcal{M}})\tilde{P}_{ij}. \end{aligned}$$

■

Proposition 4.9 *Suppose that $P = \{P_{ij}\} \in Z(L, Q)$ and that $k_i = k_j$. Then $f_{ij,0}\tau_i = \tau_i f_{ji,0}^T$.*

Proof The condition $PQ = QP^T$ implies in particular that $P_{ij}Q_j = Q_i P_{ji}^T$. Since $k_i = k_j$ we have $Q_i = Q_j$ and compute that

$$\begin{aligned} Q_i P_{ji}^T &= P_{ij} Q_i \\ &= P_{ij}(Y_{k_i} \otimes 1_{\mathcal{M}})(I_{k_i} \otimes \tau_i) \\ &= (Y_{k_i} \otimes 1_{\mathcal{M}})\tilde{P}_{ij}(I_{k_i} \otimes \tau_i) \\ &= Q_i(I_{k_i} \otimes \tau_i^{-1})\tilde{P}_{ij}(I_{k_i} \otimes \tau_i). \end{aligned}$$

Hence

$$(I_{k_i} \otimes \tau_i)P_{ji}^T = \tilde{P}_{ij}(I_{k_i} \otimes \tau_i),$$

or

$$\sum_{s=0}^{k_i-1} (N_{k_i}^s)^T \otimes \tau_i f_{ji,s}^T = \sum_{s=0}^{k_i-1} \sigma_2^s (N_{k_i}^s)^T \otimes f_{ij,s}\tau_i.$$

Using the linear independence of the matrices $N_{k_i}^s$ and comparing coefficients when $s = 0$ we have the required condition on $f_{ij,0}$. ■

4.2 Block-diagonalization of matrices in $Z(L, Q)$

Suppose that (L, Q) is a W-sum with summands (L_i, Q_i) of size k_i, where $k_1 \geq \cdots \geq k_r$. Let $G \in Z(L, Q)$ and suppose that G_{ii} is a diagonal block with leading coefficient g_{ii}.

Corollary 4.10 *$g_{ii} \in Z(\pi, \tau_i)$ and G_{ii} is nonsingular if and only if $g_{ii} \neq 0$.*

Proof By Proposition 4.7, we have $g_{ii} \in Z(\pi)$ and Proposition 4.9 implies that $g_{ii}\tau_i = \tau_i g_{ii}^T$ so that $g_{ii} \in Z(\pi, \tau_i)$. Now it follows from Proposition 4.7 that G_{ii} is an upper-triangular matrix with diagonal entries g_{ii}. Hence G_{ii} is nonsingular if and only if g_{ii} is nonsingular. But (H5) states that $Z(\pi, \tau_i)$ lies inside a division ring, so the only singular element is zero. ■

Define numbers k and c so that

$$k = k_1 = \cdots = k_c, \quad k_{c+1} < k.$$

Lemma 4.11 *Suppose that $G \in Z(L, Q)$ is nonsingular. If G_{ii} is singular, $1 \le i \le c$, then $G \sim H$ where $H \in Z(L, Q)$ has the property that H_{11} is nonsingular.*

Proof Suppose that $1 \le i, j \le c$. Then $G_{ij} \in Z(L_1)$ and $G_{ii} \in Z(L_1, Q_1)$. Let g_{ij} denote the leading coefficients of these G_{ij}. By Corollary 4.10 we have that $g_{11} = \cdots = g_{cc} = 0$. In addition, by Proposition 4.9, $g_{ij}\tau_1 = \tau_1 g_{ji}^T$, for $1 \le i, j \le c$.

Since G is nonsingular, there must be a nonzero element in the first column. Now it follows from Proposition 4.7 that a block G_{i1} has a nonzero element in the first column only if $i \le c$ and the leading coefficient is nonzero. Hence $g_{i1} \ne 0$ for some $i = 2, \ldots c$. By interchanging the 2nd and ith rows and columns we may suppose that $g_{21} \ne 0$.

We shall consider transformation matrices of the form $P = I + \tilde{P}$ where $\tilde{P}$ is the zero matrix except for one block corresponding to the block G_{12} of G. This block of $\tilde{P}$ has the form $I_k \otimes p$ where $p \in Z(\pi)$. Then P is a nonsingular matrix commuting with L.

Let $H = PGQP^TQ^{-1}$. Observe that $QP^T = EQ$ where $E = I + \tilde{E}$ and $\tilde{E}$ has a single nonzero block $I_k \otimes e$ corresponding to the block G_{21} of G. Here $e = \tau_1 p^T \tau_1^T$. Thus $H = PGE$ and we compute that

$$h_{11} = pg_{21} + g_{12}\tau_1 p^T \tau_1^T.$$

Our aim is to choose P so that $h_{11} \ne 0$. Then H_{11} is nonsingular by Corollary 4.10 as required.

There are two cases to consider, corresponding to the possibilities that π is nonsingular, or $Z(\pi) = \mathcal{M}$ (cf. (H1)). Suppose first that π is nonsingular and consider the choices $p = 1_{\mathcal{M}}$ and $p = \pi$. The corresponding matrices e are $1_{\mathcal{M}}$ and $-\pi$. We compute that $h_{11} = g_{21} + g_{12}$ or $h_{11} = -\pi g_{21} + g_{12}\pi$. Since π is nonsingular and commutes with g_{ij} we have that $h_{11} = 0$ in both cases if and only if $g_{12} + g_{21} = g_{12} - g_{21} = 0$. But then $g_{21} = 0$ which is a contradiction.

It remains to consider the case when $Z(\pi) = \mathcal{M}$. Suppose that $\mathcal{M} \ne \mathbb{R}$. By the third requirement on $\mathcal{M}$, there exists a matrix $\tilde{p} \in Z(\pi)$ such that $\tilde{p}g_{21}$ is neither symmetric nor skew-symmetric. Choose $p = \tau_1\tilde{p}$ and compute that

$$\begin{aligned} h_{11} &= pg_{21} + g_{12}\tau_1 p^T \tau_1^T \\ &= \tau_1 \tilde{p} g_{21} + \tau_1 g_{21}^T (\tau_1 \tilde{p})^T \tau_1^T \\ &= \tau_1 \left(\tilde{p} g_{21} + g_{21}^T \tilde{p}^T (\tau_1^T)^2 \right) \\ &= \tau_1 \left(\tilde{p} g_{21} + \sigma_1 (\tilde{p} g_{21})^T \right), \end{aligned}$$

where $\sigma_1 = \pm 1$, see (H2). In particular, $h_{11} = 0$ only if $\tilde{p} g_{21}$ is symmetric or skew-symmetric which we have assumed not to be the case. Finally, if $\mathcal{M} = \mathbb{R}$, take $\tilde{p} = 1$. Then $h_{11} = 2\tau_1 g_{21} = 0$ if and only if $g_{21} = 0$. ∎

Corollary 4.12 *If $G \in Z(L, Q)$ is nonsingular, then $G \sim H$ where H has the form $H_1 \oplus \cdots \oplus H_r$.*

Proof Suppose that G_{11} is singular. If one of the matrices G_{ii}, $1 \le i \le c$, is nonsingular then we may rearrange rows and columns so that G_{11} is nonsingular. (Such a transformation does not change L or Q since $L_1 = \cdots = L_c$ and $Q_1 = \cdots = Q_c$.) On the other hand, if each of the matrices G_{ii}, $1 \le i \le c$ is singular, we may apply Lemma 4.11 and perform a transformation so that the resulting matrix has a nonsingular block G_{11}.

Thus we may assume that G_{11} is nonsingular. By Lemma 3.2 we may replace G by a matrix with G_{11} unchanged, but $G_{1j} = G_{j1} = 0$ for $2 \le j \le r$. Working inductively we may block-diagonalize G into a matrix of the required form. ∎

4.3 Reduction of a single W-summand

It follows from Corollary 4.12 that to prove parts (a) and (b) of Theorem 4.6 it is sufficient to work one summand at a time. Hence we may suppose that (L, Q) consists of one W-summand of size k. Let $G \in Z(L, Q)$ be nonsingular. By Proposition 4.7, G has the expansion

$$G = \sum_{s=0}^{k-1} N_k^s \otimes g_s,$$

where $g_s \in Z(\pi)$. Moreover, $g_0 \in Z(\pi, \tau)$ is nonsingular.

Proposition 4.13 *$G \sim I_k \otimes h$, where $h \in Z_1(\pi, \tau)$.*

Proof First we prove by induction that $G \sim I_k \otimes g_0$. Suppose that $G = I_k \otimes g_0 + N_k^i \otimes g_i \bmod N_k^{i+1}$. We shall find a nonsingular matrix $P \in Z(L)$ so that $PGQP^TQ^{-1} = H$ where $H = I_k \otimes g_0 \bmod N_k^{i+1}$. Indeed set $P = I_k \otimes 1_{\mathcal{M}} - \frac{1}{2}N_k^i \otimes (g_i g_0^{-1})$. Then P is nonsingular and commutes with L (by Remark 4.3(b)). Define $E = I_k \otimes 1_{\mathcal{M}} - \frac{1}{2}N_k^i g_0^{-1} g_i$ and compute that $QP^T = EQ$. Hence $H = PGE = I_k \otimes g_0 \bmod N_k^{i+1}$ as required.

It remains to show that the element g_0 can be replaced by an element with norm one. Let $p = \frac{1}{\sqrt{\|g_0\|}} 1_{\mathcal{M}}$ and define $P = I_k \otimes p$. Then P is a nonsingular matrix in $Z(L)$ and the leading coefficient g_0 is transformed to

$$h = pg_0\tau p^T\tau^{-1} = p^2 g_0 = \frac{g_0}{\|g_0\|}.$$

■

We have proved part (a) of Theorem 4.6. Part (b) follows from the next proposition.

Proposition 4.14 *Suppose that (L,Q) is a W-summand and that G is a nonsingular matrix in $Z(L,Q)$.*

(a) If (L,Q) is of type I, then $G \sim I_k \otimes 1_{\mathcal{M}}$.

(b) If (L,Q) is of type II, then $G \sim \pm I_k \otimes 1_{\mathcal{M}}$.

Proof By Proposition 4.13 we may assume that $G = I_k \otimes g$ where $g \in Z_1(\pi,\tau)$. Suppose first that (L,Q) is of type I. Then $g \cong \begin{pmatrix} a & 0 \\ 0 & \bar{a} \end{pmatrix}$ where $\|a\| = 1$. Also $\tau \cong \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $p \cong \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} \in Z(\pi)$. Let $P = I_k \otimes p$ and compute that

$$PGQP^T = Y_k \otimes (pg\tau p^T) = Y_k \otimes \tau = (I_k \otimes 1_{\mathcal{M}})Q,$$

as required to prove part (a).

Next suppose that (L,Q) is of type II. Then $Z(\pi) \subset Z(\tau)$ and it follows that $Z(\pi,\tau)$ consists of symmetric matrices. By (H5) $Z(\pi,\tau) \cong \mathbb{R}$. In particular, $g = \pm 1$ and G is already as claimed in part (b). ■

4.4 Uniqueness

It remains to prove part (c) of Theorem 4.6. Suppose that (L, Q) is a W-sum with summands (L_i, Q_i) of size k_i so that $k_1 \geq \cdots \geq k_r$. Suppose further that $G, H \in Z(L, Q)^0$ are block-diagonal matrices with i'th summands $G_i = \rho_i I_{k_i} \otimes 1_{\mathcal{M}}$ and $H_i = \sigma_i I_{k_i} \otimes 1_{\mathcal{M}}$ respectively, $\rho_i, \sigma_i = \pm 1$. Let $k = k_i$ for some i and define A to consist of those indices α, $1 \leq \alpha \leq r$ such that $k_\alpha = k$. Write $\tau = \tau_i = \tau_\alpha$ for $\alpha \in A$.

Recall that the indices $\operatorname{ind}_k(G)$ and $\operatorname{ind}_k(H)$ are given by

$$\operatorname{ind}_k(G) = \sum_{\alpha \in A} \rho_\alpha, \quad \operatorname{ind}_k(H) = \sum_{\alpha \in A} \sigma_\alpha$$

We must show that if $G \sim H$ then $\operatorname{ind}_k(G) = \operatorname{ind}_k(H)$. We may assume that (L_i, Q_i) is of type II for otherwise the indices are both equal to the cardinality $|A|$ of A.

Suppose that $G \sim H$. Then there is a nonsingular matrix $P \in Z(L)$ such that $PGQP^T = HQ$. Write $P = \{P_{ij}\}$.

Proposition 4.15 *Let $\alpha, \beta \in A$. There are numbers $\epsilon_{\alpha j} = \pm 1$ such that*

$$\sum_{j=1}^{r} \epsilon_{\alpha j} \tilde{P}_{\alpha j} (I_{k_j} \otimes \rho_j \tau_j) P_{\beta j}^T = \delta_{\alpha\beta} I_k \otimes \sigma_\alpha \tau, \tag{4.1}$$

where $\delta_{\alpha\beta}$ is the Kronecker delta. If $j \in A$, $\epsilon_{\alpha j} = 1$.

Proof Equating α, β'th entries in $PGQP^T = HQ$ yields the equation

$$\sum_{j=1}^{r} P_{\alpha j} G_j Q_j P_{\beta j}^T = \delta_{\alpha\beta} H_\alpha Q_\alpha = \delta_{\alpha\beta} \sigma_\alpha Y_k \otimes \tau_\alpha. \tag{4.2}$$

Using Proposition 4.8 we compute that provided $k_j \leq k_\alpha = k$,

$$P_{\alpha j} G_j Q_j = P_{\alpha j} Y_{k_j} \otimes \rho_j \tau_j = (Y_k \otimes 1_{\mathcal{M}}) \tilde{P}_{\alpha j} (I_{k_j} \otimes \rho_j \tau_j).$$

If $k_j > k_\alpha$ then $j < \alpha$ and we cannot apply Proposition 4.8 directly. However, taking transposes in part (c) of Definition 4.2 we find that if A is a $k \times k$ matrix then $\begin{pmatrix} 0 & A \end{pmatrix} Y_{k_j} = \pm \begin{pmatrix} AY_k & 0 \end{pmatrix}$. It follows that

$$P_{\alpha j} G_j Q_j = \pm (Y_k \otimes 1_{\mathcal{M}}) \tilde{P}_{\alpha j} (I_{k_j} \otimes \rho_j \tau_j).$$

Substitute into equation (4.2) and cancel the common factor $Y_k \otimes 1_{\mathcal{M}}$ to obtain equation (4.1). ■

Proposition 4.16 *Regard $\tilde{P}_{\alpha j}(I_{k_j} \otimes \rho_j \tau_j) P^T_{\beta j}$ as a $k \times k$ matrix with entries in $\mathcal{M}$. If $j \notin A$ then the top-left entry is zero. If $j \in A$ then the top-left entry is given by $f_{\alpha j,0} \rho_j \tau f^T_{\beta j,0}$.*

Proof We shall assume that the top-left entry is nonzero and show that this implies that $k_j = k$ so that $j \in A$ and that the entry has the required form. First observe that the first row of $\tilde{P}_{\alpha j}$ must have nonzero entries. It follows from the definition of $\tilde{P}_{\alpha j}$ immediately after Proposition 4.7 that $k_j \geq k_\alpha = k$ and that the first row of $\tilde{P}_{\alpha j}$ is $(f_{\alpha j,0}\ 0\ \cdots\ 0)$. Hence the first row of $\tilde{P}_{\alpha j}(I_{k_j} \otimes \rho_j \tau_j)$ is $(f_{\alpha j,0} \rho_j \tau_j\ 0\ \cdots\ 0)$.

It follows in turn that the top-left entry of $P^T_{\beta j}$ must be nonzero. Applying Proposition 4.7 we find that $k_j \leq k_\alpha = k$, and so $k_j = k$. Moreover, the top-left entry of $P^T_{\beta j}$ is $f^T_{\beta j,0}$ as required. ■

Define ρ to be the $|A| \times |A|$ diagonal matrix with entries ρ_α, $\alpha \in A$. Similarly, define σ with entries σ_α.

Proposition 4.17 *Suppose that $G \sim H$. Then there is a matrix f such that $f \rho f^T = \sigma$.*

Proof Equating top-left entries in equation (4.1) and using Proposition 4.16 we obtain the equation

$$\sum_{j \in A} f_{\alpha j,0} \rho_j \tau f^T_{\beta j,0} = \delta_{\alpha\beta} \sigma_\alpha \tau.$$

Since (L_i, Q_i) is of type II, $\tau = \tau_i \in Z(\pi)$ and we can cancel factors of τ to obtain the equation

$$\sum_{j \in A} f_{\alpha j,0} \rho_j f^T_{\beta j,0} = \delta_{\alpha\beta} \sigma_\alpha.$$

Define f to be the matrix with entries $f_{\alpha\beta,0}$. Then $f \rho f^T = \sigma$ as required. ■

Corollary 4.18 *Suppose that (L, Q) is a W-sum with a summand (L_i, Q_i) of size k. If $G, H \in Z(L, Q)^0$ are equivalent then* $\mathrm{ind}_k(G) = \mathrm{ind}_k(H)$.

Proof The matrices ρ and σ are diagonal with entries $\rho_\alpha, \sigma_\alpha = \pm 1$. Since σ is nonsingular, the matrix f in Proposition 4.17 is nonsingular. The signature of symmetric matrices is invariant under congruent transformations and it follows that the number of $+1$ entries is the same in ρ and σ. In particular

$$\mathrm{ind}_k(G) = \mathrm{tr}\,\rho = \mathrm{tr}\,\sigma = \mathrm{ind}_k(H).$$

■

5 Proof of Williamson's Theorem

In this section we use the ideas described in Section 3 and the abstract version, Theorem 4.6, of Williamson's Theorem to reduce symplectic pairs (M, R) to normal form, thus proving Theorem 2.4.

There are three subsections. In Subsection 5.1 we use eigenvalue restrictions and Jordan normal form theory to reduce M to a simpler matrix L. Also we define a matrix Q associated to L. The case of zero eigenvalues over the division ring $\mathbb{R}$ is different from the others. Apart from this case the pairs (L, Q) are W-sums and are closely related to the putative normal forms. In Subsection 5.2 we exploit this structure and apply Theorem 4.6. This proves Theorem 2.4 over $\mathbb{C}$ and $\mathbb{H}$ and leaves only the zero eigenvalue case over $\mathbb{R}$ in Subsection 5.3.

5.1 Eigenvalue restrictions and a normal form for M

Suppose that (M, R) is a symplectic pair with entries in the real division ring $\mathcal{D}$. There are restrictions on the eigenvalues and Jordan blocks of M due to the condition $MR + RM^T = 0$. Since M is a real matrix, eigenvalues and the corresponding Jordan blocks occur in complex conjugate pairs. The extra structure leads to quadruplets of eigenvalues and Jordan blocks.

Lemma 5.1 *If M has p Jordan blocks of size k corresponding to the eigenvalues μ, $\bar{\mu}$ then M has p Jordan blocks of size k corresponding to the eigenvalues $-\mu$, $-\bar{\mu}$.*

Proof It is sufficient to show that M and $-M$ are related by a real similarity transformation (with entries not necessarily in $\mathcal{D}$). The condition $MR + RM^T = 0$ can be rearranged to show that M^T and $-M$ are similar matrices. But it follows from the (real) Jordan normal form theorem that M^T is similar to M so we have that M is similar to $-M$ as required. ■

Remark 5.2 There are restrictions also on the Jordan blocks of zero eigenvalues when $\mathcal{D} = \mathbb{R}$: blocks of odd size must occur with even multiplicity. It is hard to find a direct proof of this fact, but it is a consequence of the proof of Williamson's Theorem, see [17] and also Lemma 5.7(a) in this paper.

We now use Lemma 5.1 to complete Step 1 described in Section 3.

Theorem 5.3 *Suppose that (M, R) is a symplectic pair over $\mathcal{D}$. Then $(M, R) \sim (L, S)$ where L is a direct sum of summands $I_k \otimes \pi + N_k \otimes \phi$ and*

(a) If $\mathcal{D} = \mathbb{R}$, then π and ϕ are as in entries (i)–(iii) of Table 4 or $\pi = 0$, $\phi = 1$.

(b) If $\mathcal{D} = \mathbb{C}$, then π and ϕ are as in entries (iv) and (v) of the table.

(c) If $\mathcal{D} = \mathbb{H}$, then π and ϕ are as in entries (vi)–(ix) of the table.

Moreover the direct sum is unique up to the order of summands.

Proof This follows immediately from Lemma 5.1 and Theorem 1.3. ■

We proceed to Step 2.

Proposition 5.4 *Suppose that $L = I_k \otimes \pi + N_k \otimes \phi$ and $Q = Y_k \otimes \tau$ where the matrices π, ϕ, τ and Y_k are taken from a row in Table 4. Then (L, Q) is a W-summand.*

Proof This is a routine verification of hypotheses (H1)–(H5) in Definition 4.1. Parts (ii)-(iv) of Proposition 1.1 are useful in this verification. ■

Definition 5.5 If L and Q are as in the proposition, then we call Q the matrix *associated* to L. If $L = N_k$, the *associated matrix* is $Q = X_k$.

This definition associates to each summand L' of L in Theorem 5.3 a matrix Q'. Define Q to be the direct sum of the matrices Q'. We call Q the matrix *associated* to L. Recall that V_μ denotes the sum of the generalized eigenspaces corresponding to an eigenvalue μ of L.

Proposition 5.6 *The matrix Q is nonsingular, $Q(V_\mu) \subset V_\mu$ for each eigenvalue μ of L, and $LQ + QL^T = 0$,*

Proof The first two statements of the proposition are immediate from the definition of Q. It is sufficient to prove the final statement for a single summand (L, Q). If $L = N_k$ then $Q = X_k$ and we can apply Proposition 1.1(iv). Otherwise (L, Q) is a W-summand and the result follows from Proposition 4.4. ■

5.2 Normal forms in the abstract framework

Suppose that (M, R) is a symplectic pair. Carrying out steps 1 and 2 in Subsection 5.1 we have that $(M, R) \sim (L, GQ)$ where L has the form described in Theorem 5.3 and Q is the matrix associated to L (Definition 5.5. It remains to carry out Step 3, that is to reduce G to a simple form.

As described in Section 3 it is sufficient to carry out Step 3 under the assumption that L has a single quadruplet of eigenvalues. Hence $L = L_1 \oplus \cdots \oplus L_r$ where $L_i = I_{k_i} \otimes \pi + N_{k_i} \otimes \phi$ and π, ϕ are entries in Table 4 or $\pi = 0$, $\phi = 1$, $\mathcal{D} = \mathbb{R}$. This last case is different from the others and is considered in Subsection 5.3. Here we assume that π and ϕ are entries in Table 4. Associated to each summand L_i is a matrix $Q_i = Y_{k_i} \otimes \tau_i$ where τ_i and Y_{k_i} are the corresponding entries in the table.

We have already verified (Proposition 5.4) that (L_i, Q_i) is a W-summand for each $i = 1, \ldots, r$. Now arrange the summands so that $k_1 \geq \cdots \geq k_r$. Then it is easily checked that (L, Q) is a W-sum (Definition 4.2).

Consider first the (easiest) case $\mathcal{D} = \mathbb{C}$. We must make the correspondence between rows (iv) and (v) of Table 4 and rows 1–3 of Table 2. Now row (iv) yields a W-sum (L, Q) of type I. By Theorem 4.6(b), $(M, R) \sim (L, GQ) \sim (L, Q)$. The summands of (L, Q) have the form $(I_k \otimes \pi + N_k \otimes \phi, Y_k \otimes \tau)$. By Proposition 1.2(f), this is equivalent to $(\pi \otimes I_k + \phi \otimes N_k, \tau \otimes Y_k)$. But this is precisely the normal form summand in row 1 of Table 2.

The W-sum (L, Q) in row (v) of Table 4 is of type II. Again by Theorem 4.6(b), $(M, R) \sim (L, GQ) \sim (L, HQ)$ where summands of (L, HQ) have the form $(I_k \otimes \pi + N_k \otimes \phi, \pm Y_k \otimes \tau)$. This is the normal form summand in either row 2 or 3 of Table 2 depending on whether k is even or odd.

The case $\mathcal{D} = \mathbb{R}$ is similar, except when there are zero eigenvalues. Other than this we can make the correspondence between rows (i)–(iii) of Table 4 and rows 1–4 of Table 1.

It remains to consider the case $\mathcal{D} = \mathbb{H}$. The W-sums in rows (vi) and (vii) of Table 4 are of type I and lead to the normal form summands in row 1 of Table 3 with $\beta > 0$ and $\beta = 0$ respectively. Row (viii) yields a W-sum of type II and corresponds to rows 2 and 3 of Table 3 with $\beta > 0$. Row (ix) is slightly different. If k is even, the W-sum is of type II and corresponds to row 2 with $\beta = 0$. However if k is odd, the W-sum (L, Q) is not of type I or type II. Nevertheless, we claim that (L, Q) behaves like a W-sum of type I, that is $(M, R) \sim (L, GQ) \sim (L, Q)$. This is the normal form summand in row 3 with $\beta = 0$.

By Theorem 4.6(a), G can be block-diagonalized into the form $G = G_1 \oplus \cdots \oplus G_r$ where $G_i = I_{k_i} \otimes g_i$ and $g_i \in Z_1(\pi, \tau_i)$. Hence to verify the claim we may assume that there is a single summand of size k odd so that $(L, Q) = (N_k \otimes 1_{\mathbb{H}}, X_k \otimes i_{\mathbb{H}})$ and $G = I_k \otimes g$ where $g = (a + jc + kd)_{\mathbb{H}}$. Let $P = I_k \otimes p$ where p is a nonsingular matrix in $Z(\pi) \cong \mathbb{H}$. Then $PGQP^TQ^{-1} = I_k \otimes h$ where $h = pg\tau p^T \tau^{-1} = -pgi\bar{p}i$. First we choose p so that $h \in \mathbb{R}$. Suppose that $c \neq 0$. Take $p = \eta - j$ where $\eta = (a + \sqrt{a^2 + c^2})/c$, Then $h = pgp = a' + d'k$, where $d' = (\eta^2 + 1)d$. If $d' \neq 0$, then take $p = \eta' - k$ where η' is defined like η replacing c by d'. By normalizing p we may assume that $h = \pm 1$. If $h = -1$ take $p = j$ and compute that $-j(-1)i(-j)i = 1$ thus verifying the claim.

Aside from the case of zero eigenvalues when $\mathcal{D} = \mathbb{R}$, we have shown that any pair (M, R) is equivalent to a normal form composed of the normal form summands listed in Tables 1, 2 and 3. It remains to prove uniqueness. By the uniqueness statement of Theorem 5.3, the only possibility of nonuniqueness must stem from the choices $\rho = \pm 1$. But these choices only occur when the corresponding W-sum is of type II. Now apply Theorem 4.6(c).

5.3 The zero eigenvalue case over $\mathbb{R}$

In this section we consider the exceptional case of zero eigenvalues when $\mathcal{D} = \mathbb{R}$. Our methods are similar to those used in Section 4 but a priori we must use a matrix Q that is not skew-symmetric.

	π		ϕ	τ	Y_k
(i)	$\begin{pmatrix} \alpha+i\beta & 0 \\ 0 & -\alpha+i\beta \end{pmatrix}_{\mathbb{C}}$	$\alpha,\beta>0$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_{\mathbb{C}}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}_{\mathbb{C}}$	T_k
(ii)	$\begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix}$	$\alpha>0$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	T_k
(iii)	$i\beta_{\mathbb{C}}$	$\beta>0$	$1_{\mathbb{C}}$	$e_{\mathbb{C}}$	X_k
(iv)	$\begin{pmatrix} \alpha+i\beta & 0 \\ 0 & -\alpha+i\beta \end{pmatrix}_{\mathbb{C}}$	$\alpha>0$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_{\mathbb{C}}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}_{\mathbb{C}}$	T_k
(v)	$i\beta_{\mathbb{C}}$		$1_{\mathbb{C}}$	$e_{\mathbb{C}}$	X_k
(vi)	$\begin{pmatrix} \alpha+i\beta & 0 \\ 0 & -\alpha+i\beta \end{pmatrix}_{\mathbb{H}}$	$\alpha,\beta>0$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_{\mathbb{H}}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}_{\mathbb{H}}$	T_k
(vii)	$\begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix}_{\mathbb{H}}$	$\alpha>0$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_{\mathbb{H}}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}_{\mathbb{H}}$	T_k
$(viii)$	$i\beta_{\mathbb{H}}$	$\beta>0$	$1_{\mathbb{H}}$	$e_{\mathbb{H}}$	X_k
(ix)	$0_{\mathbb{H}}$		$1_{\mathbb{H}}$	$e_{\mathbb{H}}$	X_k

Table 4: W-sums over $\mathbb{R}$, rows (i)–(iii), over $\mathbb{C}$, rows(iv) and (v), and over $\mathbb{H}$, rows (vi)–(ix); $e=1$ if k even, $e=i$ if k odd

	$Z(\pi)$	$Z(\pi,\tau)$
(i)	$\begin{pmatrix} a_1+ib_1 & 0 \\ 0 & a_2+ib_2 \end{pmatrix}_{\mathbb{C}}$	$\begin{pmatrix} a+ib & 0 \\ 0 & a-ib \end{pmatrix}_{\mathbb{C}}$
(ii)	$\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$	aI_2
(iii)	$a+ib_{\mathbb{C}}$	$a_{\mathbb{C}}$
(iv)	$\begin{pmatrix} a_1+ib_1 & 0 \\ 0 & a_2+ib_2 \end{pmatrix}_{\mathbb{C}}$	$\begin{pmatrix} a+ib & 0 \\ 0 & a-ib \end{pmatrix}_{\mathbb{C}}$
(v)	$a+ib_{\mathbb{C}}$	$a_{\mathbb{C}}$
(vi)	$\begin{pmatrix} a_1+ib_1 & 0 \\ 0 & a_2+ib_2 \end{pmatrix}_{\mathbb{H}}$	$\begin{pmatrix} a+ib & 0 \\ 0 & a-ib \end{pmatrix}_{\mathbb{H}}$
(vii)	$\begin{pmatrix} a_1+ib_1+jc_1+kd_1 & 0 \\ 0 & a_2+ib_2+jc_2+kd_2 \end{pmatrix}_{\mathbb{H}}$	$\begin{pmatrix} a+ib+jc+kd & 0 \\ 0 & a+ib+jc+kd \end{pmatrix}_{\mathbb{H}}$
$(viii)$	$a+ib_{\mathbb{H}}$	$a_{\mathbb{H}}$
(ix)	$a+ib+jc+kd_{\mathbb{H}}$	$f_{\mathbb{H}}$

Table 5: Algebraic data corresponding to the W-sums in Table 4; $f = a$ if k even, $f = a + jc + kd$ if k odd

Suppose that (M, R) is a symplectic pair over $\mathbb{R}$ and that the eigenvalues of M are zero. By the Jordan normal form theorem, M is similar to a matrix $L = L_1 \oplus \cdots \oplus L_q$ where $L_j = N_{k_j}$. Hence $(M, R) \sim (L, S)$ for some $S \in \mathbf{sk}$. Define $Q = Q_1 \oplus \cdots \oplus Q_q$ where $Q_j = X_{k_j}$. Observe that Q is nonsingular, and $LQ + QL^T = 0$ so by Proposition 3.1, $S = GQ$ where $G \in Z(L)$ is a nonsingular matrix satisfying $GQ = -Q^T G^T$.

As usual, assume that $k_1 \geq k_2 \geq \cdots \geq k_q$. Let $c(k)$ be the number of $j's$ for which $k_j = k$.

Lemma 5.7 *Suppose that $G \in Z(L)$ is nonsingular and $GQ = -Q^T G^T$.*

(a) If k_j is odd, then $c(k_j)$ is even.

(b) $G \sim H = H_1 \oplus \cdots \oplus H_r$ where the blocks H_i correspond to the blocks of L: N_{k_i}, k_i even, or $N_{k_i} \oplus N_{k_i}$, k_i odd.

Proof Let $k = k_1$, $c = c(k_1)$ so that

$$k_1 = \cdots = k_c = k, \quad k_{c+1} < k.$$

We claim the following:

1. If k is even we can obtain a nonsingular block of size k in the top-left-hand corner.

2. If k is odd then $c \geq 2$ and we can obtain a nonsingular block of size $2k$ in the top-left-hand corner.

In case 1, Q_1 is skew-symmetric, and in case 2, $Q_1 \oplus Q_1$ is symmetric. In either case, we may apply Lemma 3.2 to obtain a new matrix G that has the same nonsingular block, but with zero entries below and to the right of this block. The lemma follows by induction.

It remains to verify the claim. Even though (L, Q) is not a W-sum, certain results are still valid. In particular, Q is nonsingular, $LQ + QL^T = 0$, and $Q_i^T = (-1)^{k_i - 1} Q_i$. In addition, Proposition 4.7 holds.

When k is even we just follow the proof of Lemma 4.11 (the case $S = \mathbb{R}$ and $\tilde{p} = 1$). This leaves the case k odd. The $k \times k$ blocks G_{ij}, $1 \leq i, j \leq c$, have expansions

$$G_{ij} = \sum_{s=0}^{k-1} g_{ij,s} N_k^s,$$

where $g_{ij,s} \in \mathbb{R}$, $0 \leq s \leq k-1$. Since $Q_1 = \cdots = Q_c$ is symmetric, the equation $GQ = -Q^T G^T$ implies that

$$\begin{aligned} g_{ij,s} N_k^s X_k &= -g_{ji,s} X_k (N_k^s)^T \\ &= (-1)^{s+1} g_{ji,s} N_k^s X_k. \end{aligned}$$

In particular the leading coefficients $g_{ij} = g_{ij,0}$ satisfy

$$g_{ij} = -g_{ji}.$$

Hence $g_{11} = 0$ and G_{11} has zeros in the first column. In addition, for $j > c$, G_{j1} has zeros in the first column. Since G is nonsingular, it follows that $c \geq 2$ and $g_{j1} \neq 0$ for some j, $2 \leq j \leq c$. Without loss of generality, we may assume that $g_{21} \neq 0$. Also we have $g_{12} = -g_{21}$ and $g_{22} = 0$. Define H_{11} to be the block $\begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}$. The matrices $G_{11}, G_{12}, G_{21}, G_{22}$ commute and the determinant of H_{11} is given by $(g_{21})^{2k} \neq 0$. Thus H_{11} is the required nonsingular block of size $2k$. ■

Corollary 5.8 *The symplectic pair (M,R) is equivalent to a direct sum of the normal form summands in rows 5 and 6 of Table 1.*

Proof By Lemma 5.7, $(M,R) \sim (L,GQ)$ where (L,GQ) has summands of the form $(N_{k_i}, G_i X_{k_i})$, k_i even, and $(N_{k_i} \oplus N_{k_i}, G_i(X_{k_i} \oplus X_{k_i}))$, k_i odd. Suppose that k_i is even. Then (N_{k_i}, X_{k_i}) is a W-sum of type II (with $\pi = 0$, $\phi = 1$, $\tau = 1$) and $H_i \sim \pm I_k$. This produces the normal form summand in row 5 of Table 1.

Next suppose that k_i is odd. By Proposition 1.2(f),

$$(N_{k_i} \oplus N_{k_i}, X_{k_i} \oplus X_{k_i}) = (I_2 \otimes N_{k_i}, I_2 \otimes X_{k_i}) \sim (N_{k_i} \otimes I_2, X_{k_i} \otimes I_2).$$

Hence we can replace the original summand by

$$(N_{k_i} \otimes 1_{\mathbb{C}}, G_i X_{k_i} \otimes 1_{\mathbb{C}}) = (N_{k_i} \otimes 1_{\mathbb{C}}, G_i' X_{k_i} \otimes i_{\mathbb{C}}),$$

where $G_i'(I_{k_i} \otimes i_{\mathbb{C}}) = G_i$. But $(N_{k_i} \otimes 1_{\mathbb{C}}, X_{k_i} \otimes i_{\mathbb{C}})$ is a W-sum of type I $\left(\text{with } \pi = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \phi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tau = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right)$. By Theorem 4.6(b), $G' \sim I_{k_i}$ yielding the normal form summand in row 6 of Table 1. ■

It remains to prove uniqueness of the normal form described in Corollary 5.8. Rewrite the summands of (L, Q) in the form (N_{k_i}, X_{k_i}) and $(1_{\mathbb{C}} \otimes N_{k_i}, i_{\mathbb{C}} \otimes X_{k_i})$ and order so that $k_1 \geq \cdots \geq k_r$. Define $Z(L, Q)^0$ to consist of those block-diagonal matrices $H = H_1 \oplus \cdots \oplus H_r$ where $H_i = \rho_i I_{k_i}$ or $1_{\mathbb{C}} \otimes I_{k_i}$ respectively. Let $\mathrm{ind}_k = \sum_{k_i=k} \rho_i$.

Proposition 5.9 *Suppose that $G, H \in Z(L, Q)^0$ and that $G \sim H$. Then $\mathrm{ind}_k(G) = \mathrm{ind}_k(H)$ for all k.*

Proof This has content only for k even. The proof is almost identical to that of Theorem 4.6(c) viewing blocks of the form $(1_{\mathbb{C}} \otimes N_{k_i}, i_{\mathbb{C}} \otimes X_{k_i})$ as four blocks. ∎

6 Canonical symplectic forms

Suppose that R is a nonsingular skew-symmetric real $m \times m$ matrix. Then m is even, $m = 2n$ say, and there is an orthogonal matrix P such that $PRP^T = J_n$ where $J_n = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$. The matrix J_n induces the standard symplectic form ω on $\mathbb{R}^m$ via the equation $\omega(x, y) =< x, J_n y >$ where $< , >$ is the usual inner product. The matrix P transforms the symplectic form induced by R into ω and the symplectic forms are said to be isomorphic.

The corresponding result for symplectic forms that are invariant under the action of a compact Lie group was first stated in Montaldi, Roberts and Stewart [12] and follows from Lie-theoretic results in [10]. The result is also an immediate consequence of Theorem 2.4, as we show below. Of course the Lie-theoretic proof is more direct and intrinsic.

Suppose that Γ is a compact Lie group acting on a vector space V and that R and R' are nonsingular skew-symmetric matrices commuting with Γ, so that $R, R' \in \mathbf{sk}_\Gamma$. The matrices R and R' are *congruent* if there is a matrix P commuting with Γ such that $PRP^T = R'$. If R and R' are congruent, the Γ-invariant symplectic forms induced by R and R' on V are said to be *isomorphic*. Our goal is to define *canonical* representatives in $\mathbf{sk}_\Gamma$ so that every matrix in $\mathbf{sk}_\Gamma$ is congruent to precisely one canonical representative.

Using the results in Subsection 2.1, in particular Proposition 2.1 and Corollary 2.2, we can reduce to a single isotypic component $\mathcal{D}^m$ of dimension m. Recall that $\mathbf{sk} = \mathbf{sk}_{\mathcal{D}^m}$ denotes the set of nonsingular skew-symmetric

$m \times m$ matrices with entries in the division ring $\mathcal{D}$. This time two matrices R and R' are *congruent* if there is a matrix P with entries in $\mathcal{D}$ such that $PRP^T = R'$.

Theorem 6.1 *Suppose that $R \in \mathbf{sk}$. Then R is congruent to precisely one of the following matrices.*

$\mathcal{D} = \mathbb{R}$, $(iI_n)_{\mathbb{C}}$, $n = m/2$.

$\mathcal{D} = \mathbb{C}$, $(iI_p \oplus -iI_q)_{\mathbb{C}}$, $p, q \geq 0$, $p + q = m$.

$\mathcal{D} = \mathbb{H}$, $(iI_m)_{\mathbb{H}}$.

Proof The main observation is that two matrices $R, R' \in \mathbf{sk}$ are congruent if and only if the symplectic pairs $(0, R)$ and $(0, R')$ are equivalent. Hence we need only write down the normal forms to which $(0, R)$ can be equivalent. By Theorem 2.4, such a normal form is a direct sum of normal form summands with modulus/eigenvalue $\mu = 0$ and size $k = 1$. In the case $\mathcal{D} = \mathbb{R}$ there is a single such summand $(0, \rho(iX_1)_{\mathbb{C}})$ where $\rho = 1$. Since $X_1 = -1$ it is convenient to take instead $\rho = -1$ so that the summands take the form $(0, i_{\mathbb{C}})$. Taking a direct sum of summands yields the required form for R.

The case $\mathcal{D} = \mathbb{H}$ is similar. Finally, in the case $\mathcal{D} = \mathbb{C}$ there are additional factors $\rho = \pm 1$ in each summand, and the sum of these is the index $\mathrm{ind}_{0,1}(0, R)$ which is an invariant. Hence the skew-symmetric matrices listed in this case are not congruent. ∎

Definition 6.2 ([12]) A Γ-*symplectic representation* of a compact Lie group Γ on a vector space V consists of an ordinary representation of Γ on V, called the *underlying representation*, and a Γ-invariant symplectic form. Two Γ-symplectic representations on V are *isomorphic* if both the underlying representation and the symplectic form are isomorphic.

In this terminology, Theorem 6.1 takes the following form.

Corollary 6.3 *Suppose that Γ is a compact Lie group acting on V. There exists a symplectic representation with this underlying representation if and only if each real isotypic component has even dimension. In this case, the symplectic representation is determined uniquely up to isomorphism if and only if there are no complex isotypic components.*

Remark 6.4 (a) It follows from Corollary 6.3 that the equivariant version of Darboux's theorem described in [7] is incorrect whenever there are complex isotypic components in the representation of Γ. Theorem 6.1 implies in particular that there are $m+1$ nonisomorphic symplectic forms on each complex isotypic component of dimension m. The two nonisomorphic symplectic representations on $\mathbb{C}$ are called *complex duals* [12].
(b) In the cases $\mathcal{D} = \mathbb{R}$ and $\mathcal{D} = \mathbb{H}$, Theorem 6.1 states that all matrices in **sk** are congruent. Hence in these cases we may choose any matrix in **sk** instead of the one given in the theorem. When $\mathcal{D} = \mathbb{R}$ it is usual to take J_n as the *canonical* nonsingular skew-symmetric matrix. In the remaining cases, there is no precedent for this choice. In Corollary 6.5 below we suggest some choices that are convenient for our purposes.

Corollary 6.5 *If $R \in \mathbf{sk}$ then R is congruent to $(J_r \oplus \rho i I_s)_{\mathcal{D}}$, where $r, s \geq 0$, $2r + s = m$, $\rho = \pm 1$. There is a unique choice of r, s and ρ satisfying the following restrictions: if $\mathcal{D} = \mathbb{R}$, $s = 0$, if $\mathcal{D} = \mathbb{H}$, $s = 0$ or 1 and $\rho = 1$.*

Proof If $\mathcal{D} = \mathbb{R}$ or $\mathbb{H}$, then all matrices in **sk** are congruent by Theorem 6.1 so it suffices to verify that the matrix $(J_r \oplus \rho i I_s)_{\mathcal{D}}$ is indeed skew-symmetric.

Suppose that $\mathcal{D} = \mathbb{C}$ and that $R \in \mathbf{sk}$ is congruent to $(iI_p \oplus -iI_q)_{\mathbb{C}}$. Let $s = |p-q|$, $r = (m-s)/2$ and $\rho = \operatorname{sgn}(p-q)$. Rearranging rows and columns we can transform R into the matrix $(iI_r \oplus -iI_r \oplus \rho i I_s)_{\mathbb{C}}$. Now we use the fact that $(iI_r \oplus -iI_r)_{\mathbb{C}}$ is congruent to $(J_r)_{\mathbb{C}}$. The matrix $P = \frac{1}{\sqrt{2}} \begin{pmatrix} iI_r & I_r \\ I_r & iI_r \end{pmatrix}_{\mathbb{C}}$ defines such a congruency. ■

Suppose that $K = (J_r \oplus \rho i I_s)_{\mathcal{D}}$, and s and ρ satisfy the conditions in Corollary 6.5. Then we say that K and the symplectic form induced on $\mathcal{D}^m$ by K are *canonical over $\mathcal{D}$*. More generally, a Γ symplectic form on V is Γ-*canonical* if the restriction to each isotypic component is canonical.

Proposition 6.6 *Suppose that (M, R) is a normal form summand over $\mathcal{D}$. Let $K = (J_r \oplus \rho i I_s)_{\mathcal{D}}$ be the unique canonical skew-symmetric matrix over $\mathcal{D}$ congruent to R. Then $s = 0$ or 1, and $s = 1$ if and only if the summand lies in row 3 of Table 2 or 3.*

Proof The results for $\mathcal{D} = \mathbb{R}$ and $\mathbb{H}$ follow from the uniqueness in Theorem 6.1 and, in the case of $\mathbb{H}$, a dimension count. In particular, the matrices

$\begin{pmatrix} 0 & -T_k \\ T_k & 0 \end{pmatrix}$ and $\rho(X_{2k})$ are congruent to J_k (over $\mathbb{R}$) so that the same is true over $\mathbb{C}$ if the suffix $\mathbb{C}$ is added to each matrix. Hence we obtain the required result when (M, R) lies in rows 1 and 2 of Table 2.

This leaves the case when (M, R) lies in row 3 of Table 2. The summand has odd size $k = 2\ell + 1$ and

$$R = \rho(iX_k)_{\mathbb{C}} = \rho \begin{pmatrix} 0 & 0 & iX_\ell \\ 0 & (-1)^{\ell-1}i & 0 \\ (-1)^{\ell-1}iX_\ell & 0 & 0 \end{pmatrix}_{\mathbb{C}}.$$

We can rearrange rows and columns so that R has the form

$$R = \rho \begin{pmatrix} 0 & iX_\ell \\ (-1)^{\ell-1}iX_\ell & 0 \end{pmatrix}_{\mathbb{C}} \oplus \rho(-1)^{\ell-1}i_{\mathbb{C}}.$$

Define $P = \begin{pmatrix} I_\ell & 0 \\ 0 & -\rho iX_\ell \end{pmatrix}_{\mathbb{C}} \oplus 1_{\mathbb{C}}$. Then (taking care with minus signs) a computation shows that $PRP^T = (J_\ell \oplus \rho(-1)^{\ell-1}i)_{\mathbb{C}}$. ■

Corollary 6.7 *If (M, R) is a normal form over $\mathcal{D}$ and R is congruent to $K = (J_r \oplus \rho i I_s)_{\mathcal{D}}$ where K is canonical, then (M, R) contains at least s summands from row 3 of Table 2 or 3.*

7 Cyclospectral representations

Suppose that Γ is a compact Lie group acting on V and that $R \in \mathbf{sk}_\Gamma$ is a nonsingular skew-symmetric matrix commuting with the action of Γ.

Definition 7.1 The symplectic representation of Γ is *cyclospectral* if the eigenvalues of matrices $M \in \mathbf{sp}_\Gamma(R)$ are forced to lie on the imaginary axis. The representation is *weakly cyclospectral* if every matrix $M \in \mathbf{sp}_\Gamma(R)$ has some purely imaginary eigenvalues

Remark 7.2 The cyclospectral representations were defined and classified in Montaldi, Roberts and Stewart [12]. They are of importance in the study of linear stability of relative equilibria for equivariant Hamiltonian systems. In particular, if p is a Γ-invariant equilibrium and the action of Γ is cyclospectral, then p is automatically spectrally (and even linearly) stable.

If the action of Γ is weakly cyclospectral, then the equilibrium p need not be spectrally stable, but generically there will exist periodic solutions near p. To see this, observe that the guaranteed purely imaginary eigenvalues will generically be nonzero [5] and nonresonant apart from resonances forced by symmetry [3]. Hence the hypotheses of the equivariant version of the Liapunov Center Theorem are satisfied (see for example [12]).

In this section we obtain classifications of both the cyclospectral and weakly cyclospectral representations as easy consequences of our results in Section 2. Recall from Subsection 2.1 that the space of Γ-equivariant matrices has the block-diagonal structure

$$\mathrm{Hom}_\Gamma(V) \cong \mathrm{Hom}(\mathcal{D}_1^{m_1}) \oplus \cdots \oplus \mathrm{Hom}(\mathcal{D}_\ell^{m_\ell}),$$

where for each j, $\mathcal{D}_j = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ and m_j is the dimension of the isotypic component $\mathcal{D}_j^{m_j}$. The matrix R has the corresponding block-diagonalization $R = R_1 \oplus \cdots \oplus R_\ell$ where $R_j \in \mathbf{sk}_{\mathcal{D}_j^{m_j}}$.

Theorem 7.3 *The representation of Γ is cyclospectral if and only if*

(a) $\mathcal{D}_j \neq \mathbb{R}$ *for each* j,

(b) if $\mathcal{D}_j = \mathbb{H}$ *then* $m_j = 1$, *and*

(c) if $\mathcal{D}_j = \mathbb{C}$ *then* R_j *is congruent to* $\rho(iI_{m_j})_{\mathbb{C}}$ *where* $\rho = \pm 1$.

Theorem 7.4 *The representation of Γ is weakly cyclospectral if and only if there is at least one j, $1 \leq j \leq \ell$ such that either*

(i) $\mathcal{D}_j = \mathbb{H}$ *and* m_j *is odd, or*

(ii) $\mathcal{D}_j = \mathbb{C}$ *and* R_j *is* **not** *congruent to* $(J_{m_j/2})_{\mathbb{C}}$.

Observe that the representation on V is cyclospectral if and only if each of the isotypic components $\mathcal{D}_j^{m_j}$ is cyclospectral. Also the representation is weakly cyclospectral if and only if at least one of the isotypic components is weakly cyclospectral. Hence to prove Theorems 7.3 and 7.4 it is sufficient to prove the following lemma.

Lemma 7.5 *(a) The isotypic component $\mathcal{D}^m$ is cyclospectral if and only if either $\mathcal{D} = \mathbb{H}$ and $m = 1$, or $\mathcal{D} = \mathbb{C}$ and R is congruent to $\rho(iI_m)_{\mathbb{C}}$, $\rho = \pm 1$.*

(b) *The isotypic component $\mathcal{D}_m$ is weakly cyclospectral if and only if either $\mathcal{D} = \mathbb{H}$ and m is odd, or $\mathcal{D} = \mathbb{C}$ and R is* **not** *congruent to $(J_{m/2})_{\mathbb{C}}$.*

Proof By Corollary 6.5, R is congruent to a matrix of the form $(J_r \oplus \rho i I_s)_{\mathcal{D}}$ for a unique choice of integers $r, s \geq 0$ satisfying $2r + s = m$ and $s = 0$ if $\mathcal{D} = \mathbb{R}$, $s = 0$ or 1 if $\mathcal{D} = \mathbb{H}$. We claim that the representation is cyclospectral if and only if $r = 0$ and is weakly cyclospectral if and only if $s \geq 1$. The lemma follows immediately from the claim.

Suppose that $M \in \mathbf{sp}(R)$. Corollary 6.7 states that the normal form of (M, R) must have at least s summands from row 3 of Table 2 or 3. Such summands have purely imaginary eigenvalues proving sufficiency of the conditions on r and s in the claim.

Next we prove necessity of these conditions. By Theorem 6.1 the matrix $S_r = \begin{pmatrix} 0 & -T_r \\ T_r & 0 \end{pmatrix}$ is congruent over $\mathbb{R}$ to J_r. Hence $(S_r)_{\mathcal{D}}$ is congruent over $\mathcal{D}$ to $(J_r)_{\mathcal{D}}$. Thus we may assume that $R = (S_r \oplus \rho i I_s)_{\mathcal{D}}$. Consider the matrix

$$M = \begin{pmatrix} I_r + N_r & 0 \\ 0 & -I_r - N_r \end{pmatrix}_{\mathcal{D}} \oplus (iI_s)_{\mathcal{D}} \in \mathbf{sp}(R).$$

If $r \geq 1$ then the eigenvalues of M include ± 1 and if $s = 0$ these are the only eigenvalues as required. ■

8 Normal forms for infinitesimally symplectic matrices

As a special case of the normal theorem for symplectic pairs, Williamson obtained a list of normal forms for infinitesimally symplectic matrices. In this section we describe the normal forms in the equivariant context. Suppose that Γ is a compact Lie group acting on V and that (M, R) is a Γ-symplectic pair in normal form. By an orthogonal equivariant change of coordinates, we can transform R into a matrix K as described in Corollary 6.5. The matrix K induces on V a canonical Γ-symplectic form. If (M, R) is transformed into (A, K), then A is the required normal form. We say that the Γ-symplectic pair (A, K) is a Γ-*canonical pair.*

As usual we may factor out the symmetry and reduce to a problem over a real division ring $\mathcal{D}$. Accordingly we may assume that (M,R) is a symplectic pair over $\mathcal{D}$ and is in normal form. By Theorem 2.4, (M,R) is a direct sum of the normal form summands listed in Tables 1, 2 and 3. In Lemma 8.1 below we show how to reduce each normal form summand into a canonical pair over $\mathcal{D}$.

Lemma 8.1 *Suppose that (M,R) is a normal form summand in Table 1, 2 or 3. Then $(M,R) \sim (A,K)$ where the canonical pair (A,K) lies in the corresponding row of Table 6, 7 or 8.*

Proof The proof divides into four cases.
Case 1: rows 1 and 2 of Table 1 and row 1 of Tables 2 and 3
We have $R = \begin{pmatrix} 0 & -T_k \\ T_k & 0 \end{pmatrix}_{\mathcal{R}}$ where $\mathcal{R}$ is a real division ring (not necessarily equal to $\mathcal{D}$). Define $P = \begin{pmatrix} I_k & 0 \\ 0 & T_k \end{pmatrix}_{\mathcal{R}}$. Then $PRP^T = (J_k)_{\mathcal{R}}$ and $A = PMP^{-1}$ have the required form.
Case 2: rows 3 and 5 of Table 1 and row 2 of Tables 2 and 3
Since $k = 2\ell$, $X_k = T_\ell \otimes \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. By Proposition 1.2(f), X_k can be transformed by an orthogonal change of coordinates into $\begin{pmatrix} 0 & -T_\ell \\ T_\ell & 0 \end{pmatrix}$. Similarly $N_k = I_\ell \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + N_\ell \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ becomes $\begin{pmatrix} 0 & I_\ell \\ N_\ell & 0 \end{pmatrix}$ and we can write $I_k = \begin{pmatrix} I_\ell & 0 \\ 0 & I_\ell \end{pmatrix}$. The pair

$$M = (i\beta I_k + N_k)_{\mathcal{R}}, \quad R = (\rho X_k)_{\mathcal{R}},$$

is transformed into

$$M = \begin{pmatrix} i\beta I_\ell & I_\ell \\ N_\ell & i\beta I_\ell \end{pmatrix}_{\mathcal{R}}, \quad R = \begin{pmatrix} 0 & -\rho T_\ell \\ \rho T_\ell & 0 \end{pmatrix}_{\mathcal{R}}.$$

Now apply the transformation $P = \begin{pmatrix} I_\ell & 0 \\ 0 & \rho T_\ell \end{pmatrix}_{\mathcal{R}}$.
Case 3: rows 4 and 6 of Table 1
Write $R = \rho(iX_k)_{\mathbb{C}} = \rho X_k \otimes \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and use Proposition 1.2(f) to

transform R into $\begin{pmatrix} 0 & -\rho X_k \\ \rho X_k & 0 \end{pmatrix}$. At the same time $M = (i\beta I_k + N_k)_{\mathbb{C}}$ is transformed into $\begin{pmatrix} N_k & -\beta I_k \\ \beta I_k & N_k \end{pmatrix}$. Let $P = \begin{pmatrix} I_k & 0 \\ 0 & \rho X_k \end{pmatrix}$.

Case 4: row 3 of Tables 2 and 3

In the proof of Proposition 6.6 we described how to transform $R = \rho(iX_k)_{\mathbb{C}}$, where $k = 2\ell$, into $(J_\ell \oplus \rho(-1)^{\ell-1}i)_{\mathbb{C}}$. Applying the same transformations to M produces the matrix $S_{\ell,(-1)^{\ell-1}\rho}$. (The indices ρ in row 3 of Tables 2 and 7 differ by a factor of $(-1)^{\ell-1}$.) The case $\mathcal{D} = \mathbb{H}$ is identical. ■

	A	K	Size	μ	ρ
1	$\begin{pmatrix} \mu I_k + N_k & 0 \\ 0 & -\bar{\mu} I_k - N_k^T \end{pmatrix}_{\mathbb{C}}$	J_{2k}	$k \in \mathbb{N}$	$\alpha + i\beta$	1
2	$\begin{pmatrix} \alpha I_k + N_k & 0 \\ 0 & -\alpha I_k - N_k^T \end{pmatrix}$	J_k	$k \in \mathbb{N}$	α	1
3	$\begin{pmatrix} i\beta I_\ell & \rho T_\ell \\ \rho T_\ell N_\ell & i\beta I_\ell \end{pmatrix}_{\mathbb{C}}$	J_k	$k = 2\ell$	$i\beta$	± 1
4	$\begin{pmatrix} N_k & -\rho\beta X_k \\ \rho\beta X_k & -N_k^T \end{pmatrix}$	J_k	k odd	$i\beta$	± 1
5	$\begin{pmatrix} 0 & \rho T_\ell \\ \rho T_\ell N_\ell & 0 \end{pmatrix}$	J_ℓ	$k = 2\ell$	0	± 1
6	$\begin{pmatrix} N_k & 0 \\ 0 & -N_k^T \end{pmatrix}$	J_k	k odd	0	1

Table 6: Canonical pairs over $\mathbb{R}$ of size k, modulus μ and index ρ; $\alpha, \beta > 0$

The matrix $(S_{\ell,\rho})_{\mathcal{D}}$ that appears in row 3 of Tables 7 and 8 is defined for

$\mathcal{D} = \mathbb{C}$ or $\mathbb{H}$ to be

$$S_{\ell,\rho} = \left(\begin{array}{c|c|c} i\beta I_\ell + N_\ell & 0 & \begin{matrix} 0 \\ \vdots \\ 0 \\ 1 \end{matrix} \\ \hline 0 & i\beta I_\ell - N_\ell^T & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \ \cdots \ 0 & 0 \ \cdots \ 0 \ \ -\rho i & i\beta \end{array}\right)_{\mathcal{D}} .$$

Now suppose that (M, R) is a symplectic pair over $\mathcal{D}$ and that (M, R) is in normal form, with normal form summands (M_i, R_i), $i = 1, \ldots, \ell$. Applying Lemma 8.1, $(M, R) \sim (A, K)$ where (A, K) is a direct sum of canonical summands (A_i, K_i). Unfortunately, the direct sum of canonical pairs is not canonical and further work is required to transform (A, K) into a canonical pair. This is relatively straightforward for summands that do not lie in row 3 of Tables 7 or 8. Suppose that (A, K) has no summands in these rows. Then $K_i = J_{\mathcal{D}} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}_{\mathcal{D}}$ and A_i has the corresponding block form $\begin{pmatrix} B_i & C_i \\ D_i & E_i \end{pmatrix}_{\mathcal{D}}$. The symplectic pair (A, K) has 2ℓ rows and columns of blocks. Rearrange these in the order $1, 3, \ldots, 2\ell - 1; 2, 4, \ldots, 2\ell$. Then K is transformed into $J_{\mathcal{D}}$ and A is transformed into the matrix

$$\left(\begin{array}{cc|cc} B_1 & 0 & C_1 & 0 \\ & \ddots & & \ddots \\ 0 & B_\ell & 0 & C_\ell \\ \hline D_1 & 0 & E_1 & 0 \\ & \ddots & & \ddots \\ 0 & D_\ell & 0 & E_\ell \end{array}\right)_{\mathcal{D}} .$$

	A	K	Size	μ	ρ
1	$\begin{pmatrix} \mu I_k + N_k & 0 \\ 0 & -\bar{\mu} I_k - N_k^T \end{pmatrix}_{\mathbb{C}}$	$(J_k)_{\mathbb{C}}$	$k \in \mathbb{N}$	$\alpha + i\beta$	1
2	$\begin{pmatrix} i\beta I_\ell & \rho T_\ell \\ \rho T_\ell N_\ell & i\beta I_\ell \end{pmatrix}_{\mathbb{C}}$	$(J_\ell)_{\mathbb{C}}$	$k = 2\ell$	$i\beta$	± 1
3	$(S_{\ell,\rho})_{\mathbb{C}}$	$(J_\ell \oplus \rho i)_{\mathbb{C}}$	$k = 2\ell + 1$	$i\beta$	± 1

Table 7: Canonical pairs over $\mathbb{C}$ of size k and modulus μ; $\alpha > 0$, $\beta \in \mathbb{R}$. The index is given by ρ in rows 1 and 2, and by $(-1)^{\ell-1}\rho$ in row 3.

The situation is a great deal more complicated when there are summands from row 3 of Tables 7 and 8 present. Suppose there is one such summand, (A_ℓ, K_ℓ) say. Then $K_\ell = \begin{pmatrix} 0 & -I & 0 \\ I & 0 & 0 \\ 0 & 0 & \rho i \end{pmatrix}_{\mathcal{D}}$. We can regard K as a block matrix with $2\ell + 1$ rows and columns, and applying the same operations as before to all but the last of these rows and columns we transform (A, K) into a canonical pair with $K = (J \oplus \rho i)_{\mathcal{D}}$.

Next suppose that (A_1, K_1), (A_2, K_2) are two summands from row 3 of Table 8 so that $A_j = S_{\ell_j,\rho_j}$ and $K_j = (J_{\ell_j} \oplus \rho_j i)_{\mathbb{H}}$ for $j = 1, 2$. We can combine the two summands to form a canonical pair (A, K) with $K = (J_{\ell_1+\ell_2+1})_{\mathbb{H}}$ as follows. First rearrange rows and columns to transform $K_1 \oplus K_2$ into $(J_{\ell_1} \oplus J_{\ell_2} \oplus \rho_1 i \oplus \rho_2 i)_{\mathbb{H}}$. If $\rho_1 = \rho_2$ then we can apply the transformation $P = (I_{\ell_1+\ell_2+1} \oplus j)_{\mathbb{H}}$ so suppose that $\rho_1 = -\rho_2$. Now apply the transformation $P = \left(I_{\ell_1+\ell_2} \oplus \frac{1}{\sqrt{2}} \begin{pmatrix} \rho_1 i & 1 \\ 1 & \rho_1 i \end{pmatrix} \right)_{\mathbb{H}}$ to obtain the matrix $(J_{\ell_1} \oplus J_{\ell_2} \oplus J_1)_{\mathbb{H}}$. This can be rearranged to produce the canonical matrix $(J_{\ell_1+\ell_2+1})_{\mathbb{H}}$. Applying these transformations to the matrix $A_1 \oplus A_2$ yields the required normal form.

Finally, we consider the case of summands from row 3 of Table 7. Summands of opposite sign ρ (ρ as in Table 7, not the index ρ in Table 2) can be paired off as in the case $\mathcal{D} = \mathbb{H}$. The remaining summands (suppose there are s of them) have the same sign ρ and can be combined in the obvious way to produce a canonical pair (A, K) where $K = (J \oplus \rho i I_s)_{\mathbb{C}}$.

	A	K	Size	μ	ρ
1	$\begin{pmatrix} \mu I_k + N_k & 0 \\ 0 & -\bar{\mu} I_k - N_k^T \end{pmatrix}_{\mathbb{H}}$	$(J_k)_{\mathbb{H}}$	$k \in \mathbb{N}$	$\alpha + i\beta$	1
2	$\begin{pmatrix} i\beta I_\ell & \rho T_\ell \\ \rho T_\ell N_\ell & i\beta I_\ell \end{pmatrix}_{\mathbb{H}}$	$(J_\ell)_{\mathbb{H}}$	$k = 2\ell$	$i\beta$	± 1
3	$(S_{\ell,\rho})_{\mathbb{H}}$	$(J_\ell \oplus \rho i)_{\mathbb{H}}$	$k = 2\ell + 1$	$i\beta$	$\pm 1;\ \beta > 0$ $1;\ \beta = 0$

Table 8: Canonical pairs over $\mathbb{H}$ of size k and modulus μ; $\alpha > 0$, $\beta \geq 0$. The index is given by ρ in rows 1 and 2, and by $(-1)^{\ell-1}\rho$ in row 3.

By way of example, suppose that $\mathcal{D} = \mathbb{C}$ and that (A_1, K_1), (A_2, K_2) are canonical pairs of index 1 from row 3 of Table 7, one of size 5, modulus $2i$ and the other of size 3, modulus $-7i$. Then $K_1 = (J_2 \oplus -i)_{\mathbb{C}}$ and $K_2 = (J_1 \oplus i)_{\mathbb{C}}$. If we perform the transformations described above we end up with the canonical pair (A, K) where $K = (J_4)_{\mathbb{C}}$ and

$$A = \left(\begin{array}{cccc|cccc} 2i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2i & 0 & 1/\sqrt{2} & 0 & 0 & 0 & -i/\sqrt{2} \\ 0 & 0 & -7i & -i/\sqrt{2} & 0 & 0 & 0 & 1/\sqrt{2} \\ 0 & 0 & 0 & -5i/2 & 0 & i/\sqrt{2} & 1/\sqrt{2} & 9/2 \\ \hline 0 & 0 & 0 & 0 & 2i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -7i & 0 \\ 0 & 0 & 0 & -9/2 & 0 & -1/\sqrt{2} & -i/\sqrt{2} & -5i/2 \end{array}\right)_{\mathbb{C}}.$$

A Appendix

In this appendix we prove Proposition 4.7. Recall that L is a direct sum of summands $L_i = I_{k_i} \otimes \pi + N_{k_i} \otimes \phi$ where π and ϕ are $p \times p$ matrices with entries in $\mathcal{D}$ satisfying certain properties, see (H1)–(H5). In particular, π is semisimple, ϕ is nonsingular, and $Z(\pi) \subset Z(\phi)$.

Lemma A.1 *Suppose that g and h are matrices with entries in $\mathcal{D}$ such that*

$$\pi g = g\pi \text{ and } \pi h - h\pi = \phi g.$$

Then $g = 0$.

Proof The condition $Z(\pi) \subset Z(\phi)$ implies in particular that π commutes with ϕ. Replacing h by ϕh and using the fact that ϕ is nonsingular we may reduce to the case where ϕ is the identity.

Regardless of $\mathcal{D}$, we may regard π, g and h as $m \times m$ complex matrices, where $m = p$ if $\mathcal{D} = \mathbb{R}$ or $\mathbb{C}$, and $m = 2p$ if $\mathcal{D} = \mathbb{H}$. Since π is semisimple, we may assume that π is diagonal with entries $\pi_1, \dots, \pi_m$. Writing $g = \{g_{ij}\}$ and $h = \{h_{ij}\}$ we have that

$$(\pi_i - \pi_j)g_{ij} = 0, \quad (\pi_i - \pi_j)h_{ij} = g_{ij}.$$

If $\pi_i \neq \pi_j$ then $g_{ij} = 0$ by the first equation. On the other hand, if $\pi_i = \pi_j$, then $g_{ij} = 0$ by the second equation. Hence each entry $g_{ij} = 0$ as required. ■

Proof of Proposition 4.7 Suppose that $P = \{P_{ij}\} \in Z(L)$. We shall compute the form of P_{ij} for $k_i \leq k_j$ The case $k_i \geq k_j$ can be treated in a similar way. To simplify the notation we set $m = k_i$, $n = k_j$ and denote the entries of P_{ij} by $p_{r,s}$, $r = 1, \dots, m$, $s = 1, \dots, n$. The condition $P_{ij}L_j = L_iP_{ij}$ can be written as

$$\pi p_{r,s} - p_{r,s}\pi = -\phi p_{r+1,s} + p_{r,s-1}\phi, \quad r = 1, \dots, m, \quad s = 1, \dots, n, \tag{A.3}$$

where we have set $p_{m+1,s} = p_{r,0} = 0$.

First, we show that $p_{r,s} = 0$ if $r > s$, and that $p_{r,r} \in Z(\pi)$. We proceed by induction on s. When $s = 1$ it follows from (A.3) and Lemma A.1 that

$$p_{m,1} = p_{m-1,1} = \cdots = p_{2,1} = 0 \quad \text{and} \quad p_{1,1} \in Z(\pi).$$

Now assume that $p_{m,s} = p_{m-1,s} = \cdots = p_{s+1,s} = 0$ and $p_{s,s} \in Z(\pi)$ for $s = 1, \dots, k$. Then by (A.3)

$$\begin{aligned}
\pi p_{m,k+1} - p_{m,k+1}\pi &= 0 \\
\pi p_{m-1,k+1} - p_{m-1,k+1}\pi &= -\phi p_{m,k+1} \\
&\vdots \\
\pi p_{k+2,k+1} - p_{k+2,k+1}\pi &= -\phi p_{k+3,k+1} \\
\pi p_{k+1,k+1} - p_{k+1,k+1}\pi &= -\phi p_{k+2,k+1}
\end{aligned}$$

and applying Lemma A.1 we have that $p_{r,k+1} = 0$ for $r = k+2, \dots, m$, and that $p_{k+1,k+1} \in Z(\pi)$ as required.

Next we apply (A.3) to the case where $s = r+1$ and, using the fact that $Z(\pi) \subset Z(\phi)$, we find that

$$\pi p_{r,r+1} - p_{r,r+1}\pi = \phi(p_{r,r} - p_{r+1,r+1}).$$

Since $p_{r,r}, p_{r+1,r+1} \in Z(\pi)$ we can conclude that $p_{1,1} = p_{2,2} = \cdots = p_{m,m}$. Moreover, $\pi p_{m,m+1} - p_{m,m+1}\pi = p_{m,m}\phi$ and

$$p_{1,1} = p_{2,2} = \ldots = p_{m,m} = 0.$$

Again we argue inductively and assume that $p_{1,s} = p_{2,s+1} = \cdots = p_{m,s+m-1} = 0$ for $s = 1, \dots, k$. Then for $r \leq m$

$$\pi p_{r,k+r} - p_{r,k+r}\pi = -\phi p_{r+1,k+r} + p_{r,k+r-1}\phi = 0$$

and hence $p_{r,k+r} \in Z(\pi)$, $r = 1, \dots, m$. Moreover

$$\pi p_{r-1,k+r} - p_{r-1,k+r}\pi = -\phi p_{r,k+r} + p_{r-1,k+r-1}\phi$$

and therefore $p_{1,k+1} = p_{2,k+2} = \cdots = p_{m,k+m}$. Finally, $\pi p_{m,k+m+1} - p_{m,k+m+1}\pi = p_{m,k+m}\phi$ and we obtain $p_{1,k+1} = p_{2,k+2} = \cdots = p_{m,k+m} = 0$ as desired.

So far we have shown that

$$p_{r,s} = 0 \text{ for } s < r + n - m. \tag{A.4}$$

In particular, part (a) of Proposition 4.7 is proved. It remains to show that the matrix F_{ij} defined in Proposition 4.7 has the required structure. Again we proceed inductively and show first that

$$p_{1,n-m+1} = p_{2,n-m+2} = \cdots = p_{m,n} \in Z(\pi).$$

By (A.3) and (A.4)

$$\begin{aligned} \pi p_{r,n-m+r} - p_{r,n-m+r}\pi &= -\phi p_{r+1,n-m+r} + p_{r,n-m+r-1}\phi \\ &= \phi(p_{r,r+(n-m-1)} - p_{r+1,r+1+(n-m-1)}) \\ &= 0 \end{aligned}$$

and therefore $p_{r,n-m+r} \in Z(\pi)$, $r = 1, \dots, m$. Moreover,

$$\pi p_{r,n-m+r+1} - p_{r,n-m+r+1}\pi = \phi(p_{r,n-m+r} - p_{r+1,n-m+r+1}), \quad r = 1, \dots, m-1$$

and hence $p_{r,n-m+r} = p_{r+1,n-m+r+1}$ by Lemma A.1.

Now assume that

$$p_{1,n-m+s} = p_{2,n-m+s+1} = \cdots = p_{m+1-s,n} \in Z(\pi)$$

for $s = 1, \ldots, k-1$. We have to show that $p_{1,n-m+k} = p_{2,n-m+k+1} = \cdots = p_{m+1-k,n} \in Z(\pi)$. We have for $r = 1, \ldots, m+1-k$

$$\begin{aligned} \pi p_{r,n-m+k+r-1} - p_{r,n-m+k+r-1}\pi &= -\phi p_{r+1,n-m+k+r-1} + p_{r,n-m+k+r-2}\phi \\ &= \phi(p_{r,n-m+(k-1)+r-1} - p_{r+1,n-m+(k-1)+(r+1)-1}) \\ &= 0 \end{aligned}$$

and therefore $p_{r,n-m+k+r-1} \in Z(\pi)$. Finally,

$$\pi p_{r,n-m+k+r} - p_{r,n-m+k+r}\pi = \phi(p_{r,n-m+k+r-1} - p_{r+1,n-m+k+r})$$

from which we can conclude that $p_{r,n-m+k+r-1} = p_{r+1,n-m+k+r}$. ■

References

[1] N. Burgoyne and R. Cushman. Normal forms for real linear Hamiltonian systems, in 'The 1976 NASA Conference on Geometric Control Theory' pp. 483-529, Math. Sci. Press, Brookline, Mass., 1977.

[2] N. Burgoyne and R. Cushman. Conjugacy classes in linear groups, *J. Algebra* **44** (1977) 339-362.

[3] M. Dellnitz, I. Melbourne and J.E. Marsden. Generic bifurcation of Hamiltonian vector fields with symmetry, *Nonlinearity* **5** (1992) 979-996.

[4] D.M. Galin. Versal deformations of linear Hamiltonian systems, *AMS Transl. (2)* **118** (1982) 1-12. (*Trudy Sem. Petrovsk.* **1** (1975) 63-74.)

[5] M. Golubitsky and I. Stewart. Generic bifurcation of Hamiltonian systems with symmetry, *Physica D* **24** (1987) 391-405.

[6] M. Golubitsky, I. Stewart and D. Schaeffer. *Singularities and Groups in Bifurcation Theory.* Vol. 2, Springer 1988.

[7] V. Guillemin and S. Sternberg. *Symplectic Techniques in Physics.* Camb. Univ. Press, 1984.

[8] H. Koçak. Normal forms and versal deformations of linear Hamiltonian systems, *J. Diff. Eq.* **51** (1984) 359-407.

[9] A.J. Laub and K. Meyer. Canonical forms for symplectic and Hamiltonian matrices, *Celestial Mech.* **9** (1974) 213-238.

[10] I. MacDonald. Algebraic structure of Lie groups, in *Representation Theory of Lie Groups* (ed. M.F. Atiyah), London Math. Soc. Lecture Notes **34**, Camb. Univ. Press, 1979.

[11] I. Melbourne. Versal unfoldings of equivariant linear Hamiltonian vector fields. *Math. Proc. Camb. Phil. Soc.* **114** (1993) 559-573.

[12] J. Montaldi, M. Roberts and I. Stewart. Periodic solutions near equilibria of symmetric Hamiltonian systems, *Phil. Trans. R. Soc.* **325** (1988) 237-293.

[13] G.E. Wall. On the conjugacy classes in the unitary, symplectic and orthogonal groups, *J. Austral. Math. Soc.* **3** (1963) 1-62.

[14] Y.H. Wan. Normal forms for infinitesimally symplectic transformations with involutions. Preprint, SUNY at Buffalo, 1989.

[15] Y.H. Wan. Versal deformations of infinitesimally symplectic transformations with involutions. Preprint, SUNY at Buffalo, 1989.

[16] N.A. Wiegmann. Some theorems on matrices with real quaternionic elements, *Canad. J. Math.* **7** (1955) 191-201.

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