Canad. Math. Bull. Vol. **48** (3), 2005 pp. 394–404

Diagonal Plus Tridiagonal Representatives for Symplectic Congruence Classes of Symmetric Matrices

D. Ž. Đoković, F. Szechtman and K. Zhao

Abstract. Let $n = 2m$ be even and denote by $\mathrm{Sp}_n(F)$ the symplectic group of rank m over an infinite field F of characteristic different from 2. We show that any $n \times n$ symmetric matrix A is equivalent under symplectic congruence transformations to the direct sum of $m \times m$ matrices B and C, with B diagonal and C tridiagonal. Since the $\mathrm{Sp}_n(F)$-module of symmetric $n \times n$ matrices over F is isomorphic to the adjoint module $\mathfrak{sp}_n(F)$, we infer that any adjoint orbit of $\mathrm{Sp}_n(F)$ in $\mathfrak{sp}_n(F)$ has a representative in the sum of $3m - 1$ root spaces, which we explicitly determine.

1 Introduction

A *bilinear space* over a field F is an ordered pair (V, φ) consisting of a finite-dimensional F-vector space V and a bilinear form $\varphi\colon V \times V \to F$. In this context the first basic question is to decide when two bilinear spaces, say (V, φ) and (W, ψ), are *equivalent*, *i.e.*, they admit a linear isomorphism $f\colon V \to W$ satisfying $\psi(f(x), f(y)) = \varphi(x, y)$ for all $x, y \in V$. P. Gabriel [7] showed how to reduce this problem to the non-degenerate case. Work by J. Williamson [12] and C. Riehm [9] reduces the latter to the equivalence of quadratic and hermitian forms and the similarity of matrices. The equivalence problem of sesquilinear forms and its relationship to the determination of conjugacy classes in the classical linear groups is considered by G. E. Wall in [11].

Let us mention one interesting application of the structure theory of bilinear spaces. If (V, φ) is a bilinear space, define its *transpose* to be the bilinear space (V, φ') where φ' is defined by $\varphi'(x, y) = \varphi(y, x)$ for all $x, y \in V$. It is a rather non-trivial fact that the bilinear spaces (V, φ) and (V, φ') are equivalent. Moreover, the corresponding form-preserving linear isomorphism can be chosen to be an involution. This was first proved by R. Gow [8] under the restriction that φ is non-degenerate, but the extension to the general case follows at once from Gabriel's work. An alternative proof is given in [3]. An algorithm to compute the aforementioned involutory isomorphism is presented in [6].

From now on we shall assume that the characteristic of F is not 2. Let (V, φ) be a bilinear space. Given a basis $\mathfrak{B}$ of V, we write $M_{\mathfrak{B}}(\varphi)$ for the Gram matrix of φ

Received by the editors October 3, 2003; revised May 4, 2004.
The first author was supported in part by the NSERC Grant A-5285 and the third by the NSF of China (Grants 10371120 and 10431040).
AMS subject classification: Primary: 11E39; secondary: 15A63, 17B20.

relative to $\mathfrak{B}$. We define the symmetric and alternating parts of φ to be the forms

$$\varphi^+ = (\varphi + \varphi')/2 \quad \text{and} \quad \varphi^- = (\varphi - \varphi')/2.$$

In this paper we shall deal exclusively with bilinear spaces whose associated alternating form is non-degenerate. This implies that the dimension of our bilinear spaces is even, say $n = 2m$. Let (V, φ) and (W, ψ) be such bilinear spaces. Let us write J_m, or simply J, for the matrix

$$\begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}.$$

The symplectic group $\mathrm{Sp}_n(F)$ is the stabilizer of J under the action of $\mathrm{GL}_n(F)$ by congruence transformations, *i.e.*,

$$\mathrm{Sp}_n(F) = \{X \in \mathrm{GL}_n(F) : X'JX = J\}.$$

(Here X' denotes the transpose of X.) It is well known that V and W admit bases $\mathfrak{B}$ and $\mathfrak{C}$ such that

$$M_{\mathfrak{B}}(\varphi^-) = J = M_{\mathfrak{C}}(\psi^-).$$

It follows that (V, φ) and (W, ψ) are equivalent if and only if the symmetric matrices $M_{\mathfrak{B}}(\varphi^+)$ and $M_{\mathfrak{C}}(\psi^+)$ are equivalent under the congruence action of $\mathrm{Sp}_n(F)$.

Thus we are led to the study of normal forms of symmetric matrices under symplectic congruence. This matrix problem is exactly the one considered and solved by J. Williamson [12]. Although he worked in characteristic 0, his results and proofs remain valid for any field of characteristic different from 2. Explicit normal forms in the case of the real field were computed by D. M. Galin (see [2, Appendix 6]).

The present work aims at further contributing to this theory by exhibiting representatives for the congruence action of $\mathrm{Sp}_n(F)$ on symmetric matrices which are very much unlike any others previously considered.

We point out that the module, $\mathrm{Sym}_n(F)$, of symmetric n by n matrices over F under the congruence action of $\mathrm{Sp}_n(F)$ is isomorphic to the adjoint module

$$\mathfrak{sp}_n(F) = \{Z \in M_n(F) : Z'J_m + J_m Z = 0\}.$$

An explicit isomorphism from the former to the latter is given by $A \to Z = J_m A$. Hence, the problem solved by Williamson is equivalent to the problem of finding normal forms for the adjoint action of $\mathrm{Sp}_n(F)$ on its Lie algebra $\mathfrak{sp}_n(F)$. In the case of the real field, one can consult [4] or [5] for the description of normal forms. The case when Z is nilpotent has been described in detail, for arbitrary F of characteristic different from 2, by Springer and Steinberg [10, Chapter IV].

Let us consider the $\mathbf{Z}_2$-gradation $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1$ of $\mathfrak{g} = \mathfrak{sp}_n(F)$, where

$$\mathfrak{g}_0 = \left\{ \begin{pmatrix} X & 0 \\ 0 & -X' \end{pmatrix} : X \in M_m(F) \right\},$$

and

$$\mathfrak{g}_1 = \left\{ \begin{pmatrix} 0 & Y \\ Z & 0 \end{pmatrix} : X, Y \in \mathrm{Sym}_m(F) \right\}.$$

https://doi.org/10.4153/CMB-2005-036-x Published online by Cambridge University Press

(Here $M_n(F)$ is the algebra of all $n \times n$ matrices over F.)

By generalizing a special case of a theorem of L. V. Antonyan [1, Theorem 2], we obtained the following result in [6] (see Theorem 1.6, part (i)).

Theorem 1.1 *Let F be any field of characteristic different from 2 and let $n = 2m$ be even. Then every adjoint orbit of $\mathrm{Sp}_n(F)$ in $\mathfrak{g} = \mathfrak{sp}_n(F)$ meets the subspace $\mathfrak{g}_1$. Equivalently, for any $A \in \mathrm{Sym}_n(F)$, there exists $X \in \mathrm{Sp}_n(F)$ such that*

$$X'AX = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix},$$

where $B, C \in \mathrm{Sym}_m(F)$.

It is obvious that in this theorem we can further require B or C to be a diagonal matrix. This is a consequence of the fact that every symmetric matrix over F is congruent to a diagonal one and the observation that $X \oplus (X')^{-1} \in \mathrm{Sp}_n(F)$ for $X \in \mathrm{GL}_m(F)$. It is not possible to demand that both B and C above are always diagonal. However, it is pleasantly surprising that, under the further assumption that F is infinite, we can make B diagonal and C tridiagonal.

Theorem 1.2 *Let F be an infinite field of characteristic different from 2, and A a symmetric matrix of size $n = 2m$. Then A is congruent under the symplectic group $\mathrm{Sp}_n(F)$ to the direct sum of a diagonal $m \times m$ matrix B and a tridiagonal $m \times m$ matrix C.*

It may be of interest to reformulate this theorem in terms of the adjoint module $\mathfrak{g} = \mathfrak{sp}_n(F)$. The subspace $\mathfrak{h} \subset \mathfrak{g}$ consisting of the diagonal matrices

$$\mathrm{diag}(\xi_1, \xi_2, \dots, \xi_m, -\xi_1, -\xi_2, \dots, -\xi_m), \quad \xi_1, \xi_2, \dots, \xi_m \in F$$

is a Cartan subalgebra of $\mathfrak{g}$. Let $\varepsilon_i \colon \mathfrak{h} \to F$ be the linear function which takes the value ξ_i at the above diagonal matrix. The root system of $(\mathfrak{g}, \mathfrak{h})$ consists of the functions $\pm\varepsilon_i \pm \varepsilon_j$ for $1 \le i < j \le m$ and the functions $\pm 2\varepsilon_i$ for $1 \le i \le m$. The roots

$$\alpha_1 = \varepsilon_1 - \varepsilon_2,\ \alpha_2 = \varepsilon_2 - \varepsilon_3, \dots, \alpha_{m-1} = \varepsilon_{m-1} - \varepsilon_m, \alpha_m = 2\varepsilon_m$$

form a base of this root system (of type C_m).

We denote by $\mathfrak{g}^\alpha$ the (1-dimensional) root space corresponding to the root α:

$$\mathfrak{g}^\alpha = \{Z \in \mathfrak{g} : [H, Z] = \alpha(H)Z,\ \forall H \in \mathfrak{h}\}.$$

If $A = B \oplus C$ where $B, C \in \mathrm{Sym}_m(F)$, with B diagonal and C tridiagonal, then the matrix $Z = J_m A \in \mathfrak{g}$ lies in the sum of root spaces corresponding to the roots

$$\pm 2\alpha_1,\ \pm 2\alpha_2, \dots, \pm 2\alpha_m$$

and

$$\alpha_1 + \alpha_2,\ \alpha_2 + \alpha_3, \dots, \alpha_{m-1} + \alpha_m.$$

Hence, the above theorem can be reformulated as follows.

https://doi.org/10.4153/CMB-2005-036-x Published online by Cambridge University Press

Theorem 1.3 *For any infinite field F of characteristic different from 2, any adjoint orbit of $\mathrm{Sp}_n(F)$ in $\mathfrak{g} = \mathfrak{sp}_n(F)$ meets the sum of the root spaces $\mathfrak{g}^\alpha$ where α runs through the $3m - 1$ roots listed above.*

2 Preliminary Results

This section gathers subsidiary results to our main theorem. Recall that a square matrix over F is said to be indecomposable if it is not congruent to the direct sum of square matrices over F of smaller size. Let * be the involution of $F[t]$ which sends $t \mapsto -t$. Thus if $f \in F[t]$, then $f^*(t) = f(-t)$. A proof for the following theorem may be extracted from [10, Chapter IV]. More precisely, the assertions (a) and (b) are well known (and easy to prove) and for (c) and (d) see [10, 2.19 and Corollary 2.20, p. 259].

Theorem 2.1 *Let F be a field of characteristic different from 2, let $n = 2m$ be an even positive integer, and let $A \in \mathrm{Sym}_n(F)$. Assume that $J + A$ is indecomposable (under congruence) and set $Z = JA$. Then one of the following holds:*

(a) *Z is invertible and has only one elementary divisor. This divisor is of the form f^k where f is monic, irreducible, and $f^* = f$.*
(b) *Z is invertible and has two elementary divisors. These divisors have the form f^k and $(f^*)^k$, where f is monic, irreducible, and $f^* \neq f$.*
(c) *Z is nilpotent and has only one elementary divisor, t^n.*
(d) *The integer m is odd and Z is nilpotent with exactly two elementary divisors, t^m and t^m. Moreover, if $B \in \mathrm{Sym}_n(F)$ and the elementary divisors of $Y = JB$ are t^m and t^m then $XYX^{-1} = Z$ for some $X \in \mathrm{Sp}_n(F)$.*

Our requirement that $\operatorname{char} F \neq 2$ is not needed in the next result.

Lemma 2.2

Let $\langle\ ,\ \rangle \colon V \times V \to F$ be a symmetric bilinear form. Let $v_1, \dots, v_s$ be vectors in V. Define the vectors $e_1, \dots, e_s$ of V by expanding the following "determinants" along the last column:

$$e_1 = v_1,\ e_2 = \begin{vmatrix} \langle v_1, v_1\rangle & v_1 \\ \langle v_2, v_1\rangle & v_2 \end{vmatrix}, \dots, e_s = \begin{vmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \dots & \langle v_1, v_{s-1}\rangle & v_1 \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle & \dots\, . & \langle v_2, v_{s-1}\rangle & v_2 \\ \vdots & \vdots & & \vdots & \vdots \\ \langle v_s, v_1\rangle & \langle v_s, v_2\rangle & \dots & \langle v_s, v_{s-1}\rangle & v_s \end{vmatrix}.$$

Then $e_1, \dots, e_s$ are orthogonal. Moreover, $\mathrm{span}\{e_1, \dots, e_s\} = \mathrm{span}\{v_1, \dots, v_s\}$ provided each of

$$\langle v_1, v_1\rangle,\ \begin{vmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle \end{vmatrix}, \dots, \begin{vmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \dots & \langle v_1, v_{s-1}\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle & \dots & \langle v_2, v_{s-1}\rangle \\ \vdots & \vdots & & \vdots \\ \langle v_{s-1}, v_1\rangle & \langle v_{s-1}, v_2\rangle & \dots & \langle v_{s-1}, v_{s-1}\rangle \end{vmatrix}$$

is nonzero.

Proof Since $\text{span}\{e_1, \dots, e_i\} \subseteq \text{span}\{v_1, \dots, v_i\}$ for all $1 \le i \le s$, it suffices to verify $\langle v_1, e_i\rangle = \dots = \langle v_{i-1}, e_i\rangle = 0$ for all $2 \le i \le s$ in order to show the orthogonality of $e_1, \dots, e_s$. If $1 \le j < i \le s$ then $\langle v_j, e_i\rangle$ is the determinant of the $i \times i$ matrix defining e_i, except that its last column must be replaced by column j. Having two equal columns, this determinant is 0. As the last assertion of the lemma is obvious, the proof is complete. ∎

We adopt the following conventions for the remainder of the paper. We fix a matrix $A \in \text{Sym}_n(F)$ and define the symmetric and alternating bilinear forms $\langle\ ,\ \rangle$ and $\langle\ ,\ \rangle_1$ on the column space $V = F^n$, by means of

$$\langle x, y\rangle = x'Ay, \quad \langle x, y\rangle_1 = x'Jy, \quad x, y \in V.$$

Set $Z = JA$. We view V as module over the polynomial algebra $F[t]$ by letting t act as left multiplication by Z. Observe that

$$AZ = -Z'A \text{ and } JZ = -A, \tag{2.1}$$

or equivalently

$$\langle x, ty\rangle = -\langle tx, y\rangle, \quad x, y \in V, \tag{2.2}$$

$$\langle x, ty\rangle_1 = -\langle x, y\rangle, \quad x, y \in V. \tag{2.3}$$

Repeated application of (2.2) yields

$$\langle x, py\rangle = \langle p^* x, y\rangle, \quad x, y \in V, p \in F[t]. \tag{2.4}$$

Observe the following immediate consequence of (2.1) and Theorem 2.1: the characteristic polynomial, say f, of Z is even, *i.e.*, $f^* = f$.

Let us say that $v \in V$ is a *generator* if it generates V as $F[t]$-module.

Lemma 2.3 *Suppose that the characteristic and minimal polynomials of Z are equal, say to f. Given $v \in V$, let $W = \text{span}\{v, Z^2 v, \dots, Z^{2(m-1)} v\}$. Then*

(i) $\langle W, ZW\rangle = 0$.
(ii) *The vector v is a generator if and only if $\dim W = m$ and the restriction of $\langle\ ,\ \rangle$ to W is non-degenerate.*
(iii) *If v is a generator, then $V = W \oplus ZW$, $\langle W, ZW\rangle = 0$, $\langle W, W\rangle_1 = 0$ and $\langle ZW, ZW\rangle_1 = 0$.*

Proof By (2.2)

$$\langle p(t^2)v, tq(t^2)v\rangle = -\langle tp(t^2)v, q(t^2)v\rangle$$

for all $p, q \in F[t]$. On the other hand, the symmetry of $\langle\ ,\ \rangle$ and (2.4) give

$$\langle p(t^2)v, tq(t^2)v\rangle = \langle tq(t^2)v, p(t^2)v\rangle = \langle tp(t^2)v, q(t^2)v\rangle.$$

https://doi.org/10.4153/CMB-2005-036-x Published online by Cambridge University Press

Since the characteristic of F is not 2, (i) is established.

We next turn our attention to (ii). If v is a generator, then $V = W \oplus ZW$, which by (i) is an orthogonal decomposition of $(V, \langle\ ,\ \rangle)$. Hence $\dim(W) = m$ and $\ker(Z) \cap W = 0$. As $\ker(Z)$ coincides with the radical of $\langle\ ,\ \rangle$, the restriction of this form to W is non-degenerate.

Suppose conversely that $\dim W = m$ and the restriction of $\langle\ ,\ \rangle$ to W is non-degenerate. Then $W \cap ZW = 0$ by (i). If $w \in W$ and $Zw = 0$ then $\langle W, w\rangle = 0$, so $w = 0$. Thus $\dim ZW = m$, whence $V = W \oplus ZW$, *i.e.*, v is a generator.

To prove (iii), suppose that v is a generator. By (i) and (ii), we know that $V = W \oplus ZW$ is an orthogonal decomposition of $(V, \langle\ ,\ \rangle)$. Since f is even, W and ZW are Z^2-invariant. Let i, j be non-negative integers, with $j > 0$. Then by (2.3), (2.4) and (i),

$$\langle t^{2i}v, t^{2j}v\rangle_1 = -\langle t^{2i}v, t^{2j-1}v\rangle = 0.$$

It follows that $\langle W, W\rangle_1 = 0$. We may now use (2.3) to infer that $\langle ZW, ZW\rangle_1 = 0$, as well. ∎

Proposition 2.4 *Let $A \in \mathrm{Sym}_n(F)$ and suppose the characteristic and minimal polynomials of $Z = JA$ coincide. Denote the field of rational functions on algebraically independent variables $x_1, \dots, x_n$ over F by $K = F(x_1, \dots, x_n)$. Write also $\langle\ ,\ \rangle$ for the bilinear form on the column space K^n, defined by*

$$\langle y, z\rangle = y'Az, \quad y, z \in K^n.$$

Let $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$. Then $\langle\ ,\ \rangle$ has non-degenerate restriction to the subspaces $\mathrm{span}\{x\}, \mathrm{span}\{x, Z^2x\}, \dots, \mathrm{span}\{x, Z^2x, \dots, Z^{2(m-1)}x\}$. *In other words, denoting by $P_k \in F[x_1, \dots, x_n]$ the determinant of the $k \times k$ left upper corner of*

$$M = \begin{pmatrix} \langle x, x\rangle & \langle x, Z^2x\rangle & \cdots & \langle x, Z^{2(m-1)}x\rangle \\ \langle Z^2x, x\rangle & \langle Z^2x, Z^2x\rangle & \cdots & \langle Z^2x, Z^{2(m-1)}x\rangle \\ \vdots & \vdots & & \vdots \\ \langle Z^{2(m-1)}x, x\rangle & \langle Z^{2(m-1)}x, Z^2x\rangle & \cdots & \langle Z^{2(m-1)}x, Z^{2(m-1)}x\rangle \end{pmatrix},$$

then P_k is not zero for all k, $1 \le k \le m$.

Proof By hypothesis there exists v in V which generates V as an $F[Z]$-module. Thus by Lemma 2.3, the matrix obtained from M by making the substitution $x \to v$ is invertible. Therefore $\det M = P_m$ is not zero. Clearly P_1 is not zero either.

In order to show the other P_k are not zero, it will be convenient to replace M by a similar matrix. Consider the matrix $N \in M_m(F[x_1, \dots, x_n])$, defined by

$$N = \begin{pmatrix} \langle x, x\rangle & \langle Zx, Zx\rangle & \cdots & \langle Z^{m-1}x, Z^{m-1}x\rangle \\ \langle Zx, Zx\rangle & \langle Z^2x, Z^2x\rangle & \cdots & \langle Z^mx, Z^mx\rangle \\ \vdots & \vdots & & \vdots \\ \langle Z^{m-1}x, Z^{m-1}x\rangle & \langle Z^mx, Z^mx\rangle & \cdots & \langle Z^{2(m-1)}x, Z^{2(m-1)}x\rangle \end{pmatrix}.$$

https://doi.org/10.4153/CMB-2005-036-x Published online by Cambridge University Press

Denote by $M(k)$ and $N(k)$ the $k \times k$ left upper corners of M and N, respectively. Let $D(k)$ be the $k \times k$ diagonal matrix with alternating diagonal entries 1 and -1, in this order. Since $AZ = -Z'A$, we have as before,

$$\langle y, Zz\rangle = -\langle Zy, z\rangle, \quad y, z \in K^n.$$

It follows easily that

$$D(k)M(k)D(k) = N(k), \quad 1 \le k \le m.$$

Thus it suffices to show the polynomials $R_k = \det N(k)$ are non-zero.

Given k, $1 \le k \le m$, and l, $1 \le l \le m - k + 1$, let $N(k, l)$ be the $k \times k$ submatrix of N in the intersection of rows $1, \dots, k$ and columns $l, \dots, l + k - 1$. Write $R_{k,l} \in F[x_1, \dots, x_n]$ for the determinant of $N(k, l)$. Thus $R_{k,1}$ is nothing but R_k. In fact, the very definition of N yields the following fundamental relation:

$$R_{k,l}(x) = R_k(Z^{l-1}x). \tag{2.5}$$

Suppose our result is false and choose k as small as possible satisfying $R_k = 0$. We know that $k > 1$. As R_{k-1} is not zero, it follows that the first $k - 1$ rows of $N(k)$ are linearly independent over K and row k is a linear combination of the preceding rows. Thus, by subtracting from row k a suitable linear combination of rows $1, \dots, k - 1$ we may transform the first k entries on row k of N to zero.

Denote this new matrix by P. Since N is invertible and P is row equivalent to N, some entry along row k of P must be non-zero. Suppose this entry, say a, occurs in the position s, where s is as small as possible.

For $l = s-k+1$, consider the submatrix T of P in the intersections of rows $1, \dots, k$ and columns $l, \dots, l + k - 1 = s$. Its last row is $0, \dots, 0, a$ and the intersection of its first $k - 1$ rows and columns constitute the matrix $N(k - 1, l)$. Since R_{k-1} is not zero, it follows from (2.5) that the determinant of $N(k - 1, l)$ is not zero. Hence T is invertible. But T is row equivalent to $N(k, l)$, which by (2.5) is singular. This contradiction completes the proof. ■

3 Main Result

We are ready to prove our main result.

Theorem 3.1 *Let F be an infinite field of characteristic different from 2, let $n = 2m$ be an even integer, and $A \in \mathrm{Sym}_n(F)$. Then there exists $X \in \mathrm{Sp}_n(F)$ and $B, C \in \mathrm{Sym}_m(F)$ such that B is diagonal, C is tridiagonal, and*

$$X'AX = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}. \tag{3.1}$$

https://doi.org/10.4153/CMB-2005-036-x Published online by Cambridge University Press

Proof Clearly, without any loss of generality, we may assume that $J_m + A$ is indecomposable. Let $V = F^n$ be the space of column vectors. We make V into an $F[t]$-module by letting t act as the matrix $Z = J_m A$. By Theorem 2.1, there are four possibilities, (a)–(d), for the matrix Z.

We assume first that V is cyclic as an $F[t]$-module, *i.e.,* we exclude case (d).

Let $P_1, \ldots, P_m$ be the non-zero polynomials defined in Proposition 2.4. Consider their product $P = P_1 \cdots P_m$. Since F is infinite and P is not zero, there exists $v \in F^n$ such that $P(v) \neq 0$, that is, none of $P_1, \ldots, P_m$ vanish on v.

Consider the subspace W of V spanned by the vectors $v, Z^2 v, \ldots, Z^{2(m-1)} v$. Since $P_m(v) \neq 0$, Lemma 2.3 ensures that

$$V = W \oplus ZW, \quad \langle W, ZW \rangle = 0, \quad \langle W, W \rangle_1 = 0, \quad \langle ZW, ZW \rangle_1 = 0,$$

with the restriction of $\langle\ ,\ \rangle$ to W non-degenerate.

Construct the vectors $e_1, \ldots, e_m$ by applying Lemma 2.2 to the form $\langle\ ,\ \rangle$ and the vectors $v, Z^2 v, \ldots, Z^{2(m-1)} v$. It follows that $e_1, \ldots, e_m$ are orthogonal relative to $\langle\ ,\ \rangle$. As the restriction of $\langle\ ,\ \rangle$ to W is non-degenerate, the vectors $e_1, \ldots, e_m$ are non-isotropic. Since none of $P_1, \ldots, P_{m-1}$ vanish on v, Lemma 2.2 ensures that $e_1, \ldots, e_m$ span W. Note that the matrix B of the restriction of $\langle\ ,\ \rangle$ to W is diagonal and nonsingular.

Let $f_i = -\langle e_i, e_i \rangle^{-1} Z e_i$, for $1 \leq i \leq m$. It is not difficult to see at this point that $\{e_1, \ldots, e_m, f_1, \ldots, f_m\}$ must be a symplectic basis of V.

We still have to show that the matrix C of the restriction of $\langle\ ,\ \rangle$ to ZW is tridiagonal. This means that $\langle Ze_i, Ze_j \rangle = 0$ if $j - i > 1$. Suppose that $j - i > 1$. We have

$$\langle Ze_i, Ze_j \rangle = -\langle Z^2 e_i, e_j \rangle.$$

Since $e_i \in \operatorname{span}\{v, Z^2 v, \ldots, Z^{2(i-1)} v\}$, it follows that

$$Z^2 e_i \in \operatorname{span}\{v, Z^2 v, \ldots, Z^{2(i-1)} v, Z^{2i} v\}.$$

At this crucial point we recall that none of $P_1, \ldots, P_i$ vanishes at v. Hence, by Lemma 2.2,

$$\operatorname{span}\{v, Z^2 v, \ldots, Z^{2(i-1)} v, Z^{2i} v\} = \operatorname{span}\{e_1, e_2, \ldots, e_i, e_{i+1}\}.$$

But e_j is orthogonal to $e_1, \ldots, e_i, e_{i+1}$ relative to $\langle\ ,\ \rangle$, whence $\langle Ze_i, Ze_j \rangle = 0$.

It remains to consider case (d) of Theorem 2.1.

As m is odd, we set $m = 2s+1$. It is well known that there exists a nilpotent matrix $Y_1 \in \mathfrak{sp}_{2s}(F)$ of rank $2s - 1$ (a principal nilpotent element of this Lie algebra). For instance, we can take

$$Y_1 = \begin{bmatrix} Y_{11} & Y_{12} \\ 0 & -Y_{11}' \end{bmatrix},$$

where Y_{11} is the upper triangular nilpotent Jordan block of size s and all the entries of Y_{12} are 0 except the one in the lower right hand corner which is equal to 1.

https://doi.org/10.4153/CMB-2005-036-x Published online by Cambridge University Press

Set $A_1 = -J_sY_1$. Assume that $J_s + A_1$ is decomposable. Then there exists an invertible matrix X such that

$$X'(J_s + A_1)X = (J_p + A_{11}) \oplus (J_q + A_{12})$$

for some positive integers p, q (with $p + q = s$) and some symmetric matrices A_{11} and A_{12}. We have

$$X'J_sX = J_p \oplus J_q, \quad X'A_1X = A_{11} \oplus A_{12}.$$

We obtain that

$$Y_1 = -J_s^{-1}A_1 = X(J_p \oplus J_q)X'(X')^{-1}(A_{11} \oplus A_{12})X^{-1}$$

giving the contradiction $X^{-1}Y_1X = J_pA_{11} \oplus J_qA_{12}$. We conclude that $J_s + A_1$ is indecomposable.

By the first part of the proof (covering case (c) of Theorem 2.1), A_1 is symplectically congruent to a matrix $B_1 \oplus C_1$, where $B_1, C_1 \in \mathrm{Sym}_s(F)$ with B_1 diagonal and nonsingular and C_1 tridiagonal and of rank $s - 1$. As $J_s + A_1$ is indecomposable, all entries of C_1 lying just above the diagonal must be non-zero. In particular, the first $s - 1$ rows of C_1 are linearly independent. Since $J_s(B_1 \oplus C_1)$ is nilpotent, we have $(B_1C_1)^s = (C_1B_1)^s = 0$.

Let $g_1, g_2, \dots, g_s$ denote the standard basis of the column space F^s. Let S be the $s \times s$ permutation matrix representing the transformation $g_1 \leftrightarrow g_s$, $g_2 \leftrightarrow g_{s-1}$, *etc.* We take B to be the diagonal matrix $B = B_1 \oplus [0] \oplus (-SB_1S)$ of size m and rank $m - 1$, and define the tridiagonal matrix $C \in \mathrm{Sym}_m(F)$ by

$$C = \begin{bmatrix} C_1 & g_s & 0 \\ g_s' & 0 & g_1' \\ 0 & g_1 & -SC_1S \end{bmatrix}.$$

We claim that the matrix $Y := J_m(B \oplus C)$ is similar to the direct sum of two nilpotent Jordan blocks, each of size m.

First we prove that $Y^m = 0$. Clearly we have

$$Y^m = (-1)^s \begin{bmatrix} 0 & (CB)^sC \\ -(BC)^sB & 0 \end{bmatrix}.$$

As

$$BC = \begin{bmatrix} B_1C_1 & B_1g_s & 0 \\ 0 & 0 & 0 \\ 0 & -SB_1Sg_1 & SB_1C_1S \end{bmatrix}$$

https://doi.org/10.4153/CMB-2005-036-x Published online by Cambridge University Press

and $(B_1C_1)^s = 0$, we have

$$(BC)^s = \begin{bmatrix} 0 & (B_1C_1)^{s-1}B_1g_s & 0 \\ 0 & 0 & 0 \\ 0 & -S(B_1C_1)^{s-1}B_1Sg_1 & 0 \end{bmatrix}.$$

Hence $(BC)^sB = 0$. Now B_1 is invertible, so $C_1(B_1C_1)^{s-1} = 0$. This fact together with the last displayed equation and $Sg_1 = g_s$ yield $C(BC)^s = 0$, or equivalently $(CB)^sC = 0$. This proves that $Y^m = 0$.

Next we shall prove that C has rank $m-1$. Recall that the first $s-1$ rows of C_1 are linearly independent. As C_1 is singular, its last row is a linear combination of the first $s-1$ rows. Consequently, by subtracting a suitable linear combination of the first $s-1$ rows of C from its s-th row, we obtain a row having all entries 0 except the central entry which is equal to 1. A similar argument is applicable to the last s rows of C, *i.e.,* by subtracting a suitable linear combination of the last $s-1$ rows of C from the $s+2$-nd row, we obtain the same row as above. This shows that C is singular. On the other hand, by deleting the last row and the first column of C, we obtain an invertible lower triangular matrix. We conclude that C has rank $m-1$.

Since $Y^m = 0$ and Y has rank $n-2$, Y has to be similar to the direct sum of two nilpotent Jordan blocks of size m each, *i.e.,* our claim is proved.

As we are dealing with case (d) of Theorem 2.1, both Y and Z belong to $\mathfrak{sp}_n(F)$ and are nilpotent and have exactly two elementary divisors, t^m and t^m. As m is odd, we know from Theorem 2.1 that Y and Z are symplectically similar. Consequently, the matrices A and $B \oplus C$ are symplectically congruent.

This completes the proof of the theorem. ∎

Final Comment It would be worthwhile to determine if Theorem 3.1 remains valid for finite fields of characteristic different from 2. Our method yields a positive answer provided the polynomial function associated to $P_1 \cdots P_m$ (as defined in Proposition 2.4) is not identically zero, which is a problem of independent interest.

References

[1] L. V. Antonyan, *On the classification of homogeneous elements of* $\mathbf{Z}_2$*-graded semisimple Lie algebras.* (in Russian). Vestnik Moskov. Univ. Ser. I Mat. Mekh. (1982), 29–34.

[2] V. I. Arnold, *Mathematical Methods of Classical Mechanics.* Graduate Texts in Mathematics 60, Springer-Verlag, New York, 1978.

[3] C. S. Ballantine and E. L. Yip, *Congruence and conjunctivity of matrices.* Linear Algebra Appl. **32**(1980), 159–198.

[4] N. Bourgoyne and R. Cushman, *Conjugacy classes in linear groups.* J. Algebra **44**(1977), 339–362.

[5] D.Ž. Đoković, J. Patera, P. Winternitz, and H. Zassenhaus, *Normal forms of elements of classical real and complex Lie and Jordan algebras.* J. Math. Phys. **24**(1983), 1363–1374.

[6] D.Ž. Đoković, F. Szechtman and K. Zhao, *An algorithm that carries a square matrix into its transpose by an involutory congruence transformation.*, Electron. J. Linear Algebra **10**(2003), 320–340.

[7] P. Gabriel, *Appendix: degenerate bilinear forms.* J. Algebra, **31**(1974), 67–72.

[8] R. Gow, *The equivalence of an invertible matrix to its transpose.* Linear and Multilinear Algebra, **8**(1979/80), 329–336.

[9] C. Riehm, *The equivalence of bilinear forms.* J. Algebra **31**(1974), 45–66.

[10] T. A. Springer and R. Steinberg, *Conjugacy classes.* In: Seminar on Algebraic Groups and Related Finite Groups, Lecture Notes in Mathematics 131, Springer, Berlin, 1970, pp. 167–266.

[11] G. E. Wall, *On the conjugacy classes in the unitary, symplectic and orthogonal groups.* J. Austral. Math. Soc. **3**(1963), 1–62.
[12] J. Williamson, *On the algebraic problem concerning the normal forms of linear dynamical systems.* Amer. J. Math. **58** (1936), 141–163.

Department of Pure Mathematics
University of Waterloo
Waterloo, ON
N2L 3G1
e-mail: djokovic@uwaterloo.ca

Department of Mathematics and Statistics
University of Regina
Regina, SK
S4S 0A2
e-mail: szechtf@math.uregina.ca

Department of Mathematics
Wilfrid Laurier University
Waterloo, ON
N2L 3C5
e-mail: kzhao@wlu.ca

and

Mathematics Department
Henan University
Henan,China