THE SEMISIMPLE CONJUGACY CLASSES
IN THE SYMPLECTIC GROUPS

G.E. WALL

Abstract. We determine the conjugacy classes of semisimple elements in the symplectic groups Sp_{2m}(F), where F is an arbitrary field of characteristic not 2. This note was originally a letter dated 23 March, 2006, from G.E. Wall to Cheryl Praeger, and has been reproduced with his kind permission.

1. The general problem

The problem in question is to determine the conjugacy classes in the symplectic groups Sp_{2m}(F) over a field F. The general method proposed in the present section is used in Section 2 to give a detailed (and elementary) account of the conjugacy classes of semisimple elements in the case where char F \neq 2. (A semisimple element is one whose minimal polynomial is separable. These include all elements of finite order when char F = 0.)

Denote by \mathcal{F} the set of all non-degenerate alternating bilinear forms

f : F^{2m} \times F^{2m} \to F

and by \mathcal{G} the general linear group GL_{2m}(F) of all nonsingular linear mappings T : F^{2m} \to F^{2m}.

The natural permutation action of \mathcal{G} on \mathcal{F} is defined by

(fT)(u, v) = f(uT, vT) for all u, v \in F^{2m}.

The subgroup of \mathcal{G} formed by those elements that fix a given f is the symplectic group Sp(f). Since \mathcal{G} acts transitively on \mathcal{F}, these symplectic groups form a complete set of conjugate subgroups of \mathcal{G} (thereby justifying the generic notation Sp_{2m}(F)).

In order to put forms and linear mappings on the same footing, we introduce the set of pairs

\mathcal{P} = \{(f, T) \mid f \in \mathcal{F}, T \in \text{Sp}(f)\}

and define the action of \mathcal{G} on \mathcal{P} by

(f, T)S = (fS, S^{-1}TS).

The crucial observation is this:

Observation 1.1. For fixed f_0 \in \mathcal{F}, the elements T_1, T_2, \ldots \in \mathcal{G} are a set of representatives for the conjugacy classes of Sp(f_0) if and only if (f_0, T_1), (f_0, T_2), \ldots \in \mathcal{P} are a set of representatives for the orbits under the action of \mathcal{G} on \mathcal{P}.

Date: December 16, 2015.
In short, the original conjugacy class problem can be reformulated as one about orbits on \( \mathcal{P} \). It is from this new viewpoint that the problem will be treated from here. We now describe an alternative way of constructing a set of orbit representatives.

**Step 1:** We first choose representative elements \( R_1, R_2, \ldots \) from the conjugacy classes \( K_1, K_2, \ldots \) of \( G \), determining at the same time their centralisers \( b_1, b_2, \ldots \) in \( G \). This is a matter of standard linear algebra.

**Step 2:** We next determine, for each such representative \( R_k \), the set
\[
\mathcal{F}_k = \{ f \in \mathcal{F} \mid f R_k = f \} = \{ f \in \mathcal{F} \mid (f, R_k) \in \mathcal{P} \}.
\]
It may happen that \( \mathcal{F}_k \) is empty, which simply means that no symplectic group contains elements of \( G \) conjugate to \( R_k \). Assume now that \( \mathcal{F}_k \) is nonempty.

**Step 3:** The centraliser \( b_k \) acts naturally as a permutation group on \( \mathcal{F}_k \). The final step is to determine a set of representatives \( f_{k1}, f_{k2}, \ldots \) for the orbits of \( b_k \) in this action.

The pairs
\[
(f_{11}, R_1), (f_{12}, R_1), \ldots, (f_{21}, R_2), (f_{22}, R_2), \ldots
\]
so constructed form an alternative set of representatives for the orbits under the action of \( G \) on \( \mathcal{P} \), and thus give a new way of determining the conjugacy classes in the symplectic group.

2. **Semisimple elements**

In order to avoid exceptional cases, we assume throughout that
\[
\text{char } F \neq 2.
\]
No further restriction is imposed for the present.

The first task (Step 2 of §1) is as follows: given a nonsingular, even-dimensional linear transformation over the field \( F \), it is required to determine the nonsingular alternating bilinear forms that it leaves invariant.

In matrix terms, we are given \( X \in \text{GL}_{2m}(F) \) and are required to determine those \( A \in \text{GL}_{2m}(F) \) such that
\[
A = -A', \quad A = XAX',
\]
where \( ' \) denotes transpose. Notice that these conditions are equivalent to

(i) the form \( f_A \) given by \( f_A(u, v) = uAv' \) lies in \( \mathcal{F} \), and

(ii) \( X \) leaves \( f_A \) invariant, so that \( (f_A, X) \in \mathcal{P} \).

The second task (Step 3 of §1) arises when the set of \( A \) in (2) is nonempty. The centraliser of \( X \) in \( \text{GL}_{2m}(F) \) acts on this set by congruence:
\[
A \mapsto YA'Y' \text{ for } Y \text{ such that } Y^{-1}XY = X,
\]
and it is required to determine a set of representatives
\[
A_1, A_2, \ldots
\]
for the orbits. It is tacitly assumed from now on that $A$ and $X$ are nonsingular matrices satisfying (2).

**Lemma 2.1.** $X$ is similar to $X^{-1}$.

*Proof.* By (2), $A^{-1}X^{-1}A = X'$, so that $X^{-1}$ is similar to $X'$ and hence to $X$. □

**Notation 2.2.** If $f(t)$ is a monic polynomial with $f(0) \neq 0$ then $f^-(t)$ denotes the monic polynomial whose roots are the reciprocals of those of $f(t)$. Let $c_Y(t)$ denote the characteristic polynomial of a square matrix $Y$.

**Definition 2.3.** An elementary divisor of a square matrix $Y$ is a divisor of the minimal polynomial of $Y$ of the form $f(t) = g(t)^{\lambda}$, where $g(t)$ is monic and irreducible, which is related to the rational canonical form of $Y$.

Later we shall assume that $Y$ is semisimple. In this case, irreducible factors of the minimal polynomial of $Y$ occur with multiplicity 1.

**Corollary 2.4.** If $f(t)$ is an elementary divisor of $X$, then $f^-(t)$ is an elementary divisor of the same multiplicity.

Suppose that $X$ has block diagonal form, and $A$ has corresponding block matrix form:

(5) \[ X = \begin{pmatrix} X_1 & 0 & \cdots \\ 0 & X_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \quad A = \begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \]

Then, by (2),

(6) \[ X_iA_{ij}X_j' = A_{ij} = -A_{ji}' \text{ for all } i, j. \]

Hence $A_{ij}X_j' = X_i^{-1}A_{ij}$ and so, more generally,

(7) \[ A_{ij}f(X_j)' = f(X_i^{-1})A_{ij} \]

for any polynomial $f(t)$.

**Notation 2.5.** Let $c_Y(t)$ denote the characteristic polynomial of the square matrix $Y$.

**Lemma 2.6.** If

(8) \[ (c_{X_i^{-1}}(t), c_{X_j}(t)) = 1, \]

then $A_{ij} = 0$.

*Proof.* Taking $f(t) = c_{X_i^{-1}}(t)$ in (7), we get $A_{ij}c_{X_i^{-1}}(X_j)' = 0$. However, in view of (8), $c_{X_i^{-1}}(X_j)$ is nonsingular, whence $A_{ij} = 0$. □

Elementary divisors $f_1(t)$, $f_2(t)$ of $X$ are powers of irreducible monic polynomials $g_1(t)$, $g_2(t)$. We say that $f_1(t)$ and $f_2(t)$ are related if $g_2(t) = g_1(t)$ or $g_1^{-1}(t)$.

By the theory of elementary divisors, we may choose the blocks $X_i$ in (5) in such a way that elementary divisors $f_1(t)$, $f_2(t)$ of $X$ are elementary divisors of the same $X_i$ if, and
only if, they are related. With such a choice of the \( X_i \), Lemma 2.6 shows that \( A \) has corresponding block diagonal form

\[
\begin{pmatrix}
A_{11} & 0 & \cdots \\
0 & A_{22} & \\
\vdots & & \ddots
\end{pmatrix}.
\]

In this way the original problem for \( X \) is reduced to the same problem for the individual blocks \( X_i \). We may therefore assume:

**Assumption 2.7.** There exists a monic irreducible polynomial \( g(t) \neq t \) such that every elementary divisor of \( X \) is a power of \( g(t) \) or \( g^{-1}(t) \).

**Case 1:** \( g(t) \neq g^{-1}(t) \). We may assume in (5) that

\[
X = \begin{pmatrix}
X_1 & 0 \\
0 & X_2
\end{pmatrix},
\]

where \( c_{X_1}(t), c_{X_2}(t) \) are powers of \( g(t), g^{-1}(t) \) respectively. By Lemma 2.1, \( X_2 \) is similar to \( X_1^{-1} \). We may therefore assume further that

(9) \[
X = \begin{pmatrix}
X_1 & 0 \\
0 & (X_1^{-1})'
\end{pmatrix}, \quad A = \begin{pmatrix}
A_{11} & A_{12} \\
-A_{12}' & A_{22}
\end{pmatrix}.
\]

By Lemma 2.6, \( A_{11} = A_{22} = 0 \). Also, by (6), \( X_1A_{12}X_1^{-1} = A_{12} \), i.e. \( X_1 \) commutes with \( A_{12} \).

Now let

(10) \[
Y = \begin{pmatrix}
A_{12} & 0 \\
0 & I_m
\end{pmatrix}, \quad J = \begin{pmatrix}
0 & I_m \\
-I_m & 0
\end{pmatrix},
\]

where \( I_m \) is the \( m \times m \) unit matrix, and \( m = \deg c_{X_1}(t) \). Then

\[
Y^{-1}XY = X, \quad YJY' = A,
\]

showing that, with \( X \) as in (9), there is just one orbit under the action (3), represented by the matrix \( J \) in (10). Expressed differently, the conjugacy class of \( X \) in \( GL_{2m}(F) \) intersects each symplectic subgroup in a single conjugacy class of the latter.

**Case 2:** \( g(t) = g^{-1}(t) \).

In general, the elementary divisors of \( X \) may be arbitrary powers of \( g(t) \) with arbitrary multiplicities. We now impose the condition that \( X \) be semisimple:

**Assumption 2.8.** \( X \in GL_{2m}(F) \) has the single irreducible elementary divisor \( g(t) \neq t \) with multiplicity \( n \).

We may therefore assume that

(11) \[
X = \text{diag}(R, \ldots, R),
\]

where \( R \) is a matrix of size \( m \times m \).
where

\[(12)\]
\[c_R(t) = g(t).\]

Since \(c_R(t)\) is irreducible, the matrices
\[f(R) \quad (f(t) \in F[t])\]
form a field
\[K \cong F[t]/g(t)F[t]\]
and every matrix that commutes with \(R\) is in \(K\). It follows that the centralizer of \(X\) in \(\text{GL}_{2m}(F)\) consists of the nonsingular \(n \times n\) block matrices

\[(13)\]
\[B = (f_{ij}(R))_{i,j=1,...,n},\]
for polynomials \(f_{ij}(t) \in F(t)\). These matrices form a group that we may identify with \(\text{GL}_n(K)\).

If \(\deg g(t) = 1\), then \(g(t) = t \pm 1\) (since \(g(t) = g^{-1}(t)\)), the matrix \(R\) is a \(1 \times 1\) matrix \((\pm 1)\) and \(X = \pm I_{2m}\). The centralizer of \(X\) is \(\text{GL}_{2m}(F)\) and the nonsingular \(2m \times 2m\) skew-symmetric matrices form a single orbit under its action. We assume from now on that \(\deg g(t) \geq 2\).

Let

\[(14)\]
\[A = (A_{ij})_{i,j=1,...,n}\]
be the block form of \(A\) corresponding to \((11)\). The equation \(A = XAX'\) in \((2)\) is then equivalent to the set of equations

\[(15)\]
\[RA_{ij}R' = A_{ij}.\]

Now, since \(g(t) = g^{-1}(t)\), \(R\) is similar to \(R^{-1}\) and so

\[(16)\]
\[R' = T^{-1}R^{-1}T\]
for some \(T \in \text{GL}_{2m/n}(F)\). Thus, we may rewrite \((15)\) as
\[R(A_{ij}T^{-1}) = (A_{ij}T^{-1})R,\]
whence \(A\) has the form

\[(17)\]
\[A = (f_{ij}(R)T).\]

We write this equation as

\[(18)\]
\[A = BT \quad \text{where} \quad B = (f_{ij}(R)) \quad \text{and} \quad T = \text{diag}(T_1, \ldots, T_n).\]

Now, the mapping \(K \to K\) defined by
\[\phi(R) \mapsto \phi(R^{-1}) \quad (\phi(t) \in F[t])\]
is a field automorphism of \(K\) of order 2, since \(R \neq R^{-1}\). For a matrix
\[Y = (\phi_{ij}(R)) \in \text{GL}_n(K),\]
we define
\[Y^* = (\phi_{ij}(R^{-1}))^\text{tr},\]
where tr denotes transpose qua \( n \times n \) matrix over \( K \) and \( \text{not qua} \ 2m \times 2m \) matrix over \( F \), i.e. \( Y^* = (\phi_{ji}(R^{-1})) \). Accordingly, \( Y \) is called \textit{Hermitian} if \( Y^* = Y \) and \( Y_1, Y_2 \) are said to be \(-\text{congruent}\) if \( Y_2 = CY_1C^* \) for some \( C \in \text{GL}_n(K) \).

The following two results are proved by routine calculations:

\textbf{Lemma 2.9.}  
(i) If \( A = BT \) as in \((18)\) and \( Y \in \text{GL}_n(K) \), then \( YAY' = YBY^*T \).  
(ii) If \( A = BT \) as in \((18)\) and \( T \) is skew-symmetric, then \( A \) is skew-symmetric if, and only if, \( B \) is Hermitian.

\textit{Proof.} (i) We may write \( Y \) as a block matrix \( (\phi_{ij}(R)) \) for some \( \phi_{ij}(t) \in F[t] \). Then

\[
YAY' = (\phi_{ij}(R))(f_{ij}(R)T)(\phi_{ij}(R))' \\
= (\sum_{\lambda,\mu} \phi_{ij}(R)f_{\lambda\mu}(R)T\phi_{ji}(R')) \\
= (\sum_{\lambda,\mu} \phi_{ij}(R)f_{\lambda\mu}(R)\phi_{ji}(R^{-1})T), \\
= YBY^*T
\]

by \((16)\).

(ii) By \((18)\), \( A' = T'B' \), and since \( T \) is skew-symmetric we have

\[
-A' = (-T'f_{ji}(R')) = (Tf_{ji}(R'))' = (Tf_{ji}(R^{-1})T), \quad \text{(as \( T' = -T \))}
\]

So \( A = -A' \) if, and only if, \( f_{ij}(R)T = f_{ji}(R^{-1})T \) for all \( i, j \). Since \( T \) is invertible, this holds if, and only if,

\[
B = (f_{ij}(R)) = (f_{ji}(R^{-1})) = B^*.
\]

Thus, \textit{provided that \( T \) can be chosen skew-symmetric}, our conjugacy class problem in the present case reduces to a classification problem for Hermitian forms over the extension \( K \) of \( F \). Since \( T \) can obviously be replaced in \((17)\) by \( h(R)T \), where \( h(R) \) is any nonzero (and hence nonsingular) element of \( K \), the following result shows that such a choice of \( T \) is always possible.

\textbf{Lemma 2.10.} If \( g(t) \neq t \pm 1 \), then there exists a nonzero element \( h(R) \) of \( K \) such that \( h(R)T \) is skew-symmetric.

\textit{Proof.} \((16)\) can be written \( RTR' = T \). Transposing, we get \( RT'R' = T' \), whence

\[
(19) \quad R' = (T')^{-1}R^{-1}(T').
\]

Comparing with \((16)\), we deduce that \( T'T^{-1} \) commutes with \( R \), whence

\[
T' = f(R)T
\]

for some \( f(R) \in K \).

Now, if both \( RT \) are \( T \) were symmetric, we would have

\[
RT = (RT')' = T'R' = TR' = R^{-1}T
\]

by \((16)\). But this implies that \( R = R^{-1} \) and so \( R^2 = I \), contrary to the assumption that \( g(t) \neq t \pm 1 \).
It follows that at least one of $T$ and $RT$ — let us say $T$ itself — is not symmetric. But then
\[ T - T' = (1 - f(R))T \]
is nonzero and skew-symmetric, as required.

\[ \Box \]

3. Summary

We wish to determine a complete, irredundant set of conjugacy class representatives for the semisimple elements of the symplectic groups $\text{Sp}_{2m}(F)$, where $F$ is an arbitrary field of characteristic not 2. To do so, it suffices to consider the elements whose characteristic polynomial is divisible only by powers of $g(t)$ and $g^{-}(t)$, for some irreducible $g(t) \in F[t]$. We first choose a set $R_1, R_2, \ldots, R_k$ of representatives of conjugacy classes of such elements in $\text{GL}_{2m}(F)$, and discard any such $R_i$ that do not preserve a symplectic form.

In Case 1 of Section 2, $g(t) \neq g^{-}(t)$. Then the $\text{GL}_{2m}(F)$-conjugacy class of $R_i$ meets $\text{Sp}_{2m}(F)$ in a unique conjugacy class.

In Case 2, $g(t) = g^{-}(t)$. The only $R_i$ for which $g(t)$ has degree 1 are $\pm I_{2m}$. Let $R_i = X$ be as in (11), where $R$ has characteristic polynomial $g(t)$ of degree greater than 1, and let $K$ denote the field isomorphic to the set of all polynomials in $R$. Then each congruence class of Hermitian forms on $K^{2m/n}$ corresponds to an $\text{Sp}_{2m}(F)$-conjugacy class of matrices that are similar to $R_i$. In particular, if $F$ is finite then there is only one such class, and so once again the $\text{GL}_{2m}(F)$-conjugacy class of $R_i$ meets $\text{Sp}_{2m}(F)$ in a unique conjugacy class.

Emeritus Professor G.E. Wall, School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia.