University of Alberta

Flag Actions and Representations of the Symplectic Group

by

Jonathan Miersma

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Mathematics

Department of Mathematical and Statistical Sciences

©Jonathan Miersma
Spring 2011
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Abstract

A *flag* of a finite dimensional vector space V is a nested sequence of subspaces of V. The symplectic group of V acts on the set of flags of V. We classify the orbits of this action by defining the incidence matrix of a flag of V and showing that two flags are in the same orbit precisely when they have the same incidence matrix. We give a formula for the number of orbits of a certain type and discuss how to list the incidence matrices of all orbits. In the case in which V is a vector space over a finite field, we discuss the permutation representations of the symplectic group of V corresponding to these orbits. For the case in which $V = \mathbb{F}_q^4$, we compute the conjugacy classes of the symplectic group of V and the values of the characters of the previously discussed permutation representations.

Contents

List of Tables

Chapter 1

Introduction

Let V be a finite-dimensional vector space over a field $\mathbb{F}$. A *flag* of V is a nested sequence of subspaces of V. There is a natural action of $\mathrm{GL}(V)$, the general linear group of V, on the set of flags of V, and it is straightforward to classify the orbits of this action in terms of the dimensions of the subspaces of the flags. If the field $\mathbb{F}$ is finite, there is a permutation representation of the finite group $\mathrm{GL}(V)$ corresponding to each orbit of this action. In [11], Steinberg showed that a large number of irreducible characters of $\mathrm{GL}(V)$ may be computed by taking alternating sums of the characters of these permutation representations. In [9], James constructed the representations affording these characters, and proved a number of theorems about them.

Now if a group acts on a set, any of its subgroups acts on the set as well, so we can attempt to construct an analogous theory for any subgroup of $\mathrm{GL}(V)$. In the case in which V is an even-dimensional vector space endowed with a non-degenerate alternate bilinear form B, a particularly interesting subgroup of $\mathrm{GL}(V)$ is the symplectic group of V, the subgroup of $\mathrm{GL}(V)$ which preserves B. Our goal in this thesis is to analyze the action of the symplectic group of V on flags, and to study the corresponding representations and characters in the case in which $\mathbb{F}$ is a finite field.

In Chapter 2, we summarize some of the background information which will be needed in later chapters. We first review some of the basic definitions and theorems from the representation theory and character theory of finite groups, and then we define symplectic spaces and symplectic groups, and discuss some of their properties.

Chapter 3 contains most of the new results of this thesis. After defining flags and introducing some notation regarding them, we discuss the above-mentioned situation for $\mathrm{GL}(V)$ in more detail. We then introduce the incidence matrix of a flag of a symplectic space. After developing some properties of incidence matrices and showing how they reflect the structure of the flags, we classify the orbits of the action of the symplectic group of V ($\mathrm{Sp}(V)$) by showing that two flags are in the same orbit of the action of $\mathrm{Sp}(V)$ if and only if they have the same incidence matrix. In the next section we derive a formula for the number of orbits of a certain type, and in so doing we show how to compute recursively all the incidence matrices of all the flags of a symplectic space, allowing us to list all the orbits of the action for a given dimension of V. Finally, in the case in which $\mathbb{F}$ is a finite field, we derive formulae for the degrees of the representations of $\mathrm{Sp}(V)$ corresponding to the various orbits, and describe two situations in which distinct orbits yield isomorphic representations. Computational examples designed to clarify the results are sprinkled throughout the chapter.

In Chapter 4, we use the case in which $V = \mathbb{F}_q^4$ (the 4-dimensional vector space over the field of q elements with q odd) as a large computational example which tests the extent to which the results of Chapter 3 can be used to compute irreducible characters of $\mathrm{Sp}(V)$. After discussing some general results on canonical forms of matrices and their centralizers, we find explicit representatives for the conjugacy classes of $\mathrm{Sp}(V) = \mathrm{Sp}(4, q)$, and compute their sizes in terms of q. We then compute in terms of q the values on these class representatives of the characters of the representations discussed in Chapter 3. Using techniques which prove effective in the case of $\mathrm{GL}(V)$, we attempt to decompose these characters into irreducible characters. We show that while these techniques do produce a single non-trivial irreducible character of $\mathrm{Sp}(4, q)$ (the well-known Steinberg character), they are not nearly so effective as in the case of $\mathrm{GL}(V)$.

Chapter 2

Background Information

2.1 Representations and Characters

We begin with a brief summary of the fundamentals of the theory of group representations and characters. All the results in this section are proved in [8].

2.1.1 Definition. Let G be a group, and let $\mathbb{F}$ be a field. An $\mathbb{F}$-*representation* of G is a homomorphism $\Phi : G \to \mathrm{GL}(V)$, where V is a finite-dimensional non-zero $\mathbb{F}$-vector space and $\mathrm{GL}(V)$ is the group of invertible linear transformations of V. The *degree* of Φ is the dimension $\dim V$ of V.

By choosing a basis for the n-dimensional vector space V, we obtain an isomorphism of $\mathrm{GL}(V)$ with $\mathrm{GL}(n, \mathbb{F})$, the group of invertible $n \times n$ matrices over $\mathbb{F}$, and hence via composition the $\mathbb{F}$-representation Φ yields a group homomorphism $\widehat{\Phi} : G \to \mathrm{GL}(n, \mathbb{F})$, which we call a matrix representation of G. If $A \in \mathrm{M}(n, \mathbb{F})$, the set of all $n \times n$ matrices over $\mathbb{F}$, then the trace of A, denoted $\mathrm{Tr}\, A$, is the sum of the diagonal entries of A. A basic result from linear algebra states that if $B \in \mathrm{GL}(n, \mathbb{F})$, then $\mathrm{Tr}\, B^{-1}AB = \mathrm{Tr}\, A$. If $T \in \mathrm{GL}(V)$, and if β_1 and β_2 are any two bases for V, then the matrices for T with respect to β_1 and β_2 are conjugate under the appropriate change of basis matrix, and hence we can define the trace of T to be the trace of any matrix representing T. This justifies the following definition.

2.1.2 Definition. The *character* of Φ is the function $\chi : G \to \mathbb{F}$ given by $\chi(g) = \mathrm{Tr}\, \Phi(g)$ for $g \in G$. The *degree* of χ is the degree of Φ. A *character* of G is a function $G \to \mathbb{F}$ which is the character of some $\mathbb{F}$-representation of G.

If $h, g \in G$, then the conjugacy-invariance property of the trace function shows that $\chi(h^{-1}gh) = \chi(g)$, so that χ is really a *class function* on G; that is, it is a function $G \to \mathbb{F}$ which depends only on the conjugacy classes of G.

If $\Phi : G \to \mathrm{GL}(V)$ is an $\mathbb{F}$-representation of G on the vector space V, and if $W \leq V$ is a non-zero subspace of V such that $\Phi(g)W = W$ for each $g \in G$, then $\Phi_W : G \to \mathrm{GL}(W)$ given by $\Phi_W(g)(w) = \Phi(g)(w)$ is a representation of G on the vector space W.

2.1.3 Definition. A *subrepresentation* of Φ is an $\mathbb{F}$-representation of the form Φ_W for some non-zero subspace W of V. The subrepresentation is said to be *proper* if $W \neq V$.

An $\mathbb{F}$-representation automatically has itself as a subrepresentation, but it need not have any proper subrepresentations.

2.1.4 Definition. An $\mathbb{F}$-representation is *irreducible* if it has no proper subrepresentations. A character is *irreducible* if the representation to which it corresponds is irreducible.

We are interested in the way in which representations can be decomposed into irreducible subrepresentations. To describe this decomposition process, we need two more concepts.

2.1.5 Definition. Let Ψ and Φ be $\mathbb{F}$-representations of the group G on the vector spaces U and V. Then the homomorphism $\Psi \oplus \Phi : G \to \mathrm{GL}(U \oplus V)$ given by $(\Psi \oplus \Phi)(g)(u+v) = \Psi(g)(u) + \Phi(g)(v)$ is a representation of G on $U \oplus V$, called the *direct sum* of Ψ and Φ.

If we pick bases β_1 and β_2 of U and V and take $\beta_1 \cup \beta_2$ as a basis of $U \oplus V$, then for any $g \in G$,

$$\widehat{\Psi \oplus \Phi}(g) = \left(\begin{array}{c|c} \widehat{\Psi}(g) & 0 \\ \hline 0 & \widehat{\Phi}(g) \end{array}\right),$$

and hence the character of $\Psi \oplus \Phi$ is the sum of the characters of Ψ and Φ.

2.1.6 Definition. Ψ and Φ are said to be *isomorphic* if there exists a vector space isomorphism $T : U \to V$ such that for all $g \in G$ and all $u \in U$ $\Phi(g)(T(u)) = T(\Psi(g)(u))$.

Equivalently, Ψ and Φ are isomorphic if there exist bases of U and V such that $\widehat{\Psi} = \widehat{\Phi}$ (that is, the corresponding matrix representations are equal), and hence isomorphic representations have the same characters. Although much of what follows is valid in a broader context, we now specialize to the case $\mathbb{F} = \mathbb{C}$, the complex numbers, which is our primary concern in this thesis. Henceforward the term representation will mean a $\mathbb{C}$-representation, and G will denote a finite group.

2.1.7 Theorem (Maschke's Theorem)**.** Every representation of G is isomorphic to a direct sum of irreducible representations.

It can be shown that this decomposition is unique up to isomorphism and reordering of the irreducible representations, and also that, up to isomorphism, G has only a finite number of irreducible representations. The use of characters helps to clarify this situation.

2.1.8 Proposition. Let $\{\Phi_1, \Phi_2, \ldots, \Phi_k\}$ be a set of representatives of the isomorphism classes of irreducible representations of G. Let $\beta = \{\chi_1, \chi_2, \ldots, \chi_k\}$ be the corresponding characters. Then k is equal to the number of conjugacy classes of G, and β is a basis for the vector space of $\mathbb{C}$-valued class functions on G.

Now if Φ is a representation of G with character χ, we know that Φ is isomorphic to a unique direct sum of the form $\bigoplus_{i=1}^{k} k_i\Phi_i$, where $k_i\Phi_i = \bigoplus_{j=1}^{k_i} \Phi_i$ if $k_i \in \mathbb{N}$, and Φ_i is omitted if $k_i = 0$. Thus $\chi = \sum_{i=1}^{k} k_i\chi_i$. If we know χ and $\beta = \{\chi_1, \chi_2, \ldots, \chi_k\}$, we can use linear algebra to express χ in terms of β, and the coefficients in this expression will necessarily be the k_i. Hence χ determines Φ up to isomorphism, and further, if we have computed the characters of all the irreducible representations of G, we have an explicit computational technique for decomposing any representation into irreducible representations. Thus one of the main problems of group representation theory is determining all the irreducible characters of G. Once these are found, one can choose an ordering $\chi_1, \chi_2, \ldots, \chi_k$ of the irreducible characters and an ordering $c_1, c_2, \ldots, c_k$ of the conjugacy classes of G and form the *character table* of G, the $k \times k$ matrix with (i, j) entry $\chi_i(c_j)$.

2.1.9 Example. The character table of the group $\mathrm{GL}(3, 2)$[1] is given in Table 2.1. The top row is a list of representatives of the conjugacy classes of $\mathrm{GL}(3, 2)$, while the second row gives the sizes of these conjugacy classes. Technically, the character table is the 6×6 matrix below the line, but it is often useful to include other information, such as class sizes and orders of elements, above the line.

g	$\begin{pmatrix} 1&0&0\\0&1&0\\0&0&1 \end{pmatrix}$	$\begin{pmatrix} 1&0&0\\0&1&1\\0&0&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&0\\0&1&1\\0&0&1 \end{pmatrix}$	$\begin{pmatrix} 1&0&0\\0&0&1\\0&1&1 \end{pmatrix}$	$\begin{pmatrix} 0&0&1\\1&0&1\\0&1&0 \end{pmatrix}$	$\begin{pmatrix} 0&0&1\\1&0&0\\0&1&1 \end{pmatrix}$
$\lvert Cl(g)\rvert$	1	21	42	56	24	24
χ_1	1	1	1	1	1	1
χ_2	3	-1	1	0	$\frac{-1+i\sqrt{7}}{2}$	$\frac{-1-i\sqrt{7}}{2}$
χ_3	3	-1	1	0	$\frac{-1-i\sqrt{7}}{2}$	$\frac{-1+i\sqrt{7}}{2}$
χ_4	6	2	0	0	-1	-1
χ_5	7	-1	-1	1	0	0
χ_6	8	0	0	-1	1	1

Table 2.1: Character Table of $\mathrm{GL}(3, 2)$.

Expressing a character as a $\mathbb{Z}_{\geq 0}$-linear combination of irreducible characters can be made easier through the introduction of an inner product on the vector space of $\mathbb{C}$-valued class functions on G.

[1]If $\mathbb{F}$ is the finite field with q elements, we denote $\mathrm{GL}(n, \mathbb{F})$ by $\mathrm{GL}(n, q)$.

2.1.10 Definition. Let ρ and χ be class functions on G. The *inner product* of ρ and χ is

$$\langle \rho, \chi \rangle = \frac{1}{|G|} \sum_{g \in G} \rho(g) \overline{\chi(g)}.$$

If $c_1, c_2, \ldots, c_k$ are the conjugacy classes of G, then since ρ and χ are class functions, we can compute their inner product as

$$\langle \rho, \chi \rangle = \frac{1}{|G|} \sum_{i=1}^{k} |c_i| \rho(c_i) \overline{\chi(c_i)}.$$

The irreducible characters of G behave particularly well with respect to this inner product.

2.1.11 Proposition. If $\chi_1, \chi_2, \ldots, \chi_k$ are the irreducible characters of G then

$$\langle \chi_i, \chi_j \rangle = \delta_{ij}.$$

Hence if χ is a character of G and $\chi = \sum_{i=1}^{k} k_i \chi_i$, we have $k_i = \langle \chi, \chi_i \rangle$, which allows us to decompose χ, while $\langle \chi, \chi \rangle = \sum_{i=1}^{k} k_i^2$, which allows us to gauge how far χ is from being irreducible, even if the χ_i are unknown. For example, if $\langle \chi, \chi \rangle = 3$, then since $3 = 1^2 + 1^2 + 1^2$ is the only way of expressing 3 as a sum of squares, χ must be the sum of 3 distinct irreducible characters. In particular, $\langle \chi, \chi \rangle = 1$ if and only if χ is irreducible. The inner product is especially useful in constructing the character table of G, since if we have a few irreducible characters of G, and a reducible character χ, we can determine the coefficient of each of the irreducible characters in χ and subtract them from χ accordingly, possibly obtaining a new irreducible character.

We now turn our attention to a specific class of representations, known as permutation representations, which are of particular interest to us in this thesis. Suppose G acts on a finite set X. Then there is a representation Φ_X of G on $\mathbb{C}X$, the free $\mathbb{C}$-vector space on the set X. If $g \in G$ and $X = \{x_1, x_2, \ldots, x_n\}$, then an arbitrary element of $\mathbb{C}X$ has the form $\sum_{i=1}^{n} \lambda_i x_i$ for $\lambda_i \in \mathbb{C}$, and we define $\Phi_X(g) \left(\sum_{i=1}^{n} \lambda_i x_i \right) = \sum_{i=1}^{n} \lambda_i (g x_i)$.

2.1.12 Definition. Φ_X is the *permutation representation* of G on the set X.

By definition, the degree of Φ_X is the dimension of $\mathbb{C}X$, which is simply $|X|$. It is possible to compute the character χ_X of Φ_X based solely on a knowledge of the action of G on X.

2.1.13 Lemma. For $g \in G$, $\chi_X(g)$ is the number of $x \in X$ such that $gx = x$.

Proof. With respect to the basis X, the matrix for $\Phi_X(g)$ is a permutation matrix, having a 1 in the (i, j) entry if $g x_j = x_i$ and a 0 otherwise. $\chi_X(g) = \mathrm{Tr}(\Phi_X(g))$ is the sum of all the (i, i) entries of the matrix, which equals the number of (i, i) entries equal to 1, which equals the number of $x_i \in X$ such that $g x_i = x_i$. □

If $X = \{x\}$ is a singleton set, then G acts trivially on X. If $g \in G$, then g fixes exactly one element of X, and hence $\chi_X(g) = 1$. Now

$$\langle \chi_X, \chi_X \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_X(g) \overline{\chi_X(g)} = \frac{1}{|G|} \sum_{g \in G} 1 = 1,$$

and so we have the following proposition.

2.1.14 Proposition. The class function χ defined by $\chi(g) = 1$ for all $g \in G$ is an irreducible character of G.

We call χ the *trivial character* of G.

We will later use Lemma 2.1.13 to compute a large number of characters of the 4-dimensional symplectic group over a finite field. In the next section, we define symplectic spaces and groups and discuss some useful results concerning them.

2.2 Symplectic Spaces and Symplectic Groups

Our development in this section is based primarily on [6], which contains proofs of any results which we state without proof. The concept of a symplectic space is somewhat analogous to that of ordinary Euclidean space, in that a symplectic space consists of a vector space together with a bilinear form defined on it. In particular, there is a notion of orthogonality for vectors in a symplectic space. However, the fact that every vector in a symplectic space is orthogonal to itself shows that there are also major differences between the two concepts, and that we must be careful in transferring our intuitive geometric notions of Euclidean space to symplectic spaces. We begin with some general definitions. Throughout this section, $\mathbb{F}$ denotes an arbitrary field.

2.2.1 Definition. Let V be an $\mathbb{F}$-vector space. A *bilinear form* on V is a function $B : V \times V \to \mathbb{F}$ such that for all $u, v, w \in V$ and all $\lambda, \mu \in \mathbb{F}$ we have

1. $B(\lambda u + \mu v, w) = \lambda B(u, w) + \mu B(v, w)$
2. $B(u, \lambda v + \mu w) = \lambda B(u, v) + \mu B(u, w)$

Now suppose $\beta = \{v_1, v_2, \ldots, v_n\}$ is a basis of V, and that $u = \sum_{i=1}^{n} \lambda_i v_i$ and $v = \sum_{j=1}^{n} \mu_j v_j$ are two vectors in V. Then

$$B(u, v) = B\left(\sum_{i=1}^{n} \lambda_i v_i, \sum_{j=1}^{n} \mu_j v_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \mu_j B(v_i, v_j)$$

$$= (\lambda_1, \lambda_2, \ldots, \lambda_n) \begin{pmatrix} B(v_1, v_1) & B(v_1, v_2) & \ldots & B(v_1, v_n) \\ B(v_2, v_1) & B(v_2, v_2) & \ldots & B(v_2, v_n) \\ \vdots & \vdots & \ddots & \vdots \\ B(v_n, v_1) & B(v_n, v_2) & \ldots & B(v_n, v_n) \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}.$$

In other words, if we let $\widehat{B}$ be the matrix with (i,j)-entry $B(v_i, v_j)$, and if $[u]_\beta$ and $[v]_\beta$ are the column vectors of coordinates of u and v with respect to β, we have $B(u,v) = [u]_\beta^T \widehat{B}[v]_\beta$. Conversely, if we have a basis of V, any $A \in \mathrm{M}(n, \mathbb{F})$ can be taken as defining an $\mathbb{F}$-valued function on pairs of basis elements, and this function can be extended via bilinearity to a bilinear form on V. We will usually define bilinear forms in this way. For both computational and theoretical purposes it is important to know how the matrix $\widehat{B}$ of a bilinear form B behaves with respect to a change of basis. So suppose β and β' are two bases of V, and that C is the change of basis matrix from β' to β, so that for $u \in V$, $C[u]_{\beta'} = [u]_\beta$. Then if $\widehat{B}$ is the matrix for B with respect to β, we have that for any $u, v \in V$

$$B(u,v) = [u]_\beta^T \widehat{B}[v]_\beta = (C[u]_{\beta'})^T \, \widehat{B} \, (C[v]_{\beta'}) = [u]_{\beta'}^T \left(C^T \widehat{B} C \right) [v]_{\beta'},$$

so that the matrix for B with respect to β' is $C^T \widehat{B} C$.

We are interested in a specific type of bilinear form.

2.2.2 Definition. A bilinear form B on V is *alternate* if $B(v,v) = 0$ for all $v \in V$.

If B is alternate, it follows that $B(u,v) = -B(v,u)$. In particular, if $\beta = \{v_1, v_2, \ldots, v_n\}$ is a basis of V, $B(v_i, v_j) = -B(v_j, v_i)$, and so $\widehat{B}^T = -\widehat{B}$. Also, $B(v_i, v_i) = 0$, so all the diagonal entries of $\widehat{B}$ are 0. Conversely, if $A \in \mathrm{M}(n, \mathbb{F})$ has all diagonal entries 0 and satisfies $A^T = -A$, then the bilinear form defined by A is alternate. (If the characteristic of $\mathbb{F}$ is not 2, then $A^T = -A$ implies that all diagonal entries of A are 0, but in characteristic 2, both conditions are necessary.) If B is alternate, then $B(u,v) = 0$ if and only if $B(v,u) = 0$, which allows us to define orthogonality as a symmetric and reflexive relation on V.

2.2.3 Definition. Let B be an alternate bilinear form on V, and $u, v \in V$. We say that u and v are *orthogonal* if $B(u,v) = 0$. Two subsets X and Y of V are *orthogonal* if B restricted to $X \times Y$ is identically 0. If W is a subspace of V, then the *orthogonal complement* of W is

$$W^\perp = \{v \in V | B(v,w) = 0 \text{ for all } w \in W\}.$$

$W^\perp$ is a subspace of V. If $U \subseteq W$ are subspaces of V it follows from the definition that $W^\perp \subseteq U^\perp$, as in the Euclidean case. However, since orthogonality is a reflexive relation on V, it is not necessarily the case that $W \cap W^\perp = 0$.

2.2.4 Definition. Let B be an alternate bilinear form on the vector space V. Let W be a subspace of V. Then the *radical* of W is $\operatorname{Rad} W = W \cap W^\perp$. W is said to be *non-degenerate* if $\operatorname{Rad} W = 0$. The form B is *non-degenerate* if $\operatorname{Rad} V = 0$.

We can now define the concept of a symplectic space.

2.2.5 Definition. A *symplectic space* is a vector space V together with a non-degenerate alternate bilinear form B on V. Such a form B is called a *symplectic form.*

We now state some fundamental results, all of which are proved in [6].

2.2.6 Proposition. Let (V, B) be a symplectic space, and W a subspace of V. Then

$$\dim W^{\perp} = \dim V - \dim W \qquad \text{and} \qquad (W^{\perp})^{\perp} = W.$$

2.2.7 Proposition. Let B be an alternate bilinear form on the vector space V. If W is a non-degenerate subspace of V, then $V = W \oplus W^{\perp}$.

2.2.8 Theorem. Let B be an alternate bilinear form on the vector space V, let $n = \dim V$, and let $r = \dim \operatorname{Rad} V$. Then there exists a basis of V with respect to which the matrix $\widehat{B}$ of B has the block-diagonal form

$$\widehat{B} = \begin{pmatrix} M & & & 0 \\ & \ddots & & \\ & & M & \\ 0 & & & 0_r \end{pmatrix}, \qquad \text{where} \qquad M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

0_r is the $r \times r$ zero matrix, and M appears $\frac{n-r}{2}$ times on the diagonal.

Since $\frac{n-r}{2} \in \mathbb{Z}_{\geq 0}$ in Theorem 2.2.8, and since $r = 0$ if B is non-degenerate, we obtain the following useful result.

2.2.9 Corollary. $\dim V \equiv \dim \operatorname{Rad} V \mod 2$. In particular, if V is symplectic, $\dim V$ is even.

One more useful fact, which relates two important operations on subspaces, is used implicitly without proof in [6].

2.2.10 Lemma. Let B be an alternate bilinear form on the vector space V, and let U and W be subspaces of V. Then

$$(U + W)^{\perp} = U^{\perp} \cap W^{\perp} \qquad \text{and} \qquad U^{\perp} + W^{\perp} = (U \cap W)^{\perp}.$$

Proof. If $v \in (U + W)^{\perp}$, then $B(v, x) = 0$ for all $x \in U + W$. In particular, $B(v, u) = 0$ for all $u \in U$, so that $v \in U^{\perp}$ and $B(v, w) = 0$ for all $w \in W$, so that $v \in W^{\perp}$, and so $v \in U^{\perp} \cap W^{\perp}$. Conversely, if $v \in U^{\perp} \cap W^{\perp}$, then for all $u \in U$ and all $w \in W$, $B(v, u) = B(v, w) = 0$, so for any $x = u+w \in U+W$, we have $B(v, x) = B(v, u+w) = B(v, u)+B(v, w) = 0$, and hence $v \in (U+W)^{\perp}$.

To prove the second equality, replace U and W with $U^{\perp}$ and $W^{\perp}$ in the first equality. This shows that

$$(U^{\perp} + W^{\perp})^{\perp} = U \cap W,$$

and taking the orthogonal complement of both sides yields

$$U^\perp + W^\perp = (U \cap W)^\perp.$$

$\square$

For the remainder of this section, let V be a symplectic space of dimension $n = 2m$ over the field $\mathbb{F}$ with symplectic form B.

2.2.11 Definition. Let $g \in \mathrm{GL}(V)$. Then g is a *symplectic transformation* of V if for all $u, v \in V$ we have

$$B(gu, gv) = B(u, v).$$

The *symplectic group* of V is the subgroup $\mathrm{Sp}(V)$ of $\mathrm{GL}(V)$ consisting of all symplectic transformations of V.

Choosing a basis β for V yields an isomorphism of $\mathrm{Sp}(V)$ with a subgroup of $\mathrm{GL}(n, \mathbb{F})$. Let $\widehat{B}$ be the matrix for B with respect to β, and suppose $A \in \mathrm{GL}(n, \mathbb{F})$ is the matrix for $g \in \mathrm{GL}(V)$ with respect to β. Then to determine whether $g \in \mathrm{Sp}(V)$, we note that for $u, v \in V$,

$$B(u, v) = [u]_\beta^T \widehat{B} [v]_\beta$$

and

$$B(gu, gv) = (A[u]_\beta)^T \widehat{B} (A[v]_\beta) = [u]_\beta^T (A^T \widehat{B} A)[v]_\beta.$$

Hence $B(gu, gv) = B(u, v)$ for all $u, v \in V$ if and only if

$$[u]_\beta^T \widehat{B} [v]_\beta = [u]_\beta^T (A^T \widehat{B} A)[v]_\beta$$

for all $u, v \in V$, which holds if and only if $A^T \widehat{B} A = \widehat{B}$. Hence $\mathrm{Sp}(V)$ is isomorphic to the matrix group

$$\{A \in \mathrm{GL}(n, \mathbb{F}) | A^T \widehat{B} A = \widehat{B}\}.$$

Two different bases of V may yield different matrices $\widehat{B}$, but the corresponding matrix groups will be conjugate under the appropriate change of basis matrix. By Theorem 2.2.8 we can choose a basis of V for which $\widehat{B}$ is block diagonal, with m copies of $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ on the diagonal, which yields an isomorphism of $\mathrm{Sp}(V)$ with an unambiguously defined matrix group. (Note, however, that any invertible matrix J satisfying $J^T = -J$ defines a symplectic from on V with respect to a given basis, and so could be used in place of the specific $\widehat{B}$ mentioned above.) As a consequence, $\mathrm{Sp}(V)$ is determined up to isomorphism by n and $\mathbb{F}$, and may be denoted by $\mathrm{Sp}(n, \mathbb{F})$. If $\dim V = 2$, the preceding comments allow us to determine $\mathrm{Sp}(V)$ easily.

2.2.12 Proposition. If $\dim V = 2$, then $\mathrm{Sp}(V) = \mathrm{SL}(V)$, the subgroup of $\mathrm{GL}(V)$ consisting of those elements with determinant 1.

Proof. Sp(V) is isomorphic to the matrix group

$$\left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}(2,\mathbb{F}) | \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)^T \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \right\}.$$

But

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)^T \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right) \left(\begin{smallmatrix} c & d \\ -a & -b \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & ad-bc \\ -(ad-bc) & 0 \end{smallmatrix}\right),$$

so this matrix group consists precisely of the matrices in GL$(2,\mathbb{F})$ having determinant 1, and thus is isomorphic to SL(V). □

As shown in [6], we also have the following more general relation between the symplectic and special linear groups.

2.2.13 Proposition. For any vector space V of even dimension, $\mathrm{Sp}(V) \subseteq \mathrm{SL}(V)$.

It will be useful to know how elements of Sp(V) relate to the orthogonal complement operator on subspaces.

2.2.14 Lemma. Let W be a subspace of V, and $g \in \mathrm{Sp}(V)$. Then $(gW)^\perp = g(W^\perp)$.

Proof. If $u \in V$, then $u \in (gW)^\perp$ if and only if $B(u, gw) = 0$ for all $w \in W$. Since $g^{-1} \in \mathrm{Sp}(V)$, this is true if and only if $B(g^{-1}u, w) = 0$ for all $w \in W$, which is true if and only if $g^{-1}u \in W^\perp$, which is equivalent to $u \in g(W^\perp)$. □

We will have occasion to use the two following propositions, both of which are proved in [6].

2.2.15 Proposition. Let $\Gamma = \{(u, v) \in V \times V | B(u, v) = 1\}$. Then the action of Sp($V$) on Γ given by $g(u, v) = (gu, gv)$ for $g \in \mathrm{Sp}(V)$ is transitive.

If $\mathbb{F} = \mathbb{F}_q$ is the finite field having $q = p^r$ elements for some prime p, we will replace $\mathbb{F}_q$ with q in our notation for various types of matrix groups and algebras, and so in particular we denote $\mathrm{Sp}(n, \mathbb{F}_q)$ by $\mathrm{Sp}(n, q)$. Recall that $n = 2m$ is even.

2.2.16 Proposition. $\mathrm{Sp}(n, q)$ is a finite group of order

$$q^{m^2} \prod_{i=1}^{m} (q^{2i} - 1).$$

We conclude this section with another proposition from [6] which, while not particularly important for this thesis, demonstrates the importance of the symplectic groups within group theory.

2.2.17 Proposition. The centre of Sp(V) is $Z(\mathrm{Sp}(V)) = \{\pm I\}$. If we define $\mathrm{PSp}(V) = \mathrm{Sp}(V)/Z(\mathrm{Sp}(V))$, then PSp($V$) is a simple group unless $\mathbb{F} = \mathbb{F}_2$ and $\dim V \in \{2, 4\}$ or $\mathbb{F} = \mathbb{F}_3$ and $\dim V = 2$.

The groups $\mathrm{PSp}(2m, q)$ form one of the infinite families of finite simple groups in the classification theorem of finite simple groups.

Chapter 3

Flag Representations and Incidence Theory of $\mathrm{Sp}(n, \mathbb{F})$

3.1 Definitions and Examples

We begin by defining an object which is of central interest to us in this thesis. Throughout this section, let $\mathbb{F}$ be a field, V a finite-dimensional vector space over $\mathbb{F}$, and $n = \dim V$.

3.1.1 Definition. A *flag* of V is a sequence $\mathcal{W} = \{W_0, W_1, \ldots, W_k\}$ of subspaces of V such that $W_0 = V$, $W_k = 0$, and W_{i+1} is properly contained in W_i for $i = 0, \ldots, k-1$. To emphasize the nested character of the sequence, we write

$$\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_k\}.$$

We may also have occasion to write

$$\mathcal{W} = \{W_k \subset \cdots \subset W_1 \subset W_0\}.$$

In either case, the inclusions specify the order of the sequence. If $\mathcal{U}$ is a flag which can be obtained from $\mathcal{W}$ by deleting some of the W_i, we call $\mathcal{U}$ a $subflag$ of $\mathcal{W}$.

The concept of the orthogonal complement of a subspace extends readily to flags of symplectic spaces.

3.1.2 Definition. Suppose V is a symplectic space and let

$$\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_{k-1} \supset W_k\}$$

be a flag of V. Then the *orthogonal complement* of $\mathcal{W}$ is the flag

$$\mathcal{W}^\perp = \{W_k^\perp \supset W_{k-1}^\perp \supset \cdots \supset W_1^\perp \supset W_0^\perp\}.$$

The following definition introduces some useful notation.

3.1.3 Definition. A *composition* of a number $m \in \mathbb{N}$ is a finite sequence $\mu = \{n_1, n_2, \ldots, n_k\}$ with $n_i \in \mathbb{N}$ such that $\sum_{i=1}^{k} n_i = m$. If $n_i \leq n_{i+1}$ for $i = 1, \ldots, k-1$, then μ is called a *partition* of m.

If $\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_k\}$ is a flag of V, then by canceling adjacent terms, we see that

$$\sum_{i=1}^{k} \dim(W_{i-1}/W_i) = \sum_{i=1}^{k} (\dim W_{i-1} - \dim W_i) = \dim W_0 - \dim W_k = n.$$

This justifies the following definition.

3.1.4 Definition. The *type* of the flag $\mathcal{W}$ is the composition

$$\lambda = \{\dim(W_0/W_1), \dim(W_1/W_2), \ldots, \dim(W_{k-1}/W_k)\}$$

of n. If μ is any composition of n, a *μ-flag* of V is a flag of V of type μ. The set of all μ-flags of V will be denoted by F_μ. If $\mu = \{1, 1, \ldots, 1\}$ is the composition of n consisting entirely of ones, then a μ-flag is called a *full flag*. A flag of any other type is called a *partial flag*.

If we delete a subspace W_i from a λ-flag $\mathcal{W}$, the type of the resulting subflag is obtained by adding $\dim(W_{i-1}/W_i)$ and $\dim(W_{i+1}/W_i)$ in the composition λ. Thus a λ-flag has a subflag of type μ if and only if μ can be obtained from λ by adding certain consecutive elements of λ.

3.1.5 Definition. Let $\lambda = \{\lambda_1, \ldots, \lambda_k\}$ and $\mu = \{\mu_1, \ldots, \mu_l\}$ be compositions of n. Then we write $\mu \preceq \lambda$ if there exist $n_i \in \mathbb{Z}$ such that $0 = n_0 < n_1 < \cdots < n_l$ and

$$\mu_i = \sum_{j=n_{i-1}+1}^{n_i} \lambda_j$$

for all i.

In light of the preceding comments, we have the following proposition.

3.1.6 Proposition. Let $\mathcal{W}$ be a λ-flag of V and let μ be any other composition of V. If $\mu \preceq \lambda$, then $\mathcal{W}$ has a unique subflag of type μ, which we denote $\mathcal{W}_\mu$. Otherwise, $\mathcal{W}$ has no subflag of type μ.

We now give some examples of the geometric significance of flags of various types.

3.1.7 Example. If $0 < m < n$, then by definition an $\{n-m, m\}$-flag of V consists of a sequence $\mathcal{W} = \{V \supset W \supset 0\}$, where W is a subspace of V of dimension m, but since V and 0 add no information, we usually identify the set of $\{n-m, m\}$-flags of V with the m-dimensional subspaces of V.

3.1.8 Example. If $n = 3$, there are 4 different types of flags. The only $\{3\}$-flag is $\{V \supset 0\}$, which we regard as trivial. $\{2,1\}$-flags correspond to lines, while $\{1,2\}$-flags correspond to planes, and a full flag consists of a line within a plane.

Now if $T \in \mathrm{GL}(V)$, T permutes the m-dimensional subspaces of V, with T mapping $W \subset V$ to TW, the image of W under T. Further, if $W_1 \supset W_2$ are two subspaces of V, then $TW_1 \supset TW_2$. Hence we see that if λ is any composition of n, $\mathrm{GL}(V)$ acts on F_λ, and by restriction, so does any subgroup H of $\mathrm{GL}(V)$, including $\mathrm{Sp}(V)$. If the field $\mathbb{F} = \mathbb{F}_q$ is finite, F_λ is a finite set, and so we get a permutation representation of $\mathrm{GL}(V)$, or any of its subgroups, on the set F_λ. We refer to these representations as *flag representations*. In the case of $\mathrm{GL}(n,q)$, flag representations can be used to compute explicitly a significant number of irreducible characters. We illustrate this procedure with an example before commenting on the general situation.

3.1.9 Example. Let $G = \mathrm{GL}(3,2)$, and for a given composition λ of 3, let ρ_λ be the character of the permutation representation of G on F_λ. Thus by Lemma 2.1.13, $\rho_\lambda(g)$ is the number of λ-flags fixed by g. Table 3.1 gives the values of ρ_λ on the classes of G for $\lambda = \{3\}, \{1,2\}$, and $\{1,1,1\}$.

g	$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}$	$\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&0&1\\0&1&1\end{pmatrix}$	$\begin{pmatrix}0&0&1\\1&0&1\\0&1&0\end{pmatrix}$	$\begin{pmatrix}0&0&1\\1&0&0\\0&1&1\end{pmatrix}$
$\lvert Cl(g)\rvert$	1	21	42	56	24	24
$\rho_{\{3\}}$	1	1	1	1	1	1
$\rho_{\{2,1\}}$	7	3	1	1	0	0
$\rho_{\{1,1,1\}}$	21	5	1	0	0	0

Table 3.1: Flag Characters of GL(3, 2).

We have omitted $\rho_{\{1,2\}}$ because $\rho_{\{1,2\}} = \rho_{\{2,1\}}$. Notice that in the notation of Example 2.1.9, $\rho_{\{3\}} = \chi_1$, the trivial character of G. Since

$$\langle \chi_1, \rho_{\{2,1\}} \rangle = \frac{1}{168}(7 + 21 \cdot 3 + 42 \cdot 1 + 56 \cdot 1 + 24 \cdot 0 + 24 \cdot 0) = 1,$$

χ_1 appears once in the expression of $\rho_{\{2,1\}}$ as a sum of irreducible characters. Hence $\rho_{\{2,1\}} - \chi_1$ is a character of G with fewer irreducible constituents. In fact, the values of $\rho_{\{2,1\}} - \chi_1$ are

$$6 \quad 2 \quad 0 \quad 0 \quad -1 \quad -1\ ,$$

so $\rho_{\{2,1\}} - \chi_1 = \chi_4$ is an irreducible character of G, as may easily be confirmed by checking the $\langle \chi_4, \chi_4 \rangle = 1$. Similarly, we can compute that

$$\langle \chi_1, \rho_{\{1,1,1\}} \rangle = 1 \qquad \text{and} \qquad \langle \chi_4, \rho_{\{1,1,1\}} \rangle = 2,$$

so we obtain the reduced character $\rho_{\{1,1,1\}} - \chi_1 - 2\chi_4$ with values

$$8 \quad 0 \quad 0 \quad -1 \quad 1 \quad 1\ .$$

This character is equal to χ_6, and is also irreducible.

The general situation for $\mathrm{GL}(n,q)$ is analogous. As proven by Steinberg in [11], there is, for each partition λ of n, an irreducible character Γ_λ of $\mathrm{GL}(n,q)$ which is obtained as an alternating sum of characters of the form ρ_μ. For a given partition λ, the partitions μ for which ρ_μ appears in the alternating sum are all less than or equal to λ under a certain partial ordering of the partitions of n. Hence there is an ordering $\lambda_1, \lambda_2, \ldots, \lambda_{p(n)}$, where $p(n)$ is the number of partitions of n, such that $\rho_{\lambda_1} = \Gamma_{\lambda_1}$ is irreducible, and for each $i \geq 2$, $\Gamma_{\lambda_1}, \ldots, \Gamma_{\lambda_i}$ are the only irreducible constituents of ρ_{λ_i}, so that Γ_{λ_i} can be computed by subtracting multiples of the already determined characters $\Gamma_{\lambda_1}, \ldots, \Gamma_{\lambda_{i-1}}$ from ρ_{λ_i}, with the particular multiple being determined by taking inner products. These characters are referred to as the *unipotent* characters of $\mathrm{GL}(n,q)$, and they constitute one of the main families of irreducible characters of $\mathrm{GL}(n,q)$.

In group representation theory, one is often interested not only in finding the irreducible characters of a group G, but also in constructing the irreducible representations of G to which they correspond. This problem is discussed for the unipotent characters of $\mathrm{GL}(n,q)$ in [9], in which James shows how to construct in a natural way subrepresentations of the flag representations whose characters are the unipotent characters. He also discusses various properties of these representations, proving, for example, that if μ and λ are compositions of n which yield the same partition if their entries are arranged in increasing order, then the permutation representations on F_λ and F_μ are isomorphic. (Thus the equality $\rho_{\{1,2\}} = \rho_{\{2,1\}}$ noted in Example 3.1.9 is a special case of a more general result.)

Our goal in this thesis is to provide a foundation for a similar study of the flag representations of $\mathrm{Sp}(n,q)$, and to determine the feasibility of using the corresponding characters to compute irreducible characters of $\mathrm{Sp}(n,q)$. In attempting to study the flag representations of $\mathrm{Sp}(n,q)$, one immediately encounters a striking difference between the flag actions of general linear groups and symplectic groups.

3.1.10 Proposition. If V is a vector space of dimension n, then for any partition λ of n, the action of $\mathrm{GL}(V)$ on F_λ is transitive. However the corresponding action of $\mathrm{Sp}(V)$ (assuming n is even) is not transitive unless λ is $\{n\}, \{n-1,1\}$, or $\{1, n-1\}$.

Proof. If $\mathcal{U}$ and $\mathcal{W}$ are any partial flags of the same type, we can add additional subspaces to each to obtain full flags $\mathcal{U}'$ and $\mathcal{W}'$ having $\mathcal{U}$ and $\mathcal{W}$ as subflags. If there exists $T \in \mathrm{GL}(V)$ such that $T\mathcal{U}' = \mathcal{W}'$, then certainly we must have $T\mathcal{U} = \mathcal{W}$. Hence it suffices to prove that $\mathrm{GL}(V)$ acts transitively on full flags. To accomplish this, let $\{e_1, e_2, \ldots, e_n\}$ be a basis of V, let

$$\mathcal{E} = \{E_0 \supset E_1 \supset \cdots \supset E_n\},$$

where

$$E_i = \langle e_1, e_2, \ldots, e_{n-i} \rangle,$$

and let

$$\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_n\}$$

be any other full flag of V. Choose a basis $\{w_1\}$ of W_{n-1}, and extend this to a basis $\{w_1, w_2\}$ of W_{n-2}. Iterating this process, we eventually obtain a basis $\{w_1, w_2, \ldots, w_n\}$ of V such that $W_i = \langle w_1, w_2, \ldots, w_{n-i}\rangle$. Let T be the unique element of $\mathrm{GL}(V)$ satisfying $Te_i = w_i$. Then $TE_i = W_i$, and so $T\mathcal{E} = \mathcal{W}$, so that the orbit of $\mathcal{E}$ is the entire set of full flags of V.

If n is even, then we may also consider the action of $\mathrm{Sp}(V)$ on F_λ. Once we have developed the concept of the incidence matrix of a flag, it will be easy to see when this action is or is not transitive. The reason $\mathrm{Sp}(V)$ fails to be transitive on F_λ in most cases is that elements of $\mathrm{Sp}(V)$ preserve B, and hence preserve the dimensions of the intersections of subspaces and orthogonal complements of subspaces from the flags, and except in the three cases mentioned, there are different possibilities for these dimensions. For example, if $n = 2m \geq 4$, and B has the matrix with m copies of $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ on the diagonal, then $\dim \mathrm{Rad}\langle e_1, e_2\rangle = 0$, whereas $\dim \mathrm{Rad}\langle e_1, e_3\rangle = 2$, so the $\{n-2, 2\}$-flags corresponding to $\langle e_1, e_2\rangle$ and $\langle e_1, e_3\rangle$ are contained in different orbits.

□

In general, if a finite group G acts on a set X, and $\mathcal{O}_1, \ldots, \mathcal{O}_k$ are the orbits of this action, then G acts on each $\mathcal{O}_i$ by restriction, and so the permutation representation Φ_X is isomorphic to the direct sum $\oplus_{i=1}^k \Phi_{\mathcal{O}_i}$. Thus any attempt at decomposing a permutation representation into irreducible subrepresentations should begin with finding the orbits of the corresponding action. In this chapter, we will determine the orbits of $\mathrm{Sp}(V)$ on F_λ for any composition λ. Although we are primarily interested in the case in which $\mathbb{F}$ is a finite field, the decomposition into orbits turns out to be independent of $\mathbb{F}$, so we will present the solution over an arbitrary field. Our main tool will be introduced in the next section.

3.2 The Incidence Matrix of a Flag

Throughout this section, let $\mathbb{F}$ be any field, and let $V \simeq \mathbb{F}^{2m}$ be the vector space of dimension $n = 2m$ over $\mathbb{F}$. Let B be the symplectic form on V whose matrix with respect to the basis

$$\{e_1, \ldots, e_{2m}\}$$

is

$$\begin{pmatrix} \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & & \\ & \ddots & \\ & & \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \end{pmatrix}.$$

(Recall that any symplectic space has a basis with respect to which the symplectic form has the above matrix, so there is no loss of generality here.) As a result, if $i \leq j$, then

$$B(e_i, e_j) = \begin{cases} 1, & \text{if } (i,j) = (2l-1, 2l) \text{ for some } l, \\ 0, & \text{otherwise.} \end{cases}$$

This fact will be used throughout this chapter in determining the orthogonal complements of various subspaces. We are now ready for the central definition of this section.

3.2.1 Definition. Let λ be a composition of n, and let

$$\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_k\}$$

be a λ-flag of V. Then the *incidence matrix* of $\mathcal{W}$ is the $(k-1) \times (k-1)$ matrix $\mathcal{I}(\mathcal{W})$ with (i,j) entry given by

$$\mathcal{I}(\mathcal{W})_{i,j} = \dim(W_i^{\perp} \cap W_j).$$

If the need arises, we may subscript $\mathcal{I}(\mathcal{W})$ with the partition λ.

Note that since the intersection of any subspace with V or with 0 is determined automatically, $\mathcal{I}(\mathcal{W})$ has entries only for the proper non-trivial subspaces of $\mathcal{W}$. However, for notational convenience we may occasionally use the notation $\mathcal{I}(\mathcal{W})_{i,j}$ to refer to $\dim(W_i^{\perp} \cap W_j)$ even when it is possible that i or j is 0 or k.

The relationship between the incidence matrix of a flag and the incidence matrices of its subflags is fairly obvious: deleting a subspace from a flag corresponds to deleting the row and column involving it from the incidence matrix. It will be useful to establish some notation which describes this process.

3.2.2 Definition. Let $\lambda = \{\lambda_1, \ldots, \lambda_k\}$ and $\mu = \{\mu_1, \ldots, \mu_l\}$ be compositions of n, and suppose that $\mu \preceq \lambda$, with

$$\mu_i = \sum_{j=n_{i-1}+1}^{n_i} \lambda_j$$

for certain $n_i \in \mathbb{Z}$. If A is any $(k-1) \times (k-1)$ matrix, define $D_{\lambda,\mu}(A)$ to be the $(l-1) \times (l-1)$ matrix obtained by deleting each entry of A whose row or column index is not equal to $\sum_{j=1}^{m} n_j$ for any m. (If $\lambda = \{1, \ldots, 1\}$, we denote $D_{\lambda,\mu}$ simply by D_μ.)

The following proposition makes precise the relationship between the incidence matrices of a flag and its subflags.

3.2.3 Proposition. In the notation of Definition 3.2.2, if $\mathcal{W}$ is any λ-flag of V, we have $\mathcal{I}(\mathcal{W}_\mu) = D_{\lambda,\mu}(\mathcal{I}(\mathcal{W}))$.

Proof. The entries of $\mathcal{I}(\mathcal{W})$ with row and column indices of the form $\sum_{j=1}^{m} n_j$ are precisely the entries which correspond to subspaces of $\mathcal{W}$ contained in $\mathcal{W}_\mu$, and therefore by deleting all other entries, we obtain $\mathcal{I}(\mathcal{W}_\mu)$. □

We now give several examples of incidence matrices.

3.2.4 Example. If $\mathcal{W} = \{V \supset W \supset 0\}$ is any $\{n-r, r\}$-flag of V, then

$$\mathcal{I}(\mathcal{W}) = \dim \operatorname{Rad} W.$$

3.2.5 Example. If $n = 6$ and $\mathcal{W} =$

$$\{V \supset \langle e_1, e_2, e_3, e_4, e_5\rangle \supset \langle e_1, e_2, e_4, e_5\rangle \supset \langle e_1, e_2, e_5\rangle \supset \langle e_1, e_5\rangle \supset \langle e_5\rangle \supset 0\},$$

then to compute $\mathcal{I}(\mathcal{W})$, we first note that $\mathcal{W}^\perp =$

$$\{0 \subset \langle e_5\rangle \subset \langle e_4, e_5\rangle \subset \langle e_3, e_4, e_5\rangle \subset \langle e_1, e_3, e_4, e_5\rangle \subset \langle e_1, e_2, e_3, e_4, e_5\rangle \subset V\}.$$

Computing $\mathcal{I}(\mathcal{W})$ amounts to determining the dimensions of the intersections of all possible pairs of subspaces from $\mathcal{W}^\perp$ and $\mathcal{W}$. Since each subspace is spanned by certain e_i, the dimension of each intersection is just the number of indices appearing in the bases of both subspaces. Thus $\mathcal{I}(\mathcal{W})$ can be computed by filling in the following array, with each entry being the number of digits which appear in both the row and column labels.

	12345	1245	125	15	5
5					
45					
345					
1345					
12345					

Hence

$$\mathcal{I}(\mathcal{W}) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 & 1 \\ 3 & 2 & 1 & 1 & 1 \\ 4 & 3 & 2 & 2 & 1 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}.$$

3.2.6 Example. If $n = 6$ and $\mathcal{W} =$

$$\{V \supset \langle e_1, e_2, e_3, e_4, e_5\rangle \supset \langle e_1, e_3, e_4, e_5\rangle \supset \langle e_1, e_3, e_4\rangle \supset \langle e_3, e_4\rangle \supset \langle e_3\rangle \supset 0\},$$

then

$$\mathcal{I}(\mathcal{W}) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 \\ 4 & 3 & 2 & 1 & 1 \end{pmatrix}.$$

Taken together, the two preceding examples show that two flags of the same type need not have the same incidence matrix. It is, however, the case that two flags in the same orbit of $\operatorname{Sp}(V)$ have the same incidence matrix.

3.2.7 Theorem. If $\mathcal{W}$ is any flag of V and if $g \in \mathrm{Sp}(V)$, then $\mathcal{I}(g\mathcal{W}) = \mathcal{I}(\mathcal{W})$.

Proof. Suppose

$$\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_k\},$$

so that

$$g\mathcal{W} = \{gW_0 \supset gW_1 \supset \cdots \supset gW_k\}.$$

Let $1 \leq i, j \leq k-1$. By Lemma 2.2.14, $(gW_i)^\perp \cap gW_j = g(W_i^\perp) \cap gW_j$, and since g is a bijection, $g(W_i^\perp) \cap gW_j = g\left(W_i^\perp \cap W_j\right)$. Hence since any element of $\mathrm{GL}(V)$ preserves subspace dimensions, we have

$$\dim\left((gW_i)^\perp \cap gW_j\right) = \dim g\left(W_i^\perp \cap W_j\right) = \dim W_i^\perp \cap W_j,$$

which shows that $\mathcal{I}(g\mathcal{W})_{i,j} = \mathcal{I}(\mathcal{W})_{i,j}$ as desired.

□

We will eventually prove that the converse of Theorem 3.2.7 also holds, so that incidence matrices completely classify the orbits of the flag actions of $\mathrm{Sp}(V)$. As a first step toward this goal, we abstract some properties of incidence matrices of full flags.

3.3 Incidence Matrices of Full Flags

3.3.1 Proposition. The following properties hold for any full flag $\mathcal{W}$ of V.

1. $\mathcal{I}(\mathcal{W})_{j,i} = \mathcal{I}(\mathcal{W})_{i,j} + j - i$.
2. $\mathcal{I}(\mathcal{W})_{i,j} - \mathcal{I}(\mathcal{W})_{i,j+1} \in \{0, 1\}$.
3. $\mathcal{I}(\mathcal{W})_{i+1,j} - \mathcal{I}(\mathcal{W})_{i,j} \in \{0, 1\}$.
4. $\mathcal{I}(\mathcal{W})_{i,i} \equiv i \mod 2$. In particular, $\mathcal{I}(\mathcal{W})_{1,1} = \mathcal{I}(\mathcal{W})_{n-1,n-1} = 1$.
5. $\mathcal{I}(\mathcal{W})_{i+1,i+1} - \mathcal{I}(\mathcal{W})_{i,i} = \pm 1$.
6. If $\mathcal{I}(\mathcal{W})_{i,j} = \mathcal{I}(\mathcal{W})_{i,k}$, then $\mathcal{I}(\mathcal{W})_{l,j} = \mathcal{I}(\mathcal{W})_{l,k}$ whenever $1 \leq l \leq i$.
7. $\mathcal{I}(\mathcal{W})_{i,j} \in \{0, 1, \ldots, n-1\}$.

Throughout, we assume the various indices are such that elements of the $(n-1) \times (n-1)$ matrix $\mathcal{I}(\mathcal{W})$ are referenced.

Proof. We will use freely Proposition 2.2.6, Lemma 2.2.10, and the linear algebra result which states that if U and W are subspaces of V, then

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

We let $\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_n\}$ as usual.

1. We have

$$\mathcal{I}(\mathcal{W})_{j,i} = \dim\left(W_j^{\perp} \cap W_i\right) = n - \dim\left(W_j^{\perp} \cap W_i\right)^{\perp} = n - \dim\left(W_i^{\perp} + W_j\right)$$
$$= n - \left(\dim W_i^{\perp} + \dim W_j - \dim\left(W_i^{\perp} \cap W_j\right)\right).$$

$\dim W_j = n - j$ by construction, and $\dim W_i^{\perp} = n - \dim W_i = i$, so that

$$\mathcal{I}(\mathcal{W})_{j,i} = n - (i + n - j - \mathcal{I}(\mathcal{W})_{i,j}) = \mathcal{I}(\mathcal{W})_{i,j} + j - i.$$

2. Since $W_{j+1} \subset W_j$, we have that

$$\mathcal{I}(\mathcal{W})_{i,j+1} = \dim\left(W_i^{\perp} \cap W_{j+1}\right) = \dim\left(\left(W_i^{\perp} \cap W_j\right) \cap W_{j+1}\right)$$
$$= \dim\left(W_i^{\perp} \cap W_j\right) + \dim W_{j+1} - \dim\left(\left(W_i^{\perp} \cap W_j\right) + W_{j+1}\right).$$

Now $\dim\left(W_i^{\perp} \cap W_j\right) = \mathcal{I}(\mathcal{W})_{i,j}$. Since $\mathcal{W}$ is a full flag, W_{j+1} is a subspace of W_j of codimension 1, and hence

$$\dim\left(\left(W_i^{\perp} \cap W_j\right) + W_{j+1}\right) = \begin{cases} \dim W_{j+1} & \text{if } W_i^{\perp} \cap W_j \subset W_{j+1}, \\ \dim W_j & \text{otherwise.} \end{cases}$$

Thus

$$\mathcal{I}(\mathcal{W})_{i,j+1} = \begin{cases} \mathcal{I}(\mathcal{W})_{i,j} & \text{if } W_i^{\perp} \cap W_j \subset W_{j+1}, \\ \mathcal{I}(\mathcal{W})_{i,j} - 1 & \text{otherwise,} \end{cases}$$

which proves 2.

3. Property 3 follows by applying Property 1 to the equation in Property 2.

4. By construction, the subspaces $W_0, W_2, \ldots W_n$ have even dimension, while $W_1, W_3, \ldots W_{n-1}$ have odd dimension, so $\dim W_i \equiv i \mod 2$. But

$$\mathcal{I}(\mathcal{W})_{i,i} = \dim\left(W_i^{\perp} \cap W_i\right) = \dim \operatorname{Rad} W_i,$$

and by Corollary 2.2.9,

$$\dim \operatorname{Rad} W_i \equiv \dim W_i \mod 2,$$

so $\mathcal{I}(\mathcal{W})_{i,i} \equiv i \mod 2$. If $i \in \{1, n-1\}$, then one of W_i and $W_i^{\perp}$ has dimension 1. Hence $\mathcal{I}(\mathcal{W})_{i,i} = \dim\left(W_i^{\perp} \cap W_i\right) \leq 1$, and is odd, and so must be 1.

5. We have $\mathcal{I}(\mathcal{W})_{i+1,i+1} - \mathcal{I}(\mathcal{W})_{i,i} =$

$$\left(\mathcal{I}(\mathcal{W})_{i+1,i+1} - \mathcal{I}(\mathcal{W})_{i+1,i}\right) - \left(\mathcal{I}(\mathcal{W})_{i,i} - \mathcal{I}(\mathcal{W})_{i+1,i}\right).$$

By Properties 2 and 3, each of the two terms is 0 or -1, and hence their difference is in $\{-1, 0, 1\}$. But by 4, $\mathcal{I}(\mathcal{W})_{i+1,i+1} - \mathcal{I}(\mathcal{W})_{i,i} \equiv (i+1) - i \equiv 1 \mod 2$, and so $\mathcal{I}(\mathcal{W})_{i+1,i+1} - \mathcal{I}(\mathcal{W})_{i,i} \in \{-1, 1\}$.

6. Without loss of generality, we may assume that $j < k$. Hence $W_j \supset W_k$, and so $W_i^\perp \cap W_j \supset W_i^\perp \cap W_k$. But $\mathcal{I}(\mathcal{W})_{i,j} = \mathcal{I}(\mathcal{W})_{i,k}$, so these two subspaces have the same dimension, and hence $W_i^\perp \cap W_j = W_i^\perp \cap W_k$. Now since $1 \leq l \leq i$, $W_l \supset W_i$, and hence $W_l^\perp \subset W_i^\perp$. Using these facts, we see that

$$\mathcal{I}(\mathcal{W})_{l,j} = \dim\left(W_l^\perp \cap W_j\right) = \dim\left(W_l^\perp \cap \left(W_i^\perp \cap W_j\right)\right)$$

$$= \dim\left(W_l^\perp \cap \left(W_i^\perp \cap W_k\right)\right) = \dim\left(W_l^\perp \cap W_k\right) = \mathcal{I}(\mathcal{W})_{l,k}.$$

7. $\mathcal{I}(\mathcal{W})_{i,j}$ is the dimension of a proper subspace of the n-dimensional vector space V.

□

The properties in Proposition 3.3.1 place strong restrictions on what sort of matrix can be the incidence matrix of some full flag. In fact, we can use them to enumerate all possible incidence matrices of full flags for small dimensions.

3.3.2 Example. If $n = 4$, Properties 4 and 5 imply that the diagonal of any incidence matrix of a $\{1,1,1,1\}$-flag must be $(1,0,1)$ or $(1,2,1)$. By Property 1, the entries below the diagonal are determined by those above it, so we only need to determine the possibilities for the entries above the diagonal. The only way to fill in the above-diagonal entries of $\begin{pmatrix} 1 & & \\ & 0 & \\ & & 1 \end{pmatrix}$ so that the entries decrease along rows, and increase along columns, as required by Properties 2 and 3 is $\begin{pmatrix} 1 & 0 & 0 \\ & 0 & 0 \\ & & 1 \end{pmatrix}$. We can fill in $\begin{pmatrix} 1 & & \\ & 2 & \\ & & 1 \end{pmatrix}$ in two ways, as either $\begin{pmatrix} 1 & 1 & 0 \\ & 2 & 1 \\ & & 1 \end{pmatrix}$ or as $\begin{pmatrix} 1 & 1 & 1 \\ & 2 & 1 \\ & & 1 \end{pmatrix}$. Computing the below-diagonal entries via Property 1, we see that the only possible incidence matrices for $\{1,1,1,1\}$-flags are

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 3 & 2 & 1 \end{pmatrix}.$$

To see that these are actually incidence matrices of full flags, we note that they correspond to the flags

$$\{V \supset \langle e_1, e_2, e_3 \rangle \supset \langle e_1, e_2 \rangle \supset \langle e_1 \rangle \supset 0\},$$

$$\{V \supset \langle e_1, e_2, e_3 \rangle \supset \langle e_1, e_3 \rangle \supset \langle e_1 \rangle \supset 0\},$$

and

$$\{V \supset \langle e_1, e_2, e_3 \rangle \supset \langle e_1, e_3 \rangle \supset \langle e_3 \rangle \supset 0\}.$$

Thus in the case $n = 4$, every 3×3 matrix satisfying the properties of Proposition 3.3.1 is the incidence matrix of some full flag. We will prove shortly that this is true for any n; that is, Proposition 3.3.1 completely describes incidence matrices of full flags. In proving this we will proceed inductively; given a matrix M satisfying the properties of Proposition 3.3.1 and a flag $\{V \supset W_i \supset \cdots \supset W_n\}$ of type $\{i, 1, \ldots, 1\}$ whose incidence matrix is the $(n-i) \times (n-i)$ submatrix in the bottom right corner of M, we will show that it is always possible to choose a vector v so that the incidence matrix of the $\{i-1, 1, \ldots, 1\}$-flag $\{V \supset \langle v \rangle \oplus W_i \supset W_i \supset \cdots \supset W_n\}$ is the $(n-i+1) \times (n-i+1)$ submatrix in the bottom right corner of M. In order to do this, we need to know all the different possibilities which are allowed by Proposition 3.3.1 for those entries of the larger submatrix which correspond to $\langle v \rangle \oplus W_i$, and how to choose v so that each of these possibilities is realized. The following lemma provides precisely this information. (By Property 1 of Proposition 3.3.1, we only need to consider entries above the diagonal.)

3.3.3 Lemma. Let

$$\{W_0 \supset W_1 \supset \cdots \supset W_{i-1} \supset W_i \supset \cdots \supset W_n\}$$

be a full flag of V, and let $2 \leq i \leq n-1$. If $v \in W_{i-1} \backslash W_i$, then $W_{i-1} = \langle v \rangle \oplus W_i$, and the entries on or above the diagonal in rows $i-1$ and i of $\mathcal{I}(\mathcal{W})$ take either the form

$$\begin{matrix} k{+}1 & k & \ldots & k & k{-}1 & \ldots & k{-}1 & \ldots & 1 & \ldots & 1 & 0 & \ldots & 0 \\ & k & \ldots & k & k{-}1 & \ldots & k{-}1 & \ldots & 1 & \ldots & 1 & 0 & \ldots & 0 \end{matrix} \qquad \text{(Form 1)}$$

or the form

$$\begin{matrix} & & & & & & & & & j{-}1 & j & & & & & & & & \\ k{-}1 & k{-}1 & \ldots & k{-}1 & k{-}2 & \ldots & k{-}2 & \ldots & l{-}1 & \ldots & l{-}1 & l{-}1 & \ldots & l{-}1 & \ldots & 1 & \ldots & 1 & 0 & \ldots & 0 \\ & k & \ldots & k & k{-}1 & \ldots & k{-}1 & \ldots & l & \ldots & l & l{-}1 & \ldots & l{-}1 & \ldots & 1 & \ldots & 1 & 0 & \ldots & 0 \end{matrix} \qquad \text{(Form 2)}$$

for some l with $1 \leq l \leq k$, and with the indicated entries lying in columns $j-1$ and j. (It is possible for no zeros to be present at the end of one or both of the rows, or for the row beginning with the smaller number to consist entirely of zeros. It is also possible to have $j = n$, in which case $W_j = 0$.) Further, they take Form 1 if and only if

$$v \in (W_i + W_i^{\perp}) \backslash W_i,$$

and Form 2 if and only if

$$v \in (W_i + W_j^{\perp}) \backslash (W_i + W_{j-1}^{\perp}).$$

Proof. Since W_i has codimension 1 in W_{i-1}, it is automatic that $W_{i-1} = \langle v \rangle \oplus W_i$ for any $v \in W_{i-1} \backslash W_i$. That the above-diagonal entries of rows $i-1$ and i take one of the above-mentioned forms is a direct consequence of the properties in Proposition 3.3.1. First, since the entries of $\mathcal{I}(\mathcal{W})$ increase along columns, and since $\mathcal{I}(\mathcal{W})_{n-1,n-1} = 1$, every entry in the last column of $\mathcal{I}(\mathcal{W})$ is 0 or 1. Hence since entries decrease along rows, if $\mathcal{I}(\mathcal{W})_{i,i} = k$, then the above-diagonal portion of row i of $\mathcal{I}(\mathcal{W})$ must have the form

$$k \; \ldots \; k \; k{-}1 \; \ldots \; k{-}1 \; \ldots \; 1 \; \ldots \; 1 \; 0 \; \ldots \; 0\,,$$

with each number in $\{k, k-1, \ldots, 1\}$ occurring at least once, and 0 occurring zero or more times. Now consider the form row $i-1$ must take. By Property 3 of Proposition 3.3.1, each entry of row $i-1$ is equal to or one less than the entry below it in row i. It is possible that rows $i-1$ and i are identical in the entries in questions, in which case the rows take Form 1. Otherwise, let j be the index of the leftmost column in which rows i and $i-1$ have the same entry, and let that entry be $l-1$ for some l with $1 \leq l \leq k$. (If rows i and $i-1$ do not have the same entry in any column, we take j to be n.) By the properties of Proposition 3.3.1, we see that rows $i-1$ and i must have the same entries in all columns with index $j' \geq j$, and by construction, their entries differ by 1 in all columns with index $j' < j$. In particular, in order for Property 2 of Proposition 3.3.1 to hold, the entries in column $j-1$ must be $l-1$ and l. Thus the entries in question take Form 2 for the number l and index j.

To prove that the entries have Form 1 if and only if $v \in (W_i + W_i^{\perp}) \backslash W_i$, suppose first that $v \in (W_i + W_i^{\perp}) \backslash W_i$. Since $W_{i-1} = \langle v \rangle + W_i$, we have

$$W_{i-1}^{\perp} \cap W_i = (\langle v \rangle + W_i)^{\perp} \cap W_i = \langle v \rangle^{\perp} \cap \left(W_i^{\perp} \cap W_i\right).$$

Now

$$v \in W_i + W_i^{\perp} = \left(W_i^{\perp} \cap W_i\right)^{\perp},$$

so $W_i^{\perp} \cap W_i \subset \langle v \rangle^{\perp}$, and hence $\langle v \rangle^{\perp} \cap \left(W_i^{\perp} \cap W_i\right) = W_i^{\perp} \cap W_i$. Hence

$$W_{i-1}^{\perp} \cap W_i = W_i^{\perp} \cap W_i,$$

and so $\mathcal{I}(\mathcal{W})_{i-1,i} = \mathcal{I}(\mathcal{W})_{i,i}$, which occurs only when the entries have Form 1.

Suppose conversely that the entries in question have Form 1. Then by Property 1, $\mathcal{I}(\mathcal{W})_{i,i-1} = k+1$. Hence

$$\dim\left(W_i^{\perp} \cap W_{i-1}\right) = k+1 > \dim\left(W_i^{\perp} \cap W_i\right) = k,$$

so there is some vector $u \in W_i^{\perp} \cap W_{i-1} \backslash W_i^{\perp} \cap W_i$. Since $u \in W_{i-1} = \langle v \rangle + W_i$, there exist $\lambda \in \mathbb{F}$ and $w \in W_i$ such that

$$u = \lambda v + w.$$

Since $u \notin W_i$, $\lambda \neq 0$, and hence $v = \lambda^{-1} u - \lambda^{-1} w$, which shows that $v \in \left(W_i + W_i^{\perp}\right) \backslash W_i$.

Finally, we prove that the entries have Form 2 for the number l and index j if and only if $v \in (W_i + W_j^{\perp}) \backslash (W_i + W_{j-1}^{\perp})$. Suppose first that $v \in (W_i + W_j^{\perp}) \backslash (W_i + W_{j-1}^{\perp})$, for some j, and let $l = \mathcal{I}(\mathcal{W})_{i,j-1}$. We claim that the submatrix

$$\begin{matrix} \mathcal{I}(\mathcal{W})_{i-1,j-1} & \mathcal{I}(\mathcal{W})_{i-1,j} \\ \mathcal{I}(\mathcal{W})_{i,j-1} & \mathcal{I}(\mathcal{W})_{i,j} \end{matrix} \qquad \text{equals} \qquad \begin{matrix} l-1 & l-1 \\ l & l-1 \end{matrix}\,.$$

Since $(W_i + W_j^\perp)\backslash(W_i + W_{j-1}^\perp)$ contains v,

$$W_i + W_j^\perp \supsetneq W_i + W_{j-1}^\perp \qquad \Rightarrow \qquad W_i^\perp \cap W_j \subsetneq W_i^\perp \cap W_{j-1},$$

so that $\mathcal{I}(\mathcal{W})_{i,j} < \mathcal{I}(\mathcal{W})_{i,j-1}$, which means $\mathcal{I}(\mathcal{W})_{i,j} = l-1$ by Property 2. Once again using $W_{i-1} = \langle v \rangle + W_i$, we have

$$W_{i-1}^\perp \cap W_j = \langle v \rangle^\perp \cap \left(W_i^\perp \cap W_j\right).$$

Since $v \in W_i + W_j^\perp$, $W_i^\perp \cap W_j \subset \langle v \rangle^\perp$, and so the preceding equality reduces to

$$W_{i-1}^\perp \cap W_j = W_i^\perp \cap W_j,$$

which implies that $\mathcal{I}(\mathcal{W})_{i-1,j} = \mathcal{I}(\mathcal{W})_{i,j} = l-1$. By construction, $v \in W_{i-1}$, so $v \in W_{i-1} + W_{j-1}^\perp$. But by hypothesis, $v \notin W_i + W_{j-1}^\perp$, so $W_i + W_{j-1}^\perp \subsetneq W_{i-1} + W_{j-1}^\perp$. Taking orthogonal complements then shows that

$$W_{i-1}^\perp \cap W_{j-1} \subsetneq W_i^\perp \cap W_{j-1},$$

and so $\mathcal{I}(\mathcal{W})_{i-1,j-1} < \mathcal{I}(\mathcal{W})_{i,j-1}$, which implies that $\mathcal{I}(\mathcal{W})_{i-1,j-1} = l-1$ by Property 3. We have shown that the entries in question include the submatrix

$$\begin{matrix} l-1 & l-1 \\ l & l-1 \end{matrix},$$

and so the entries must have Form 2 for the number l, as this is the only form in which such a submatrix can occur.

Suppose conversely that the entries have Form 2 for the number l and index j. Then $\mathcal{I}(\mathcal{W})_{i-1,j-1} = \mathcal{I}(\mathcal{W})_{i-1,j} = \mathcal{I}(\mathcal{W})_{i,j} = l-1$, but $\mathcal{I}(\mathcal{W})_{i,j-1} = l$, and so

$$W_{i-1}^\perp \cap W_{j-1} = W_{i-1}^\perp \cap W_j = W_i^\perp \cap W_j,$$

and all three subspaces are properly contained in $W_i^\perp \cap W_{j-1}$. Taking orthogonal complements, we see that

$$W_{i-1} + W_{j-1}^\perp = W_{i-1} + W_j^\perp = W_i + W_j^\perp,$$

and that all three subspaces properly contain $W_i + W_{j-1}^\perp$. Hence we may choose some vector $w \in \left(W_i + W_j^\perp\right) \setminus \left(W_i + W_{j-1}^\perp\right)$, and write $w = w_1 + w_2$, with $w_1 \in W_i$ and $w_2 \in W_j^\perp \backslash W_{j-1}^\perp$. By the equality of subspaces mentioned above, we also know that

$$w \in W_{i-1} + W_{j-1}^\perp = \langle v \rangle + W_i + W_{j-1}^\perp,$$

so we can write $w = \lambda v + w_3 + w_4$ for some $\lambda \in \mathbb{F}$, $w_3 \in W_i$, and $w_4 \in W_{j-1}^\perp$. Since $w \notin W_i + W_{j-1}^\perp$, $\lambda \neq 0$, and so we can equate the two expressions for w and solve for v to obtain

$$v = \lambda^{-1}(w_1 - w_3) + \lambda^{-1}(w_2 - w_4).$$

By considering the subspace in which each w_α is contained, and noting in particular that $W_{j-1}^\perp \subset W_j^\perp$, we see that this expression for v shows that $v \in W_i + W_j^\perp$. However, if v were in $W_i + W_{j-1}^\perp$, the equation $w = \lambda v + w_3 + w_4$ would force w to be in $W_i + W_{j-1}^\perp$, contrary to our choice of w, so $v \notin W_i + W_{j-1}^\perp$, as desired. $\square$

With this technical lemma established, we can now prove that the properties of Proposition 3.3.1 completely determine incidence matrices of full flags.

3.3.4 Theorem. Any $(n-1) \times (n-1)$ matrix having the properties listed in Proposition 3.3.1 is the incidence matrix of some full flag of V.

Proof. Let M be such a matrix. Since M satisfies Property 1, we only need to show that the upper triangular portion of M is the upper triangular portion of $\mathcal{I}(\mathcal{W})$ for some flag $\mathcal{W}$. We will proceed by downward induction, and show that for any i with $1 \le i \le n-1$, there exists an $\{i, 1, \ldots, 1\}$-flag of V whose incidence matrix is $D_{\{i,1,\ldots,1\}}(M)$, the $(n-i) \times (n-i)$ submatrix in the bottom right corner of M. When $i = n-1$, this is automatic; for by Property 4, $M_{n-1,n-1} = 1$, and if w is any non-zero vector in V, the sole entry of the incidence matrix of the $\{n-1, 1\}$-flag $\{V \supset \langle w \rangle \supset 0\}$ is $\dim\left(\langle w \rangle^\perp \cap \langle w \rangle\right) = \dim\langle w \rangle = 1$. Now suppose that for some i, with $2 \le i \le n-1$, there is an $\{i, 1, \ldots, 1\}$-flag

$$\mathcal{W} = \{V \supset W_i \supset W_{i+1} \supset \cdots \supset W_n = 0\}$$

whose incidence matrix is $D_{\{i,1,\ldots,1\}}(M)$. (We have indexed the subspaces of $\mathcal{W}$ as if they were the corresponding subspaces of the same dimension in a full flag of V.) Since M satisfies the properties of Proposition 3.3.1, the above-diagonal entries of M in rows $i-1$ and i must take one of the forms described in Lemma 3.3.3. (Recall that these forms were derived entirely from the properties of Proposition 3.3.1.)

Suppose first that they take Form 1, so that $M_{i,i} = k$ and $M_{i-1,i-1} = k+1$ for some k. We claim that $k < i$. To see this, observe that since $M_{1,1} = 1$, and since by Property 5 entries increase by at most 1 from one diagonal entry to the next, $M_{j,j} \le j$ for any j, with $M_{j,j} = j$ occurring only when the diagonal entries increase by 1 at every step up to j. Since $M_{i-1,i-1} = k+1 > M_{i,i}$, entries do not increase at every step before i, and so $M_{i,i} < i$. Hence since $D_{\{i,1,\ldots,1\}}(M)$ is the incidence matrix of $\mathcal{W}$,

$$\dim\left(W_i + W_i^\perp\right) = n - \dim\left(W_i^\perp \cap W_i\right) = n - k > n - i = \dim W_i.$$

Thus $W_i + W_i^\perp$ properly contains W_i, and so there is some vector $v \in (W_i + W_i^\perp) \backslash W_i$. If we set

$$\mathcal{W}' = \{V \supset \langle v \rangle + W_i \supset W_i \supset W_{i+1} \supset \cdots \supset W_n\},$$

then by Lemma 3.3.3, the incidence matrix of $\mathcal{W}'$ is $D_{\{i-1,1,\ldots,1\}}(M)$.

Now suppose that the above-diagonal entries in rows $i-1$ and i are of Form 2 for some l and j as specified in the notation of Lemma 3.3.3. Then since $D_{\{i,1,\ldots,1\}}(M)$ is the incidence matrix of $\mathcal{W}$, we have

$$\dim\left(W_i + W_j^{\perp}\right) = n - \dim\left(W_i^{\perp} \cap W_j\right) = n - M_{i,j}$$

$$= n - (l-1) > n - l = n - \dim\left(W_i^{\perp} \cap W_{j-1}\right) = \dim\left(W_i + W_{j-1}^{\perp}\right),$$

and so there exists some vector $v \in (W_i + W_j^{\perp}) \setminus (W_i + W_{j-1}^{\perp})$. Hence by Lemma 3.3.3, if we set

$$\mathcal{W}' = \{V \supset \langle v \rangle + W_i \supset W_i \supset W_{i+1} \supset \cdots \supset W_n\},$$

then $\mathcal{I}(\mathcal{W}') = D_{\{i-1,1,\ldots,1\}}(M)$.

Thus in all cases we can find an $\{i-1, 1, \ldots, 1\}$-flag whose incidence matrix is $D_{\{i-1,1,\ldots,1\}}(M)$, and so, by induction, each $D_{\{i,1,\ldots,1\}}(M)$ is the incidence matrix of some $\{i, 1, \ldots, 1\}$-flag. In particular, we may set $i = 1$ to conclude that M itself is the incidence matrix of some full flag. □

3.4 The Orbit Classification Theorem

In proving the converse of Theorem 3.2.7, we will proceed by induction on the dimension of the space. Most of the technicalities needed in the induction step are contained in the following lemma, which gives a procedure for relating a full flag of a symplectic space V to a full flag of a non-degenerate index 2 subspace of V in a canonical way, and describes how incidence matrices behave under the procedure.

3.4.1 Lemma. Let V be the symplectic space of dimension $n = 2m$, where $n \geq 4$, and let

$$\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_n\}$$

be a full flag of V. Then there exists a unique $j \in \{0, 1, \ldots, n-2\}$ such that the last column of $\mathcal{I}(\mathcal{W})$ is $(0, \ldots, 0, 1, \ldots, 1)^T$, with the first 1 appearing in row $j+1$. There exist $w \in W_{n-1}$ and $v \in W_j$ with $B(v, w) = 1$, and for any such v, w, if we set

$$U = \langle v, w \rangle^{\perp} \qquad \text{and} \qquad U_i = U \cap W_i,$$

then $W_i = U_i \oplus \langle v, w \rangle$ if $0 \leq i \leq j$ and $W_i = U_i \oplus \langle w \rangle$ if $j+1 \leq i \leq n-1$. Further, U is isomorphic to the symplectic space of dimension $n-2$ and

$$\mathcal{U} = \{U_0 \supset U_1 \supset \cdots \supset U_j \supset U_{j+2} \supset \cdots \supset U_{n-1}\}$$

is a full flag of U whose incidence matrix has entries given by

$$\mathcal{I}(\mathcal{U})_{k,l} = \begin{cases} \mathcal{I}(\mathcal{W})_{k,l} & \text{if } k, l \leq j \\ \mathcal{I}(\mathcal{W})_{k,l+1} & \text{if } k \leq j \text{ and } l \geq j+1 \\ \mathcal{I}(\mathcal{W})_{k+1,l} - 1 & \text{if } k \geq j+1 \text{ and } l \leq j \\ \mathcal{I}(\mathcal{W})_{k+1,l+1} - 1 & \text{if } k, l \geq j+1 \end{cases}.$$

Proof. That the last column of $\mathcal{I}(\mathcal{W})$ takes the specified form for a unique $j \in \{0, 1, \ldots, n-2\}$ is a direct consequence of Properties 3 and 4 of Proposition 3.3.1. For this j we have $W_j^\perp \cap W_{n-1} = 0$, which implies that the one-dimensional subspace W_{n-1} is not contained in $W_j^\perp$. Hence there must exist $w \in W_{n-1}$ and $v \in W_j$ such that $B(v, w) \neq 0$, and by scaling w appropriately we can assume that $B(v, w) = 1$. Now let $U = \langle v, w \rangle^\perp$ and $U_i = U \cap W_i$. Since the subspace $\langle v, w \rangle$ is non-degenerate, Proposition 2.2.7 implies that

$$V = U \oplus \langle v, w \rangle.$$

Now suppose that $0 \leq i \leq j$. If $x \in W_i$, then by the direct sum decomposition given for V, there exist unique $u \in U$ and $y \in \langle v, w \rangle$ such that $x = u + y$. Since $\langle v, w \rangle \subset W_i$, $y \in W_i$, and so also $u \in W_i$. Thus $x \in U_i \oplus \langle v, w \rangle$, and we conclude that $W_i \subset U_i \oplus \langle v, w \rangle$. The opposite inclusion is automatic, so

$$W_i = U_i \oplus \langle v, w \rangle.$$

If $j + 1 \leq i \leq n - 1$, then since $\mathcal{I}(\mathcal{W})_{i,n-1} = 1$, $\dim \big(W_i^\perp \cap W_{n-1} \big) = 1$, which implies that $W_i \subset W_{n-1}^\perp = \langle w \rangle^\perp$. If $x \in W_i$, then by the direct sum decomposition of V we can write uniquely $x = u + \lambda v + \mu w$ for $u \in U$ and $\lambda, \mu \in \mathbb{F}$. Since x, u and w are all contained in $\langle w \rangle^\perp$, we must have $\lambda v \in \langle w \rangle^\perp$, and so since $B(v, w) = 1$, $\lambda = 0$. Thus $x = u + \mu w$, and so since $w \in W_i$, $u \in W_i$ as well. This shows that $x \in U_i \oplus \langle w \rangle$, and so $W_i \subset U_i \oplus \langle w \rangle$. Once again, the opposite inclusion is automatic, so

$$W_i = U_i \oplus \langle w \rangle.$$

Since $\langle v, w \rangle$ is a non-degenerate subspace of dimension 2, U is a non-degenerate subspace of dimension $n - 2$, and hence is isomorphic to the symplectic space of dimension $n - 2$. By comparing the dimensions of U_i and W_i, we see that

$$\dim U_i = \begin{cases} (n-i) - 2 & \text{if } 0 \leq i \leq j \\ (n-i) - 1 & \text{if } j+1 \leq i \leq n-1 \end{cases},$$

and the subspaces U_i are nested by definition, so noting that $U_j = U_{j+1}$, we conclude that

$$\mathcal{U} = \{U_0 \supset U_1 \supset \cdots \supset U_j \supset U_{j+2} \supset \cdots \supset U_{n-1}\}$$

is a full flag of U. It remains to justify the formula for the incidence matrix of $\mathcal{U}$. In doing this, we will need to discuss orthogonal complements of subspaces in both U and W; by the notation $U_k^\perp$ we will mean the orthogonal complement of U_k in W, so if we wish to refer to the orthogonal complement of U_k in U, we will use the fact that this subspace equals $U_k^\perp \cap U$. Let $0 \leq l \leq n - 1$. Observe first that if $0 \leq k \leq j$, then

$$W_k^\perp \cap W_l = (U_k \oplus \langle v, w \rangle)^\perp \cap W_l = U_k^\perp \cap U \cap W_l$$

$$= \left(U_k^\perp \cap U\right) \cap (U \cap W_l) = \left(U_k^\perp \cap U\right) \cap U_l.$$

If $j+1 \le k \le n-1$, then $\dim\left(W_k^\perp \cap W_{n-1}\right) = 1$, so $\langle w\rangle \subset W_k^\perp$, but also, $\langle w\rangle \subset W_k$, so $W_k^\perp \subset \langle w\rangle^\perp$. Hence since $\langle w\rangle \subset W_l$ as well,

$$\langle w\rangle \subset W_k^\perp \cap W_l \subset \langle w\rangle^\perp.$$

If $x \in W_k^\perp \cap W_l$, then, as before, there exist unique $u \in U$ and $\lambda, \mu \in \mathbb{F}$ such that $x = u + \lambda v + \mu w$. Since x, u, and w are all contained in $\langle w\rangle^\perp$, $\lambda v \in \langle w\rangle^\perp$, and so $\lambda = 0$. Hence since x and w are contained in $W_k^\perp \cap W_l$, so is u, and thus $W_k^\perp \cap W_l \subset \left(\left(W_k^\perp \cap W_l\right) \cap U\right) \oplus \langle w\rangle$. The opposite inclusion is again automatic, so

$$W_k^\perp \cap W_l = \left(\left(W_k^\perp \cap W_l\right) \cap U\right) \oplus \langle w\rangle.$$

But (noting that $U \subset \langle w\rangle^\perp$) we have

$$\left(W_k^\perp \cap W_l\right) \cap U = \left((U_k \oplus \langle w\rangle)^\perp \cap U\right) \cap U_l$$

$$= \left(U_k^\perp \cap \langle w\rangle^\perp \cap U\right) \cap U_l = \left(U_k^\perp \cap U\right) \cap U_l,$$

and so

$$W_k^\perp \cap W_l = \left(\left(U_k^\perp \cap U\right) \cap U_l\right) \oplus \langle w\rangle.$$

As a result of these expressions for $W_k^\perp \cap W_l$, we see that

$$\dim\left(\left(U_k^\perp \cap U\right) \cap U_l\right) = \begin{cases} \dim\left(W_k^\perp \cap W_l\right) & \text{if } 0 \le k \le j \\ \dim\left(W_k^\perp \cap W_l\right) - 1 & \text{if } j+1 \le k \le n-1 \end{cases}.$$

Hence noting the shift in indexing which occurs in the flag $\mathcal{U}$ as a result of the equality of U_j and U_{j+1}, we see that

$$\mathcal{I}(\mathcal{U})_{k,l} = \begin{cases} \dim\left(\left(U_k^\perp \cap U\right) \cap U_l\right) & \text{if } k, l \le j \\ \dim\left(\left(U_k^\perp \cap U\right) \cap U_{l+1}\right) & \text{if } k \le j \text{ and } l \ge j+1 \\ \dim\left(\left(U_{k+1}^\perp \cap U\right) \cap U_l\right) & \text{if } k \ge j+1 \text{ and } l \le j \\ \dim\left(\left(U_{k+1}^\perp \cap U\right) \cap U_{l+1}\right) & \text{if } k, l \ge j+1 \end{cases}$$

$$= \begin{cases} \mathcal{I}(\mathcal{W})_{k,l} & \text{if } k, l \le j \\ \mathcal{I}(\mathcal{W})_{k,l+1} & \text{if } k \le j \text{ and } l \ge j+1 \\ \mathcal{I}(\mathcal{W})_{k+1,l} - 1 & \text{if } k \ge j+1 \text{ and } l \le j \\ \mathcal{I}(\mathcal{W})_{k+1,l+1} - 1 & \text{if } k, l \ge j+1 \end{cases}.$$

□

We can now prove the converse of Theorem 3.2.7 for full flags.

3.4.2 Theorem. If

$$\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_n\} \qquad \text{and} \qquad \mathcal{W}' = \{W'_0 \supset W'_1 \supset \cdots \supset W'_n\}$$

are two full flags of V such that $\mathcal{I}(\mathcal{W}) = \mathcal{I}(\mathcal{W}')$, then there exists $g \in \mathrm{Sp}(V)$ such that $g\mathcal{W} = \mathcal{W}'$.

Proof. We proceed by induction on the dimension n of V, which we recall is necessarily even. If $n = 2$, then any full flag of V has the form $\mathcal{W} = \{V \supset \langle w \rangle \supset 0\}$ for some non-zero $w \in V$, so we can identify full flags with lines in V. Thus all full flags of V have incidence matrix (1). Further, by Proposition 2.2.12, $\mathrm{Sp}(V) = \mathrm{SL}(V)$ when $n = 2$, and $\mathrm{SL}(V)$ acts transitively on the set of lines in V for any vector space V.

Now assume that for some even number $k \geq 2$, $\mathrm{Sp}(W)$ acts transitively on the set of full flags of W with a given incidence matrix for all symplectic spaces W with $\dim W \leq k$, and suppose V is a symplectic space of dimension $k + 2$. Suppose

$$\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_{k+2}\} \qquad \text{and} \qquad \mathcal{W}' = \{W'_0 \supset W'_1 \supset \cdots \supset W'_{k+2}\}$$

are two full flags of V such that $\mathcal{I}(\mathcal{W}) = \mathcal{I}(\mathcal{W}')$. Then we can apply Lemma 3.4.1 to both $\mathcal{W}$ and $\mathcal{W}'$ to conclude that for some index j determined solely from the common incidence matrix of $\mathcal{W}$ and $\mathcal{W}'$, there exist vectors $v \in W_j$, $v' \in W'_j$, $w \in W_{k+1}$, and $w' \in W'_{k+1}$ such that $B(v, w) = B(v', w') = 1$, and if $U = \langle v, w\rangle^\perp$, $U' = \langle v', w'\rangle^\perp$, $U_i = U \cap W_i$, and $U'_i = U_i \cap W'_i$, then

$$\mathcal{W} = \{U_0 \oplus \langle v, w\rangle \supset \cdots \supset U_j \oplus \langle v, w\rangle \supset U_{j+1} \oplus \langle w\rangle \supset \cdots \supset U_{k+1} \oplus \langle w\rangle \supset 0\}$$

and

$$\mathcal{W}' = \{U'_0 \oplus \langle v', w'\rangle \supset \cdots \supset U'_j \oplus \langle v', w'\rangle \supset U'_{j+1} \oplus \langle w'\rangle \supset \cdots \supset U'_{k+1} \oplus \langle w'\rangle \supset 0\}.$$

By Proposition 2.2.15, there exists $g_0 \in \mathrm{Sp}(V)$ such that $g_0 v' = v$ and $g_0 w' = w$. Then by Lemma 2.2.14

$$g_0 U' = g_0 \left(\langle v', w'\rangle^\perp\right) = (g_0\langle v', w'\rangle)^\perp = \langle v, w\rangle^\perp = U.$$

Hence if we let $\mathcal{W}'' = g_0\mathcal{W}'$, and $U''_i = g_0 U'_i$, we have that

$$\mathcal{W}'' = \{U''_0 \oplus \langle v, w\rangle \supset \cdots \supset U''_j \oplus \langle v, w\rangle \supset U''_{j+1} \oplus \langle w\rangle \supset \cdots \supset U''_{k+1} \oplus \langle w\rangle \supset 0\}.$$

(Note that $U''_0 = U_0 = U$.) Now by Theorem 3.2.7, $\mathcal{I}(\mathcal{W}'') = \mathcal{I}(\mathcal{W}')$, which by hypothesis is $\mathcal{I}(\mathcal{W})$. Also, if we denote the subspaces of $\mathcal{W}''$ by W''_i, then $v \in W''_j$ and $w \in W''_{k+1}$ satisfy $B(v, w) = 1$ and $U''_i = U \cap W''_i$. Hence by Lemma 3.4.1, U is a symplectic space of dimension k, and both

$$\mathcal{U} = \{U_0 \supset U_1 \supset \cdots \supset U_j \supset U_{j+2} \supset \cdots \supset U_{k+1}\}$$

and

$$\mathcal{U}'' = \{U_0'' \supset U_1'' \supset \cdots \supset U_j'' \supset U_{j+2}'' \supset \cdots \supset U_{k+1}''\}$$

are full flags of U, and further, their incidence matrices may be computed from $\mathcal{I}(\mathcal{W})$ and $\mathcal{I}(\mathcal{W}'')$ by the formula given in that lemma. Thus since $\mathcal{I}(\mathcal{W}) = \mathcal{I}(\mathcal{W}'')$, $\mathcal{I}(\mathcal{U}) = \mathcal{I}(\mathcal{U}'')$, and so by our induction hypothesis, there exists $h \in \mathrm{Sp}(U)$ such that $h\mathcal{U}'' = \mathcal{U}$. Since $V = U \oplus \langle v, w\rangle$, we can extend h to an element of $\mathrm{GL}(V)$ by noting that any element $x \in V$ can be written uniquely in the form $x = u+y$, $u \in U$, $y \in \langle v, w\rangle$, and defining $h(x) = h(u)+y$. In fact, it is easily checked that under this definition, $h \in \mathrm{Sp}(V)$. Hence since $hU_i'' = U_i$, but h restricts to the identity map on $\langle v, w\rangle$, $h\mathcal{W}'' = \mathcal{W}$, so that if we define $g = hg_0 \in \mathrm{Sp}(V)$, $g\mathcal{W}' = \mathcal{W}$. Hence $\mathrm{Sp}(V)$ acts transitively on full flags of V with a given incidence matrix, and so, by induction, the result holds for symplectic spaces of any dimension.

□

We wish to extend Theorem 3.4.2 to all flags of V. The basic idea is that if λ is a composition of n, and two λ-flags of V have the same incidence matrix, then we can insert subspaces into the flags to obtain two full flags with the same incidence matrix, and then apply Theorem 3.4.2. In order to justify this argument, however, we need to be sure that there is a way of extending a λ-flag with a given incidence matrix to a full flag which is independent of the choice of flag having that incidence matrix. That such a canonical extension method exists is the content of our next lemma.

3.4.3 Lemma. Let $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_{k+1}, \ldots, \lambda_{m+1}\}$ be a composition of n such that $\lambda_{k+1} \geq 2$, but $\lambda_i = 1$ for $i > k+1$ (possibly $k = m$). Let

$$\mathcal{W} = \{V = W_0 \supset W_1 \supset \cdots \supset W_k \supset W_{k+1} \supset \cdots \supset W_{m+1} = 0\}$$

be a λ-flag of V. Then there exists $w \in W_k \backslash W_{k+1}$ such that if $\mathcal{W}'$ is the λ'-flag

$$\{W_0 \supset W_1 \supset \cdots \supset W_k \supset \langle w\rangle \oplus W_{k+1} \supset W_{k+1} \supset \cdots \supset W_{m+1}\},$$

where $\lambda' = \{\lambda_1, \lambda_2, \ldots, \lambda_{k+1}-1, 1, \ldots, \lambda_{m+1}\}$, then $\mathcal{I}(\mathcal{W}')$ is given by a formula which depends only on $\mathcal{I}(\mathcal{W})$ (and λ), not $\mathcal{W}$.

Proof. Observe that $W_0^\perp \cap W_k = W_0^\perp \cap W_{k+1} = 0$, whereas $W_{m+1}^\perp \cap W_k = W_k \neq W_{k+1} = W_{m+1}^\perp \cap W_{k+1}$. Hence there must exist a smallest j with $1 \leq j \leq m+1$ such that $W_j^\perp \cap W_k \neq W_j^\perp \cap W_{k+1}$. Since $W_i^\perp \cap W_k \supset W_i^\perp \cap W_{k+1}$ for any i, $\mathcal{I}(\mathcal{W})_{i,k} \geq \mathcal{I}(\mathcal{W})_{i,k+1}$, with equality if and only if $W_i^\perp \cap W_k = W_i^\perp \cap W_{k+1}$, and so j is the smallest row index for which the entries in columns k and $k+1$ of $\mathcal{I}(\mathcal{W})$ are distinct. (If columns k and $k+1$ of $\mathcal{I}(\mathcal{W})$ are identical, then $j = m+1$.) Thus j can be determined from $\mathcal{I}(\mathcal{W})$ alone. Now select any $w \in \left(W_j^\perp \cap W_k\right) \setminus \left(W_j^\perp \cap W_{k+1}\right)$, and consider the λ'-flag

$$\mathcal{W}' = \{W_0 \supset W_1 \supset \cdots \supset W_k \supset \langle w\rangle \oplus W_{k+1} \supset W_{k+1} \supset \cdots \supset W_{m+1}\}.$$

We need to show that $\mathcal{I}(\mathcal{W}')$ can be determined entirely from $\mathcal{I}(\mathcal{W})$. Since $\mathcal{W}'$ is constructed by adding one additional subspace to $\mathcal{W}$, any entry of $\mathcal{I}(\mathcal{W}')$ not in row or column $k+1$ corresponds naturally to an entry of $\mathcal{I}(\mathcal{W})$. Further, the entries in row $k+1$ can be determined from those in column $k+1$ using the partial flag analogue of Property 1 of Proposition 3.3.1. (Proposition 3.3.1 applies only to *full* flags, but since any partial flag is a subflag of some full flag, each property in Proposition 3.3.1 yields information about partial flags as well. Property 1 can be applied to a partial flag by determining in which entry (i, j) a particular number appears in the incidence matrix of a full flag having the partial flag as a subflag.) Thus we only need to show how to determine the numbers

$$\dim\left(W_i^{\perp} \cap (\langle w\rangle \oplus W_{k+1})\right)$$

and

$$\dim\left((\langle w\rangle \oplus W_{k+1})^{\perp} \cap (\langle w\rangle \oplus W_{k+1})\right)$$

from $\mathcal{I}(\mathcal{W})$. Now if $i \geq j$, $W_i^{\perp} \supset W_j^{\perp}$, so $w \in W_i^{\perp}$. Hence

$$W_i^{\perp} \cap (\langle w\rangle \oplus W_{k+1}) = \langle w\rangle \oplus \left(W_i^{\perp} \cap W_{k+1}\right),$$

so that

$$\dim\left(W_i^{\perp} \cap (\langle w\rangle \oplus W_{k+1})\right) = \dim\left(W_i^{\perp} \cap W_{k+1}\right) + 1.$$

If $i < j$, then by our choice of j, $W_i^{\perp} \cap W_{k+1} = W_i^{\perp} \cap W_k$. But

$$W_i^{\perp} \cap W_{k+1} \subset W_i^{\perp} \cap (\langle w\rangle \oplus W_{k+1}) \subset W_i^{\perp} \cap W_k.$$

Hence $W_i^{\perp} \cap (\langle w\rangle \oplus W_{k+1}) = W_i^{\perp} \cap W_{k+1}$, and so

$$\dim\left(W_i^{\perp} \cap (\langle w\rangle \oplus W_{k+1})\right) = \dim\left(W_i^{\perp} \cap W_{k+1}\right).$$

If $j \leq k+1$, then $W_j^{\perp} \subset W_{k+1}^{\perp}$, and so $w \in W_{k+1}^{\perp}$, and as always, $w \in \langle w\rangle^{\perp}$. Hence

$$(\langle w\rangle \oplus W_{k+1})^{\perp} \cap (\langle w\rangle \oplus W_{k+1}) = \left(\langle w\rangle^{\perp} \cap W_{k+1}^{\perp}\right) \cap (\langle w\rangle \oplus W_{k+1})$$

$$= \langle w\rangle \oplus \left(\langle w\rangle^{\perp} \cap W_{k+1}^{\perp} \cap W_{k+1}\right) = \langle w\rangle \oplus \left(W_{k+1}^{\perp} \cap W_{k+1}\right),$$

(since $W_{k+1} \subset \langle w\rangle^{\perp}$), which shows that

$$\dim\left((\langle w\rangle \oplus W_{k+1})^{\perp} \cap (\langle w\rangle \oplus W_{k+1})\right) = \dim\left(W_{k+1}^{\perp} \cap W_{k+1}\right) + 1.$$

If $j > k+1$, then by our choice of j, $W_{k+1}^{\perp} \cap W_{k+1} = W_{k+1}^{\perp} \cap W_k$. Hence

$$(\langle w\rangle \oplus W_{k+1})^{\perp} \cap (\langle w\rangle \oplus W_{k+1}) \subset W_{k+1}^{\perp} \cap W_k = W_{k+1}^{\perp} \cap W_{k+1},$$

and so since by Property 5 of Proposition 3.3.1, the difference between

$$\dim\left((\langle w\rangle \oplus W_{k+1})^{\perp} \cap (\langle w\rangle \oplus W_{k+1})\right)$$

and $\dim\left(W_{k+1}^{\perp}\cap W_{k+1}\right)$ is ± 1, we must have

$$\dim\Big((\langle w\rangle \oplus W_{k+1})^{\perp}\cap(\langle w\rangle\oplus W_{k+1})\Big) = \dim\left(W_{k+1}^{\perp}\cap W_{k+1}\right) - 1.$$

Since k, j, and $\dim\left(W_i^{\perp}\cap W_{k+1}\right)$ can be determined from $\mathcal{I}(\mathcal{W})$ and λ, we are thus able to determine $\mathcal{I}(\mathcal{W}')$ from $\mathcal{I}(\mathcal{W})$ and λ, and so $\mathcal{I}(\mathcal{W}')$ is independent of the λ-flag $\mathcal{W}$ having incidence matrix $\mathcal{I}(\mathcal{W})$.

□

We can now complete the classification of the orbits of the action of $\mathrm{Sp}(V)$ on F_λ.

3.4.4 Theorem (The Orbit Classification Theorem)**.** Let λ be a composition of n, and $\mathcal{W},\mathcal{U}\in F_\lambda$. Then there exists $g\in\mathrm{Sp}(V)$ such that $g\mathcal{W}=\mathcal{U}$ if and only if $\mathcal{I}(\mathcal{W})=\mathcal{I}(\mathcal{U})$.

Proof. The first direction is Theorem 3.2.7, so conversely, suppose that $\mathcal{I}(\mathcal{W}) = \mathcal{I}(\mathcal{U})$. If λ is not the composition $\{1,1,\ldots,1\}$, then by Lemma 3.4.3 we can construct λ' flags (where λ' is a composition of n in which 1 occurs more times than in λ) $\mathcal{W}'$ and $\mathcal{U}'$ having the same incidence matrix, and having $\mathcal{W}$ and $\mathcal{U}$ as subflags. Thus by repeatedly applying Lemma 3.4.3, we will eventually obtain full flags $\mathcal{W}''$ and $\mathcal{U}''$ having $\mathcal{W}$ and $\mathcal{U}$ as subflags and having the same incidence matrix. By Theorem 3.4.2, there exists $g\in\mathrm{Sp}(V)$ such that $g\mathcal{W}''=\mathcal{U}''$, and so certainly for the corresponding subflags we have $g\mathcal{W}=\mathcal{U}$. □

In light of this theorem, any orbit of a flag action of $\mathrm{Sp}(V)$ is completely determined by its type λ and its incidence matrix $\mathcal{I}$, and therefore we will denote it by $\mathcal{O}(\mathcal{I}_\lambda)$, or simply $\mathcal{O}(\mathcal{I})$ when λ is understood.

3.5 Orbit Counting and Listing

Now that we understand the nature of the orbits of the flag actions of $\mathrm{Sp}(V)$, we can begin to answer various questions about them. In this section we will show how to list all the orbits of these actions for a given dimension of V, and prove a general formula for the number of orbits on the set of full flags. We begin by noting that listing the orbits can be reduced to listing the orbits on the set of full flags. For any λ-flag is a subflag of some full flag, and so by Proposition 3.2.3 every incidence matrix of a partial flag is equal to $D_\lambda(\mathcal{I})$ for some full flag incidence matrix $\mathcal{I}$. Consequently, if we have listed all incidence matrices of full flags, we can obtain all incidence matrices of partial flags by appropriate deleting of rows and columns. Since full flag incidence matrices are completely determined by the properties of Proposition 3.3.1, the number of orbits of the flag actions of $\mathrm{Sp}(V)$ of each type and the incidence matrices of those orbits depend only on $n=\dim V$, and in particular are independent of the field $\mathbb{F}$.

3.5.1 Example. In Example 3.3.2 we listed all incidence matrices of full flags for $n = 4$. Thus by applying D_λ to these matrices for all compositions λ of 4, we conclude that the incidence matrices of the orbits of $\mathrm{Sp}(V)$ on the set of all flags of various types are as listed below.

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 1 & 1 \end{pmatrix}_{\{1,1,1,1\}} \quad \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix}_{\{1,1,1,1\}} \quad \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 3 & 2 & 1 \end{pmatrix}_{\{1,1,1,1\}}$$

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}_{\{1,1,2\}} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}_{\{1,1,2\}} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}_{\{1,2,1\}} \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix}_{\{1,2,1\}}$$

$$\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}_{\{2,1,1\}} \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}_{\{2,1,1\}}$$

$$(1)_{\{3,1\}} \qquad (2)_{\{2,2\}} \qquad (0)_{\{2,2\}} \qquad (1)_{\{1,3\}} \qquad \emptyset_{\{4\}}$$

($\emptyset_{\{4\}}$ corresponds to the orbit whose single element is the trivial flag $\{V \supset 0\}$.)

Our next theorem counts the number of full flag orbits of V in any dimension, and is proved in a manner that is sufficiently constructive to demonstrate how to list the corresponding incidence matrices.

3.5.2 Theorem. If V is a symplectic space of dimension $n = 2m$, then the number of orbits of the action of $\mathrm{Sp}(V)$ on the set of full flags of V is

$$\prod_{k=1}^{m} (2k-1).$$

Proof. By Theorems 3.4.4 and 3.3.4, we can identify the orbits of the action of $\mathrm{Sp}(V)$ on the set of full flags of V with the set $\mathfrak{I}_n$ of $(n-1)\times(n-1)$ matrices satisfying the properties of Proposition 3.3.1. To simplify notation, we adopt in this proof the convention that if $\mathcal{I} \in \mathfrak{I}_n$, then for any i with $0 \le i \le n$,

$$\mathcal{I}_{0,i} = 0, \qquad \mathcal{I}_{n,i} = n - i, \qquad \mathcal{I}_{i,0} = i, \qquad \text{and} \qquad \mathcal{I}_{i,n} = 0.$$

Thus if $\mathcal{W}$ is any full flag of V, $\mathcal{I}(\mathcal{W})_{k,l} = \dim\left(W_k^\perp \cap W_l\right)$, even when k or l is 0 or n.

We now proceed by induction on m. If $m = 1$, then, as noted before, there is only one orbit and only one incidence matrix, namely (1). Since $\prod_{k=1}^{1}(2k-1) = 1$, the result holds when $m = 1$. Now suppose that for some $m \ge 1$,

$$|\mathfrak{I}_{2m}| = \prod_{k=1}^{m}(2k-1).$$

We need to show that

$$\left|\mathfrak{I}_{2(m+1)}\right| = \prod_{k=1}^{m+1}(2k-1),$$

which by our induction hypothesis amounts to showing that

$$\left|\mathfrak{I}_{2(m+1)}\right| = (2m+1)\left|\mathfrak{I}_{2m}\right|.$$

Thus it is natural to construct a bijection between $\mathfrak{I}_{2(m+1)}$ and $\{0,1,\ldots,2m\}\times\mathfrak{I}_{2m}$. To define one direction of this bijection, we can use Lemma 3.4.1. That is, for $\mathcal{I}\in\mathfrak{I}_{2(m+1)}$, we define $\Phi(\mathcal{I})=(j,\mathcal{I}')$, where j is the smallest index such that the last entry of row $j+1$ of $\mathcal{I}$ is 1, and

$$\mathcal{I}'_{k,l} = \begin{cases} \mathcal{I}_{k,l} & \text{if } k,l\le j \\ \mathcal{I}_{k,l+1} & \text{if } k\le j \text{ and } l\ge j+1 \\ \mathcal{I}_{k+1,l}-1 & \text{if } k\ge j+1 \text{ and } l\le j \\ \mathcal{I}_{k+1,l+1}-1 & \text{if } k,l\ge j+1 \end{cases}.$$

By Lemma 3.4.1, $\mathcal{I}'\in\mathfrak{I}_{2m}$ whenever $\mathcal{I}\in\mathfrak{I}_{2(m+1)}$.

Conversely, for $(j,\mathcal{J})\in\{0,1,\ldots,2m\}\times\mathfrak{I}_{2m}$, define $\Psi(j,\mathcal{J})\in\mathfrak{I}_{2(m+1)}$ by

$$\Psi(j,\mathcal{J})_{k,l} = \begin{cases} \mathcal{J}_{k,l} & \text{if } 1\le k,l\le j \\ \mathcal{J}_{k,l-1} & \text{if } 1\le k\le j \text{ and } j+1\le l\le n+1 \\ \mathcal{J}_{k-1,l}+1 & \text{if } j+1\le k\le n+1 \text{ and } 1\le l\le j \\ \mathcal{J}_{k-1,l-1}+1 & \text{if } j+1\le k,l\le n+1 \end{cases}.$$

(Here we are using the notational convention mentioned at the beginning of this proof.) We must show that $\Psi(j,\mathcal{J})\in\mathfrak{I}_{2(m+1)}$. To do this, we show that it is the incidence matrix of a flag which results from a construction dual to that used in Lemma 3.4.1. So let $n=2m$ and suppose V is the symplectic space of dimension $n+2$. Fix $v,w\in V$ with $B(v,w)=1$, and let $U=\langle v,w\rangle^{\perp}$. Then $V=U\oplus\langle v,w\rangle$, and U is isomorphic to the symplectic space of dimension n, so $\mathfrak{I}_n$ is identified with the set of orbits of the action of $\mathrm{Sp}(U)$ on the set of full flags of U. So now choose a full flag

$$\mathcal{U}=\{U_0\supset U_1\supset\cdots\supset U_n\}$$

of U such that $\mathcal{I}(\mathcal{U})=\mathcal{J}$. Let $\mathcal{W}$ be the full flag of V given by

$$\mathcal{W}=\{U_0\oplus\langle v,w\rangle\supset\cdots\supset U_j\oplus\langle v,w\rangle\supset U_j\oplus\langle w\rangle\supset\cdots\supset U_n\oplus\langle w\rangle\supset 0\}.$$

We claim that $\mathcal{I}(\mathcal{W})=\Psi(j,\mathcal{J})$. To see this, first note that if A is either $\langle v,w\rangle$ or $\langle w\rangle$, then by reasoning identical to that used in the proof of Lemma 3.4.1, we have that for any k,l with $0\le k,l\le n$,

$$(U_k\oplus\langle v,w\rangle)^{\perp}\cap(U_l\oplus A)=\left(U_k^{\perp}\cap U\right)\cap U_l,$$

while

$$(U_k\oplus\langle w\rangle)^{\perp}\cap(U_l\oplus A)=\left(U_k^{\perp}\cap U\right)\cap U_l\oplus\langle w\rangle.$$

Now by definition $\mathcal{I}(\mathcal{W})_{k,l}$ equals

$$\begin{cases} \dim\Big((U_k \oplus \langle v,w\rangle)^{\perp} \cap (U_l \oplus \langle v,w\rangle)\Big) & \text{if } 1 \le k,l \le j \\ \dim\Big((U_k \oplus \langle v,w\rangle)^{\perp} \cap (U_{l-1} \oplus \langle w\rangle)\Big) & \text{if } 1 \le k \le j \text{ and } j+1 \le l \le n+1 \\ \dim\Big((U_{k-1} \oplus \langle w\rangle)^{\perp} \cap (U_l \oplus \langle v,w\rangle)\Big) & \text{if } j+1 \le k \le n+1 \text{ and } 1 \le l \le j \\ \dim\Big((U_{k-1} \oplus \langle w\rangle)^{\perp} \cap (U_{l-1} \oplus \langle w\rangle)\Big) & \text{if } j+1 \le k,l \le n+1 \end{cases},$$

which by the above-mentioned subspace equalities equals

$$\begin{cases} \dim\big(\big(U_k^{\perp} \cap U\big) \cap U_l\big) & \text{if } 1 \le k,l \le j \\ \dim\big(\big(U_k^{\perp} \cap U\big) \cap U_{l-1}\big) & \text{if } 1 \le k \le j \text{ and } j+1 \le l \le n+1 \\ \dim\big(\big(U_{k-1}^{\perp} \cap U\big) \cap U_l\big)+1 & \text{if } j+1 \le k \le n+1 \text{ and } 1 \le l \le j \\ \dim\big(\big(U_{k-1}^{\perp} \cap U\big) \cap U_{l-1}\big)+1 & \text{if } j+1 \le k,l \le n+1 \end{cases},$$

which is, by definition, $\Psi(j,\mathcal{I}(\mathcal{U}))_{k,l} = \Psi(j,\mathcal{J})_{k,l}$. Thus Ψ is, in fact, a map

$$\{0,1,\ldots,n\} \times \mathfrak{I}_n \to \mathfrak{I}_{n+2}.$$

We claim that $\Psi = \Phi^{-1}$. To show this, suppose first that $(j,\mathcal{J}) \in \{0,1,\ldots,n\} \times \mathfrak{I}_n$, and let $(j',\mathcal{J}') = \Phi(\Psi(j,\mathcal{J}))$. To show that $j' = j$, recall that j' is the smallest index for which the last entry of row $j'+1$ of $\Psi(j,\mathcal{J})$ is 1. Since $\Psi(j,\mathcal{J})_{j,n+1} = \mathcal{J}_{j,n} = 0$, whereas $\Psi(j,\mathcal{J})_{j+1,n+1} = \mathcal{J}_{j,n}+1 = 1$, this index is j. Hence using j in the definition of $\Phi(\Psi(j,\mathcal{J}))$, we can compute $\mathcal{J}'$. If $1 \le k,l \le j$, then $\mathcal{J}'_{k,l} = \Psi(j,\mathcal{J})_{k,l} = \mathcal{J}_{k,l}$. If $k \le j$ and $l \ge j+1$, then $\mathcal{J}'_{k,l} = \Psi(j,\mathcal{J})_{k,l+1} = \mathcal{J}_{k,l+1-1} = \mathcal{J}_{k,l}$. If $k \ge j+1$ and $l \le j$, then $\mathcal{J}'_{k,l} = \Psi(j,\mathcal{J})_{k+1,l} - 1 = (\mathcal{J}_{k+1-1,l}+1) - 1 = \mathcal{J}_{k,l}$. Finally, if $k,l \ge j+1$, then $\mathcal{J}'_{k,l} = \Psi(j,\mathcal{J})_{k+1,l+1} - 1 = \mathcal{J}_{k,l}$. Thus in all cases, $\mathcal{J}'_{k,l} = \mathcal{J}_{k,l}$, and so $\Phi(\Psi(j,\mathcal{J})) = (j,\mathcal{J})$.

Now suppose that $\mathcal{I} \in \mathfrak{I}_{n+2}$, and consider $\Psi(\Phi(\mathcal{I}))$. If we let $(j,\mathcal{I}') = \Phi(\mathcal{I})$, then it is automatic that j is the index which appears in the formulae for $\mathcal{I}'$ and $\Psi(j,\mathcal{I}')$. By definition,

$$\Psi(\Phi(I))_{k,l} = \begin{cases} \mathcal{I}'_{k,l} & \text{if } 1 \le k,l \le j \\ \mathcal{I}'_{k,l-1} & \text{if } 1 \le k \le j \text{ and } j+1 \le l \le n+1 \\ \mathcal{I}'_{k-1,l}+1 & \text{if } j+1 \le k \le n+1 \text{ and } 1 \le l \le j \\ \mathcal{I}'_{k-1,l-1}+1 & \text{if } j+1 \le k,l \le n+1 \end{cases}.$$

We must show that in all four cases the entry in question is $\mathcal{I}_{k,l}$. If neither k nor l is $j+1$ this follows directly from the definition of $\mathcal{I}'$ in the same manner as in the previous case. By Property 1 of Proposition 3.3.1, if $\mathcal{I}$ and $\Psi(\Phi(\mathcal{I}))$ agree in column $j+1$, they will also agree in row $j+1$, so we only need to consider the case in which $l = j+1$. By definition

$$\Psi(\Phi(\mathcal{I}))_{k,j+1} = \begin{cases} \mathcal{I}'_{k,j} = \mathcal{I}_{k,j} & \text{if } k \le j \\ \mathcal{I}'_{j,j}+1 = \mathcal{I}_{j,j}+1 & \text{if } k = j+1 \\ \mathcal{I}'_{k-1,j}+1 = \mathcal{I}_{k,j} & \text{if } k \ge j+2 \end{cases}.$$

It is not immediately obvious that this formula always yields $\mathcal{I}_{k,j+1}$. However, by the definition of j, $\mathcal{I}_{j,n+1} = 0$, while $\mathcal{I}_{j+1,n+1} = 1$. Thus by Property 1 of Proposition 3.3.1, $\mathcal{I}_{n+1,j} = \mathcal{I}_{n+1,j+1} = n+1-j$, and so by Property 6, $\mathcal{I}_{k,j} = \mathcal{I}_{k,j+1}$ for all $k \leq n+1$. This proves $\Psi(\Phi(\mathcal{I}))_{k,j+1} = \mathcal{I}_{k,j+1}$ for $k \neq j+1$, and since applying Property 1 to the equation $\mathcal{I}_{j,j} = \mathcal{I}_{j,j+1}$ shows that $\mathcal{I}_{j,j} = \mathcal{I}_{j+1,j} - 1$, we can conclude that

$$\Psi(\Phi(\mathcal{I}))_{j+1,j+1} = \mathcal{I}_{j,j} + 1 = \mathcal{I}_{j+1,j} = \mathcal{I}_{j+1,j+1}.$$

Thus $\Psi(\Phi(\mathcal{I})) = \mathcal{I}$, and we conclude that $\Psi = \Phi^{-1}$. Hence Φ is a bijection, and so

$$\left|\mathfrak{I}_{2(m+1)}\right| = |\{0,1,\ldots,2m\} \times \mathfrak{I}_{2m}| = (2m+1)\,|\mathfrak{I}_{2m}| = \prod_{k=1}^{m+1}(2k-1)$$

as desired. □

We note that the orbit formula can be written in terms of factorials, using the identity

$$\prod_{k=1}^{m}(2k-1) = \frac{(2m)!}{2^m m!}.$$

The utility of the proof of Theorem 3.5.2 compensates fully for its inelegance, for the map

$$\Psi : \{0,1,\ldots,n\} \times \mathfrak{I}_n \to \mathfrak{I}_{n+2}$$

provides a straightforward recursive algorithm for computing $\mathfrak{I}_n$ for any even dimension n. For we know that $\mathfrak{I}_2 = \{(1)\}$, and if we already know $\mathfrak{I}_n$, we can easily compute

$$\mathfrak{I}_{n+2} = \Psi\left(\{0,1,\ldots,n\} \times \mathfrak{I}_n\right)$$

using the formula provided for Ψ.

3.5.3 Example. Using the already computed set $\mathfrak{I}_4$ from Example 3.3.2, we can compute $\Psi\left(\{0,1,2,3,4\} \times \mathfrak{I}_4\right)$ to find that $\mathfrak{I}_6$ consists of the 15 incidence matrices listed in Table 3.2. If $(j,\mathcal{J}) \in \{0,1,2,3,4\} \times \mathfrak{I}_4$, we list $\Psi(j,\mathcal{J})$ in the row labelled by j and column labelled by $\mathcal{J}$.

	$\begin{pmatrix} 1&0&0 \\ 1&0&0 \\ 2&1&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&0 \\ 2&2&1 \\ 2&2&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1 \\ 2&2&1 \\ 3&2&1 \end{pmatrix}$
0	$\begin{pmatrix} 1&1&1&1&1 \\ 2&2&1&1&1 \\ 3&2&1&1&1 \\ 4&3&2&2&1 \\ 5&4&3&2&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1&1&1 \\ 2&2&2&1&1 \\ 3&3&3&2&1 \\ 4&3&3&2&1 \\ 5&4&3&2&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1&1&1 \\ 2&2&2&2&1 \\ 3&3&3&2&1 \\ 4&4&3&2&1 \\ 5&4&3&2&1 \end{pmatrix}$
1	$\begin{pmatrix} 1&1&0&0&0 \\ 2&2&1&1&1 \\ 2&2&1&1&1 \\ 3&3&2&2&1 \\ 4&4&3&2&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1&0&0 \\ 2&2&2&1&1 \\ 3&3&3&2&1 \\ 3&3&3&2&1 \\ 4&4&3&2&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1&1&0 \\ 2&2&2&2&1 \\ 3&3&3&2&1 \\ 4&4&3&2&1 \\ 4&4&3&2&1 \end{pmatrix}$
2	$\begin{pmatrix} 1&0&0&0&0 \\ 1&0&0&0&0 \\ 2&1&1&1&1 \\ 3&2&2&2&1 \\ 4&3&3&2&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1&0&0 \\ 2&2&2&1&0 \\ 3&3&3&2&1 \\ 3&3&3&2&1 \\ 4&3&3&2&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1&1&0 \\ 2&2&2&1&0 \\ 3&3&3&2&1 \\ 4&3&3&2&1 \\ 4&3&3&2&1 \end{pmatrix}$
3	$\begin{pmatrix} 1&0&0&0&0 \\ 1&0&0&0&0 \\ 2&1&1&1&0 \\ 3&2&2&2&1 \\ 4&3&2&2&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&0&0&0 \\ 2&2&1&1&0 \\ 2&2&1&1&0 \\ 3&3&2&2&1 \\ 4&3&2&2&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1&1&0 \\ 2&2&1&1&0 \\ 3&2&1&1&0 \\ 4&3&2&2&1 \\ 4&3&2&2&1 \end{pmatrix}$
4	$\begin{pmatrix} 1&0&0&0&0 \\ 1&0&0&0&0 \\ 2&1&1&0&0 \\ 3&2&1&0&0 \\ 4&3&2&1&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&0&0&0 \\ 2&2&1&0&0 \\ 2&2&1&0&0 \\ 3&2&1&0&0 \\ 4&3&2&1&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1&0&0 \\ 2&2&1&0&0 \\ 3&2&1&0&0 \\ 3&2&1&0&0 \\ 4&3&2&1&1 \end{pmatrix}$

Table 3.2: Incidence Matrices of Full Flags for $\dim V = 6$.

Having shown how to list the orbits of the flag actions of $\mathrm{Sp}(V)$, we now begin a consideration of the corresponding representations which arise when the field $\mathbb{F}$ is finite.

3.6 Flag Representations of $\mathrm{Sp}(n, q)$

Throughout this section, let $\mathbb{F} = \mathbb{F}_q$ be the field of q elements, and let V be the symplectic space of even dimension $n = 2m$ over $\mathbb{F}_q$. We denote the symplectic group of V by $\mathrm{Sp}(n, q)$. Corresponding to each orbit $\mathcal{O}(\mathcal{I}_\lambda)$ of the action of $\mathrm{Sp}(n, q)$ on the set of flags of V we get a permutation representation of $\mathrm{Sp}(n, q)$ which we denote by $\Phi(\mathcal{I}_\lambda)$, or simply $\Phi(\mathcal{I})$ when the composition λ is understood. In this section we will show how to compute the degrees of these representations and discuss certain isomorphisms which occur among them.

As is the case with any permutation representation, $\deg \Phi(\mathcal{I}_\lambda) = |\mathcal{O}(\mathcal{I}_\lambda)|$. Thus to compute the degrees of these representations, we need a technique for counting the number of flags in a given orbit. We will first determine how to count the number of full flags in a given orbit, and then extend the method to partial flags. First, we need some notation. If $\beta = \{v_1, v_2, \ldots, v_n\}$ is a basis of V, then let $\mathcal{W}(\beta)$ be the full flag of V given by

$$\{V \supset \langle v_1, \ldots, v_{n-1} \rangle \supset \cdots \supset \langle v_1, v_2 \rangle \supset \langle v_1 \rangle \supset 0\}.$$

Since it is easier to count the number of bases β of V such that $\mathcal{W}(\beta)$ is a full flag in a given orbit than to count the number of flags in the orbit directly, we will have need of the following lemma.

3.6.1 Lemma. Let $\mathcal{W}$ be a full flag of V. Then the number of bases β of V such that $\mathcal{W}(\beta) = \mathcal{W}$ is

$$\prod_{i=1}^{n} \left(q^i - q^{i-1}\right).$$

Proof. Let $\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_n\}$. We need to determine the number of bases $\beta = \{v_1, v_2, \ldots, v_n\}$ of V such that $W_{n-i} = \langle v_1, \ldots, v_i\rangle$. In choosing such a basis, we must choose v_1 so that $W_{n-1} = \langle v_1\rangle$. Since $\dim W_{n-1} = 1$, we can choose any non-zero vector in W_{n-1} for v_1, and hence there are $|W_{n-1}| - 1 = q - 1$ choices for v_1. In general, if we have already chosen $v_1, \ldots, v_k$ so that $W_{n-i} = \langle v_1, \ldots, v_i\rangle$ for $i \leq k$, then we will have $W_{n-(k+1)} = \langle v_1, \ldots, v_{k+1}\rangle$ if and only if $v_{k+1} \in W_{n-(k+1)} \backslash W_{n-k}$, so there are $\left|W_{n-(k+1)}\right| - |W_{n-k}| = q^{k+1} - q^k$ choices for v_{k+1}. Hence multiplying these quantities, we conclude that the total number of choices for β is

$$(q-1)(q^2-q)\ldots(q^n - q^{n-1}) = \prod_{i=1}^{n} \left(q^i - q^{i-1}\right).$$

□

We can now give a method for counting the number of full flags in a given orbit. Recall from Lemma 3.3.3 that if $\mathcal{I}$ is the incidence matrix of a full flag, then for some k the above-diagonal entries of any two consecutive rows of $\mathcal{I}$ take either the form

$$\begin{matrix} k{+}1 & k & \ldots & k & k{-}1 & \ldots & k{-}1 & \ldots & 1 & \ldots & 1 & 0 & \ldots & 0 \\ & k & \ldots & k & k{-}1 & \ldots & k{-}1 & \ldots & 1 & \ldots & 1 & 0 & \ldots & 0 \end{matrix} \qquad \text{(Form 1)}$$

or the form

$$\begin{matrix} k{-}1 & k{-}1 & \ldots & k{-}1 & k{-}2 & \ldots & k{-}2 & \ldots & l{-}1 & \ldots & l{-}1 & l{-}1 & \ldots & l{-}1 & \ldots & 1 & \ldots & 1 & 0 & \ldots & 0 \\ & k & \ldots & k & k{-}1 & \ldots & k{-}1 & \ldots & l & \ldots & l & l{-}1 & \ldots & l{-}1 & \ldots & 1 & \ldots & 1 & 0 & \ldots & 0 \end{matrix} \qquad \text{(Form 2)}$$

for some l with $1 \leq l \leq k$.

3.6.2 Theorem. Let $\mathcal{I}$ be the incidence matrix of some orbit of the action of $\mathrm{Sp}(n,q)$ on the set of full flags of V. Define a sequence $\{(a_i, b_i)\}_{1\leq i\leq n}$ by $(a_1, b_1) = (n, 0)$, $(a_n, b_n) = (n, n-1)$, and if $1 < i < n$, then

$$(a_i, b_i) = \begin{cases} (n-k, i-1) & \text{if rows } n-i+1 \text{ and } n-i \text{ take Form 1} \\ (n+1-l, n-l) & \text{if rows } n-i+1 \text{ and } n-i \text{ take Form 2} \end{cases},$$

with k and l being the entries of $\mathcal{I}$ corresponding to the form the rows take. Then the number of full flags with incidence matrix $\mathcal{I}$ is

$$\deg \Phi(\mathcal{I}) = |\mathcal{O}(\mathcal{I})| = \frac{\prod_{i=1}^{n} \left(q^{a_i} - q^{b_i}\right)}{\prod_{i=1}^{n} \left(q^i - q^{i-1}\right)}.$$

Proof. We begin by counting the number of bases $\beta = \{v_1, \ldots, v_n\}$ of V such that $\mathcal{I}(\mathcal{W}(\beta)) = \mathcal{I}$. Choosing such a basis roughly amounts to following the procedure used in Theorem 3.3.4. That is, we first choose a non-zero vector v_1, and then assuming we have chosen $v_1, \ldots, v_j$ so that the incidence matrix of $\{\langle v_1, \ldots, v_j\rangle \supset \cdots \supset \langle v_1\rangle \supset 0\}$ is $D_{\{n-j,1,\ldots,1\}}(\mathcal{I})$ (the $j \times j$ submatrix in the bottom right corner of $\mathcal{I}$), we use Lemma 3.3.3 to choose v_{j+1} so that the incidence matrix of the corresponding flag beginning with $\langle v_1, \ldots, v_{j+1}\rangle$ is $D_{\{n-(j+1),1,\ldots,1\}}(\mathcal{I})$. Doing this for $j = 1, 2 \ldots, n-1$ will yield a basis β with $\mathcal{I}(\mathcal{W}(\beta)) = \mathcal{I}$, and any such β can be obtained in this way.

We claim that the number of possible choices for v_j at any step in this procedure is $q^{a_j} - q^{b_j}$. Since v_1 can be any non-zero vector in V, there are $q^n - 1 = q^{a_1} - q^{b_1}$ choices for v_1. If $1 < i < n$ and we have already chosen $v_1, \ldots, v_{i-1}$, then the number of possible choices for v_i depends on the form of rows $n-i+1$ and $n-i$. If the above diagonal entries of these rows are of Form 1 for the number k, then by Lemma 3.3.3 the set of possible choices for v_i is $\left(W + W^{\perp}\right) \backslash W$, where $W = \langle v_1, \ldots, v_{i-1}\rangle$. Hence the number of choices for v_i is $q^{\dim\left(W+W^{\perp}\right)} - q^{\dim W}$. Of course $\dim W = i-1$. $\dim\left(W + W^{\perp}\right) = n - \dim\left(W + W^{\perp}\right)^{\perp} = n - \dim\left(W^{\perp} \cap W\right) = n-k$, since by our choice of $v_1, \ldots, v_{i-1}$, $\dim\left(W^{\perp} \cap W\right) = \mathcal{I}_{n-i+1,n-i+1} = k$. Thus the number of possible choices for v_i is $q^{n-k} - q^{i-1}$. If the above diagonal entries in the rows in question are of Form 2 for the numbers k and l, then by Lemma 3.3.3 the set of possible choices for v_i is $(W + U^{\perp}) \backslash (W + U'^{\perp})$, where again $W = \langle v_1, \ldots, v_{i-1}\rangle$, and if the rightmost appearance of l in row $n-i+1$ is in column $n-j$, then $U' = \langle v_1, \ldots, v_j\rangle$ and $U = \langle v_1, \ldots, v_{j-1}\rangle$ (if $j = 1$ then $U = 0$). Hence the number of possible choices for v_i is $q^{\dim\left(W+U^{\perp}\right)} - q^{\dim\left(W+U'^{\perp}\right)}$. But $\dim\left(W + U^{\perp}\right) = n - \dim\left(W^{\perp} \cap U\right) = n - \mathcal{I}_{n-i+1,n-j+1} = n - (l-1)$, and $\dim\left(W + U'^{\perp}\right) = n - \dim\left(W^{\perp} \cap U'\right) = n - \mathcal{I}_{n-i+1,n-j} = n - l$, so this means the number of possible choices is $q^{n+1-l} - q^{n-l}$. Thus in all cases if $1 < i < n$, the number of possible choices for v_i is $q^{a_i} - q^{b_i}$. Finally, once we have chosen $v_1, \ldots, v_{n-1}$, we can choose v_n to be any vector in $V \backslash \langle v_1, \ldots, v_{n-1}\rangle$, and so there are $q^n - q^{n-1} = q^{a_n} - q^{b_n}$ choices for v_n. Thus for each j, the number of possible choices for v_j is $q^{a_j} - q^{b_j}$.

Hence multiplying the number of possible choices at each step, we conclude that the number of bases β with $\mathcal{I}(\mathcal{W}(\beta)) = \mathcal{I}$ is

$$\prod_{i=1}^{n} \left(q^{a_i} - q^{b_i}\right).$$

By Lemma 3.6.1, each flag $\mathcal{U}$ with incidence matrix $\mathcal{I}$ has $\prod_{i=1}^{n}\left(q^i - q^{i-1}\right)$ different bases β with $\mathcal{W}(\beta) = \mathcal{U}$, so the total number of flags with incidence matrix $\mathcal{I}$ is

$$\frac{\prod_{i=1}^{n}\left(q^{a_i} - q^{b_i}\right)}{\prod_{i=1}^{n}\left(q^i - q^{i-1}\right)}.$$

□

We now wish to extend this method to partial flags. The basic idea is that any λ-flag of V can be extended to a full flag of V in a certain number of ways (depending only on λ), and so we can count the number of λ-flags with a given incidence matrix by counting the number of full flags which have such a flag as a subflag and dividing by the number of ways of extending a λ-flag to a full flag. Thus we need the following lemma.

3.6.3 Lemma. Let $\lambda = \{\lambda_1, \ldots, \lambda_k\}$ be a composition of n, and let $\mathcal{W}$ be a λ-flag of V. Set $\mu_i = \lambda_{k+1-i}$. Let $S_0 = 0$ and for $1 \leq i \leq k$ let

$$S_i = \sum_{j=1}^{i} \mu_j.$$

Then the number of full flags of V with $\mathcal{W}$ as a subflag is

$$E(\lambda) = \frac{\prod_{i=1}^{k} \left(\prod_{j=1}^{\mu_i} \left(q^{S_i} - q^{S_{i-1}+j-1} \right) \right)}{\prod_{i=1}^{n} \left(q^i - q^{i-1} \right)}.$$

Proof. Let $\mathcal{W} = \{W_0 \supset W_1 \supset \cdots \supset W_k\}$. We begin by counting the number of bases $\beta = \{v_1, \ldots, v_n\}$ of V such that $\mathcal{W}(\beta)$ has $\mathcal{W}$ as a subflag, which is the number of β such that $W_{k-i} = \langle v_1, \ldots, v_{S_i} \rangle$. (By construction $S_i = \dim W_{k-i}$.) To choose such a β, we must first choose $v_1, \ldots, v_{S_1}$ so that $\langle v_1, \ldots, v_{S_1} \rangle = W_{k-1}$. Hence the total number of choices for $v_1, \ldots, v_{S_1}$ is the number of bases for W_{k-1}, namely

$$\prod_{j=1}^{S_1} \left(q^{S_1} - q^{j-1} \right) = \prod_{j=1}^{\mu_1} \left(q^{S_1} - q^{S_0+j-1} \right).$$

Assuming that for some $i \geq 2$ we have already chosen $v_1, \ldots, v_{S_{i-1}}$ so that $W_{k-j} = \langle v_1, \ldots, v_{S_j} \rangle$ for all $j \leq i-1$, we must choose $v_{S_{i-1}+1}, \ldots, v_{S_i}$ so that $W_{k-i} = \langle v_1, \ldots, v_{S_i} \rangle$. For each $v_{S_{i-1}+j}$ which we choose in this process, the set of possible choices is $W_{k-i} \backslash \langle v_1, \ldots, v_{S_{i-1}+j-1} \rangle$, so the number of possibilities for $v_{S_{i-1}+j}$ is $q^{S_i} - q^{S_{i-1}+j-1}$. Hence multiplying these numbers, we conclude that the total number of possibilities for $v_{S_{i-1}+1}, \ldots, v_{S_i}$ is

$$\prod_{j=1}^{\mu_i} \left(q^{S_i} - q^{S_{i-1}+j-1} \right).$$

Multiplying these numbers from 1 to k shows that the total number of β such that $\mathcal{W}$ is a subflag of $\mathcal{W}(\beta)$ is

$$\prod_{i=1}^{k} \left(\prod_{j=1}^{\mu_i} \left(q^{S_i} - q^{S_{i-1}+j-1} \right) \right),$$

and hence dividing by the number of bases for each full flag, we conclude that the number $E(\lambda)$ of full flags of V with $\mathcal{W}$ as a subflag is

$$\frac{\prod_{i=1}^{k}\left(\prod_{j=1}^{\mu_i}\left(q^{S_i}-q^{S_{i-1}+j-1}\right)\right)}{\prod_{i=1}^{n}\left(q^i-q^{i-1}\right)}.$$

□

The preceding lemma allows us to compute the number of λ-flags with a given incidence matrix for any composition λ. Recall that $\mathfrak{I}_n$ denotes the set of all $(n-1)\times(n-1)$ incidence matrices of full flags, and that for $\mathcal{I}\in\mathfrak{I}_n$, $D_\lambda(\mathcal{I})$ is the (uniquely determined) incidence matrix of any λ-flag which occurs as a subflag of a flag in $\mathcal{O}(\mathcal{I})$. Thus for an incidence matrix $\mathcal{J}_\lambda$ of a λ-flag, $D_\lambda^{-1}(\mathcal{J}_\lambda)$ is the subset of $\mathfrak{I}_n$ consisting of all incidence matrices of full flags whose subflags of type λ have incidence matrix $\mathcal{J}_\lambda$.

3.6.4 Theorem. Let λ be a composition of n, and let $\mathcal{O}(\mathcal{J}_\lambda)$ be the orbit of the action of $\mathrm{Sp}(n,q)$ on F_λ with incidence matrix $\mathcal{J}_\lambda$. Then

$$|\mathcal{O}(\mathcal{J}_\lambda)| = \frac{1}{E(\lambda)}\sum_{\mathcal{I}\in D_\lambda^{-1}(\mathcal{J}_\lambda)}|\mathcal{O}(\mathcal{I})|\,.$$

Proof. A full flag $\mathcal{W}$ of V has a subflag of type λ with incidence matrix $\mathcal{J}_\lambda$ if and only if $\mathcal{W}\in\mathcal{O}(\mathcal{I})$ for some $\mathcal{I}\in D_\lambda^{-1}(\mathcal{J}_\lambda)$. Hence the total number of such full flags is

$$\sum_{\mathcal{I}\in D_\lambda^{-1}(\mathcal{J}_\lambda)}|\mathcal{O}(\mathcal{I})|\,.$$

Since each λ-flag in $\mathcal{O}(\mathcal{J}_\lambda)$ is a subflag of $E(\lambda)$ different full flags represented in this sum, there must be

$$\frac{1}{E(\lambda)}\sum_{\mathcal{I}\in D_\lambda^{-1}(\mathcal{J}_\lambda)}|\mathcal{O}(\mathcal{I})|$$

different λ-flags in $\mathcal{O}(\mathcal{J}_\lambda)$.

□

3.6.5 Example. In Example 3.5.1 we listed all incidence matrices of flags of a 4-dimensional symplectic space. We will now use the results of this section to compute the degrees of the corresponding representations of $\mathrm{Sp}(4,q)$ for any q (that is, the number of flags with each incidence matrix). We begin with the full flag incidence matrices, for which we use Theorem 3.6.2. In the notation of that theorem, if $\mathcal{I}=\left(\begin{smallmatrix}1&0&0\\1&0&0\\2&1&1\end{smallmatrix}\right)$, we have $(a_1,b_1)=(4,0)$ and $(a_4,b_4)=(4,3)$. Rows 3 and 2 take Form 2 with $l=1$, so $(a_2,b_2)=(4,3)$, while rows 2 and

1 take Form 1 with $k = 0$, so $(a_3, b_3) = (4, 2)$. Hence according to Theorem 3.6.2,

$$|\mathcal{O}(\mathcal{I})| = \frac{(q^4-1)(q^4-q^3)(q^4-q^2)(q^4-q^3)}{(q-1)(q^2-q)(q^3-q^2)(q^4-q^3)}$$
$$= (q^3+q^2+q+1)q^2(q+1) = q^6+2q^5+2q^4+2q^3+q^2.$$

The computations for the other two full flag incidence matrices are identical, except that for $\begin{pmatrix} 1&1&0\\2&2&1\\2&2&1 \end{pmatrix}$, we have $(a_2, b_2) = (3, 1)$ and $(a_3, b_3) = (4, 3)$, while for $\begin{pmatrix} 1&1&1\\2&2&1\\3&2&1 \end{pmatrix}$ we have $(a_2, b_2) = (3, 1)$ and $(a_3, b_3) = (3, 2)$.

Table 3.3 gives $|\mathcal{O}(\mathcal{I})|$ for each full flag incidence matrix $\mathcal{I}$.

$\mathcal{I}$	$\lvert\mathcal{O}(\mathcal{I})\rvert = \deg \Phi(\mathcal{I})$
$\begin{pmatrix} 1&0&0\\1&0&0\\2&1&1 \end{pmatrix}$	$q^6+2q^5+2q^4+2q^3+q^2$
$\begin{pmatrix} 1&1&0\\2&2&1\\2&2&1 \end{pmatrix}$	$q^5+2q^4+2q^3+2q^2+q$
$\begin{pmatrix} 1&1&1\\2&2&1\\3&2&1 \end{pmatrix}$	$q^4+2q^3+2q^2+2q+1$

Table 3.3: Full Flag Orbit Sizes for $\mathrm{Sp}(4, q)$.

To compute the degrees of the representations corresponding to the remaining orbits, we need the quantity $E(\lambda)$ given by the formula in Lemma 3.6.3 for each composition λ of 4. The formula is relatively straightforward. For example, if $\lambda = \{2, 1, 1\}$, then we have $\mu_1 = 1, \mu_2 = 1, \mu_3 = 2, S_1 = 1, S_2 = 2$, and $S_3 = 4$, and so

$$E(\{2,1,1\}) = \frac{(q-1)(q^2-q)(q^4-q^2)(q^4-q^3)}{(q-1)(q^2-q)(q^3-q^2)(q^4-q^3)} = q+1.$$

Table 3.4 gives $E(\lambda)$ for each composition λ of 4.

λ	$E(\lambda)$
$\{1,1,1,1\}$	1
$\{1,1,2\}$	$q+1$
$\{1,2,1\}$	$q+1$
$\{2,1,1\}$	$q+1$
$\{1,3\}$	q^3+2q^2+2q+1
$\{2,2\}$	q^2+2q+1
$\{3,1\}$	q^3+2q^2+2q+1
$\{4\}$	$q^6+3q^5+5q^4+6q^3+5q^2+3q+1$

Table 3.4: Values of $E(\lambda)$.

Now we can apply Theorem 3.6.4 to compute $|\mathcal{O}(\mathcal{I}_\lambda)|$ for each remaining orbit $\mathcal{O}(\mathcal{I}_\lambda)$. We will give two examples of how this is done. If $\mathcal{I}_\lambda = \begin{pmatrix} 2&1\\2&1 \end{pmatrix}_{\{2,1,1\}}$, then $D_\lambda^{-1}(\mathcal{I}_\lambda) = \left\{ \left(\begin{smallmatrix} 1&1&0\\ 2&2&1\\ 2&2&1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1&1&1\\ 2&2&1\\ 3&2&1 \end{smallmatrix}\right) \right\}$, since these are the full flag incidence matrices for which $\mathcal{I}_\lambda$ is the submatrix obtained by removing the first row and column. Hence

$$|\mathcal{O}(\mathcal{I}_\lambda)| = \frac{(q^5 + 2q^4 + 2q^3 + 2q^2 + q) + (q^4 + 2q^3 + 2q^2 + 2q + 1)}{q+1}$$
$$= q^4 + 2q^3 + 2q^2 + 2q + 1.$$

If $\mathcal{I}_\lambda = (0)_{\{2,2\}}$, then

$$D_\lambda^{-1}(\mathcal{I}_\lambda) = \left\{ \left(\begin{smallmatrix} 1&0&0\\ 1&0&0\\ 2&1&1 \end{smallmatrix}\right) \right\},$$

since this is the only full flag incidence matrix with a 0 in the $(2,2)$ entry. Hence

$$|\mathcal{O}(\mathcal{I}_\lambda)| = \frac{q^6 + 2q^5 + 2q^4 + 2q^3 + q^2}{q^2 + 2q + 1} = q^4 + q^2.$$

Table 3.5 lists $|\mathcal{O}(\mathcal{I}_\lambda)|$ for each partial flag incidence matrix $\mathcal{I}_\lambda$.

$\mathcal{I}_\lambda$	$\lvert\mathcal{O}(\mathcal{I}_\lambda)\rvert = \deg \Phi(\mathcal{I}_\lambda)$
$\left(\begin{smallmatrix} 1&0\\ 1&0 \end{smallmatrix}\right)_{\{1,1,2\}}$	$q^5 + q^4 + q^3 + q^2$
$\left(\begin{smallmatrix} 1&1\\ 2&2 \end{smallmatrix}\right)_{\{1,1,2\}}$	$q^4 + 2q^2 + 2q^2 + 2q + 1$
$\left(\begin{smallmatrix} 1&0\\ 2&1 \end{smallmatrix}\right)_{\{1,2,1\}}$	$q^5 + 2q^4 + 2q^3 + 2q^2 + q$
$\left(\begin{smallmatrix} 1&1\\ 3&1 \end{smallmatrix}\right)_{\{1,2,1\}}$	$q^3 + q^2 + q + 1$
$\left(\begin{smallmatrix} 0&0\\ 1&1 \end{smallmatrix}\right)_{\{2,1,1\}}$	$q^5 + q^4 + q^3 + q^2$
$\left(\begin{smallmatrix} 2&1\\ 2&1 \end{smallmatrix}\right)_{\{2,1,1\}}$	$q^4 + 2q^3 + 2q^2 + 2q + 1$
$(1)_{\{1,3\}}$	$q^3 + q^2 + q + 1$
$(2)_{\{2,2\}}$	$q^3 + q^2 + q + 1$
$(0)_{\{2,2\}}$	$q^4 + q^2$
$(1)_{\{3,1\}}$	$q^3 + q^2 + q + 1$
$\emptyset_{\{4\}}$	1

Table 3.5: Partial Flag Orbit Sizes for $\mathrm{Sp}(4,q)$.

Examining the computed degrees of these representations, we notice that in many cases, two or more different representations may have the same degree. For example, the representations corresponding to $(1)_{\{1,3\}}$, $(2)_{\{2,2\}}$, $(1)_{\{3,1\}}$, and $\left(\begin{smallmatrix} 1&1\\ 3&1 \end{smallmatrix}\right)_{\{1,2,1\}}$ all have degree $q^3 + q^2 + q + 1$. It is natural to ask whether this is mere coincidence, or whether these representations are, in fact, isomorphic. We will see in the next chapter that $\Phi((2)_{\{2,2\}})$ and $\Phi((1)_{\{3,1\}})$ have different characters, and thus are not isomorphic, but based on the next two results, equality of degree does imply isomorphism for all pairs of representations not including $\Phi((2)_{\{2,2\}})$. The first of these isomorphism theorems describes a situation in which one or more subspaces in a flag are completely determined by the remaining subspaces and the incidence matrix of the flag.

3.6.6 Theorem. Let $\mathcal{I}_\lambda$ be the incidence matrix of some λ-flag of V, and suppose μ is a composition of n with $\mu \preceq \lambda$. Let $\mathcal{J}_\mu = D_{\lambda,\mu}(\mathcal{I}_\lambda)$. Then the representations $\Phi(\mathcal{I}_\lambda)$ and $\Phi(\mathcal{J}_\mu)$ of $\mathrm{Sp}(n,q)$ are isomorphic if and only if they have the same degree.

Proof. Obviously two isomorphic representations have the same degree. Conversely, note that in general we can define a linear map $T_{\lambda,\mu} : \mathbb{C}\mathcal{O}(\mathcal{I}_\lambda) \to \mathbb{C}\mathcal{O}(\mathcal{J}_\mu)$ on the basis $\mathcal{O}(\mathcal{I}_\lambda)$ of $\mathbb{C}\mathcal{O}(\mathcal{I}_\lambda)$ by

$$T_{\lambda,\mu}(\mathcal{W}) = \mathcal{W}_\mu.$$

$T_{\lambda,\mu}$ is at least a homomorphism of representations, since if $g \in \mathrm{Sp}(n,q)$, we will get the same result if we apply g to a flag $\mathcal{W}$ before or after removing subspaces of certain dimensions, and thus $T_{\lambda,\mu}(g\mathcal{W}) = gT_{\lambda,\mu}(\mathcal{W})$. Further, $T_{\lambda,\mu}$ is always surjective, for if $\mathcal{U} \in \mathcal{O}(\mathcal{J}_\mu)$ and $\mathcal{W} \in \mathcal{O}(\mathcal{I}_\lambda)$ we must have $T_{\lambda,\mu}(\mathcal{W}) = \mathcal{W}_\mu \in \mathcal{O}(\mathcal{J}_\mu)$. By Theorem 3.4.4, there exists $g \in \mathrm{Sp}(n,q)$ such that $g\mathcal{W}_\mu = \mathcal{U}$, and hence $T_{\lambda,\mu}(g\mathcal{W}) = \mathcal{U}$. This shows that the image of $T_{\lambda,\mu}$ contains a basis of $\mathcal{O}(\mathcal{J}_\mu)$, and so $T_{\lambda,\mu}$ is surjective. Hence if $\deg \Phi(\mathcal{I}_\lambda) = \deg \Phi(\mathcal{J}_\mu)$, this surjective homomorphism of representations must be an isomorphism. □

The second isomorphism theorem describes an isomorphism which corresponds to taking the orthogonal complement of a flag. If A is a $k \times k$ matrix, denote by ${}^tA$ the backward transpose of A (the matrix with (i,j) entry $A_{k+1-j,k+1-i}$), and if λ is a composition of n, denote by $\overline{\lambda}$ the composition of n consisting of λ taken in reverse order.

3.6.7 Theorem. If $\mathcal{W}$ is a λ-flag of V, $\mathcal{W}^\perp$ is a $\overline{\lambda}$-flag of V and $\mathcal{I}(\mathcal{W}^\perp) = {}^t\mathcal{I}(\mathcal{W})$. The mapping $\mathcal{W} \mapsto \mathcal{W}^\perp$ induces an isomorphism between the representations $\Phi(\mathcal{I}_\lambda)$ and $\Phi({}^t\mathcal{I}_{\overline{\lambda}})$ of $\mathrm{Sp}(n,q)$.

Proof. That $\mathcal{W}^\perp$ is a flag of type $\overline{\lambda}$ follows directly from the fact that for any subspaces $U \subset W$ of V, we have $\dim(W/U) = \dim\left(U^\perp/W^\perp\right)$. To see that $\mathcal{I}(\mathcal{W}^\perp) = {}^t\mathcal{I}(\mathcal{W})$, note that if $\lambda = \{\lambda_1, \ldots, \lambda_k\}$, the ith subspace in the flag $\mathcal{W}^\perp$ is $W_{k-i}^\perp$. Hence by definition $\mathcal{I}(\mathcal{W}^\perp)_{i,j} = \dim\left(\left(W_{k-i}^\perp\right)^\perp \cap W_{k-j}^\perp\right) = \dim\left(W_{k-j}^\perp \cap W_{k-i}\right) = \mathcal{I}(\mathcal{W})_{k-j,k-i}$, which is the (i,j) entry of ${}^t\mathcal{I}(\mathcal{W})$, since $\mathcal{I}(\mathcal{W})$ is a $(k-1) \times (k-1)$ matrix.

Since $\left(\mathcal{W}^\perp\right)^\perp = \mathcal{W}$ for any flag $\mathcal{W}$, by an extension of Proposition 2.2.6, the mapping $\mathcal{W} \mapsto \mathcal{W}^\perp$ defines a bijection between $\mathcal{O}(\mathcal{I}_\lambda)$ and $\mathcal{O}({}^t\mathcal{I}_{\overline{\lambda}})$, and so induces a vector space isomorphism between $\mathbb{C}\mathcal{O}(\mathcal{I}_\lambda)$ and $\mathbb{C}\mathcal{O}({}^t\mathcal{I}_{\overline{\lambda}})$. If $g \in \mathrm{Sp}(n,q)$, we have $(g\mathcal{W})^\perp = g(\mathcal{W}^\perp)$ by an extension of Proposition 2.2.14, which implies that this vector space isomorphism is, in fact, an isomorphism of representations. □

We can now see several isomorphisms among the representations of Example 3.6.5. For example, if $\mathcal{I}_\lambda = \left(\begin{smallmatrix} 1 & 1 \\ 2 & 2 \end{smallmatrix}\right)_{\{1,1,2\}}$, ${}^t\mathcal{I}_{\overline{\lambda}} = \left(\begin{smallmatrix} 2 & 1 \\ 2 & 1 \end{smallmatrix}\right)_{\{2,1,1\}}$, so the two corresponding representations are isomorphic by Theorem 3.6.7. If $\mathcal{I}_\lambda = \left(\begin{smallmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{smallmatrix}\right)$, then $D_{\{1,2,1\}}(\mathcal{I}_\lambda) = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)_{\{1,2,1\}}$, so since the corresponding representations have the same degree, they are isomorphic by Theorem 3.6.6. Table 3.6 partitions the incidence matrices of flags of V for $n = 4$ into classes which yield isomorphic representations.

Incidence Matrices	Common Degree of Representations
$\emptyset_{\{4\}}$	1
$(1)_{\{1,3\}} (1)_{\{3,1\}} \left(\begin{smallmatrix} 1 & 1 \\ 3 & 1 \end{smallmatrix}\right)_{\{1,2,1\}}$	$q^3 + q^2 + q + 1$
$(2)_{\{2,2\}}$	$q^3 + q^2 + q + 1$
$(0)_{\{2,2\}}$	$q^4 + q^2$
$\left(\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right)_{\{1,1,2\}} \left(\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right)_{\{2,1,1\}}$	$q^5 + q^4 + q^3 + q^2$
$\left(\begin{smallmatrix} 1 & 1 \\ 2 & 2 \end{smallmatrix}\right)_{\{1,1,2\}} \left(\begin{smallmatrix} 2 & 1 \\ 2 & 1 \end{smallmatrix}\right)_{\{2,1,1\}} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 3 & 2 & 1 \end{pmatrix}_{\{1,1,1,1\}}$	$q^4 + 2q^3 + 2q^2 + 2q + 1$
$\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)_{\{1,2,1\}} \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix}_{\{1,1,1,1\}}$	$q^5 + 2q^4 + 2q^3 + 2q^2 + q$
$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 1 & 1 \end{pmatrix}_{\{1,1,1,1\}}$	$q^6 + 2q^5 + 2q^4 + 2q^3 + q^2$

Table 3.6: Isomorphism Classes of Flag Representations of $\mathrm{Sp}(4, q)$.

From Table 3.6, we see that there are actually only 8 representations of $\mathrm{Sp}(4, q)$ corresponding to the various flag actions. In the next chapter, we will compute the characters of these representations.

Chapter 4

Conjugacy Classes and Flag Characters of $\mathrm{Sp}(4, q)$

Our goal in this chapter is to provide a complete list of the conjugacy classes of $\mathrm{Sp}(4, q)$, together with their sizes, and to compute the values of the previously discussed flag characters on these classes. We will also show how the inner product defined in Section 2.1 can be used to compute an irreducible character of $\mathrm{Sp}(4, q)$ from the flag characters. We restrict our attention to the case in which q is odd, since when q is even, the list of conjugacy classes takes a slightly different (though simpler) form. First, we review some background information.

4.1 Canonical Forms and Centralizers

In this section, we discuss some general results and computational techniques which will be needed for producing a list of class representatives of $\mathrm{Sp}(4, q)$ and computing the sizes of the classes. Throughout this section let $\mathbb{F}$ be a field, and let $\overline{\mathbb{F}}$ be the algebraic closure of $\mathbb{F}$. For convenience, we also assume that $\mathbb{F}$ is a perfect field, so that any irreducible polynomial of degree n in $\mathbb{F}[x]$ has n distinct roots in $\overline{\mathbb{F}}$. Every finite field is perfect, as is every field of characteristic zero [3]. We begin by discussing canonical forms of matrices over $\mathbb{F}$ and $\overline{\mathbb{F}}$. The Jordan canonical form is well-known. See, for example, [5] page 145.

4.1.1 Proposition. Any matrix $A \in \mathrm{M}(n, \overline{\mathbb{F}})$ is similar to a block diagonal matrix with each block having the form

$$\lambda I_k + J_k = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

for some $\lambda \in \overline{\mathbb{F}}$. (Throughout this section a blank entry will indicate a 0.) Here I_k is the $k \times k$ identity matrix, and J_k is the $k \times k$ matrix with a 1 in each entry immediately above the diagonal, and a 0 in every other entry. Up to reordering of the blocks, A is similar to a unique matrix of this form.

Since we will generally be working over finite fields, which are not algebraically closed, we need the modified version of the Jordan canonical form for $\mathbb{F}$ which is used in [4]. To describe this form, we first fix some notation.

4.1.2 Definition. If $\mu = \{n_1, n_2, \ldots, n_k\}$ is a composition of n, then $\mathrm{B}(\mu, \mathbb{F})$ is the subalgebra of $\mathrm{M}(n, \mathbb{F})$ consisting of all block diagonal matrices of the form

$$\begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{pmatrix},$$

where $B_i \in \mathrm{M}(n_i, \mathbb{F})$. The above element of $\mathrm{B}(\mu, \mathbb{F})$ may be written more compactly as

$$\mathrm{B}_\mu(B_1, B_2, \ldots, B_k),$$

and in this notation the composition μ may be omitted when there is no danger of confusion about the block sizes.

4.1.3 Definition. If $f(x) = x^d + b_1 x^{d-1} + \cdots + b_{d-1}x + b_d$ is a monic polynomial in $\mathbb{F}[x]$, then the *companion matrix* of f is

$$C(f) = C_1(f) = \begin{pmatrix} & & & & -b_d \\ 1 & & & & -b_{d-1} \\ & 1 & & & -b_{d-2} \\ & & \ddots & & \vdots \\ & & & 1 & -b_1 \end{pmatrix}$$

In [5] it is shown that f is the minimal polynomial, and hence also the characteristic polynomial of $C(f)$. We define the block matrix $C_k(f)$ for $k \geq 2$ to be

$$\begin{pmatrix} C(f) & I_d & & \\ & C(f) & \ddots & \\ & & \ddots & I_d \\ & & & C(f) \end{pmatrix},$$

with $C(f)$ appearing k times on the diagonal. If $\mu = \{a_1, a_2, \ldots, a_k\}$ is a partition of some integer m, then we define $C_\mu(f)$ to be the block diagonal matrix

$$\mathrm{B}_\nu(C_{a_1}(f), C_{a_2}(f), \ldots, C_{a_k}(f)),$$

where ν is the partition $\{a_1 d, \ldots, a_k d\}$ of md. Although Green uses in [4] a slightly different matrix in place of $C(f)$, his matrix is similar to ours and the result he uses is still valid for our choice.

4.1.4 Theorem. Let $A \in \mathrm{M}(n, \mathbb{F})$ have characteristic polynomial $\prod_{i=1}^{k} f_i^{m_i}$, with each f_i an irreducible monic polynomial in $\mathbb{F}[x]$. Then there exist unique partitions μ_i of m_i such that A is similar to the block diagonal matrix

$$C = \mathrm{B}_\nu(C_{\mu_1}(f_1), C_{\mu_2}(f_2), \ldots, C_{\mu_k}(f_k)),$$

where ν is the composition $\{n_1, \ldots, n_k\}$ of n, with $n_i = m_i \deg f_i$.

Proof. If f is an irreducible polynomial of degree d which appears with multiplicity m in the characteristic polynomial of A, then if $\lambda_1, \ldots, \lambda_d$ are the distinct roots of f in $\overline{\mathbb{F}}$ there exists some partition $\mu = \{a_1, a_2, \ldots, a_k\}$ of m such that the Jordan canonical form of A includes the blocks

$$\begin{gathered}
\lambda_1 I_{a_1} + J_{a_1}, \ldots, \lambda_d I_{a_1} + J_{a_1}, \\
\lambda_1 I_{a_2} + J_{a_2}, \ldots, \lambda_d I_{a_2} + J_{a_2}, \\
\vdots \\
\lambda_1 I_{a_k} + J_{a_k}, \ldots, \lambda_d I_{a_k} + J_{a_k}.
\end{gathered}$$

(Note in particular that to each block $\lambda_1 I_j + J_j$ there corresponds a block $\lambda_i I_j + J_j$ of the same size for each root λ_i of f.) We claim that the block matrix

$$A_{f,a_j} = \mathrm{B}\left(\lambda_1 I_{a_j} + J_{a_j}, \ldots, \lambda_d I_{a_j} + J_{a_j}\right)$$

is similar to $C_{a_j}(f)$. To see this, note that if we conjugate A_{f,a_j} by the appropriate permutation matrix, we obtain the matrix

$$\begin{pmatrix} D & I_d & & \\ & D & \ddots & \\ & & \ddots & I_d \\ & & & D \end{pmatrix},$$

where D is the $d \times d$ diagonal matrix

$$\mathrm{B}(\lambda_1, \ldots, \lambda_d).$$

Now D is the Jordan canonical form of $C(f)$, the companion matrix of f, since D has the distinct eigenvalues of f, the characteristic polynomial of $C(f)$, along its diagonal. Hence there exists $B \in \mathrm{GL}(d, \overline{\mathbb{F}})$ such that $BDB^{-1} = C(f)$. But then

$$\begin{pmatrix} B & & & \\ & B & & \\ & & \ddots & \\ & & & B \end{pmatrix} \begin{pmatrix} D & I_d & & \\ & D & \ddots & \\ & & \ddots & I_d \\ & & & D \end{pmatrix} \begin{pmatrix} B^{-1} & & & \\ & B^{-1} & & \\ & & \ddots & \\ & & & B^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} BDB^{-1} & BI_dB^{-1} & & \\ & BDB^{-1} & \ddots & \\ & & \ddots & BI_dB^{-1} \\ & & & BDB^{-1} \end{pmatrix} = \begin{pmatrix} C(f) & I_d & & \\ & C(f) & \ddots & \\ & & \ddots & I_d \\ & & & C(f) \end{pmatrix}$$

$$= C_{a_j}(f),$$

so A_{f,a_j} and $C_{a_j}(f)$ are similar over $\overline{\mathbb{F}}$. Applying this result for each $a_j \in \mu$ then shows that the matrix consisting of the blocks corresponding to f in the Jordan form of A is similar to $C_\mu(f)$, and applying this to each irreducible factor of the characteristic polynomial of A shows that the Jordan form of A is similar to the matrix C. By transitivity of similarity, A is also similar to C over $\overline{\mathbb{F}}$. But A and C both have entries in $\mathbb{F}$, so they must also be similar over $\mathbb{F}$ (see, for example, [5], page 143). Uniqueness follows from the uniqueness of the Jordan canonical form up to reordering, and the fact that the assumed factorization of the characteristic polynomial and the requirement that partitions be increasing sequences specifies an ordering of the blocks of C. $\square$

To illustrate Theorem 4.1.4 we now use it to describe a set of representatives of the conjugacy classes of $\mathrm{GL}(4,\mathbb{F})$.

4.1.5 Example. According to Theorem 4.1.4, every matrix in $\mathrm{GL}(4,\mathbb{F})$ is conjugate to a matrix from the following list. Here $\lambda_i \in \mathbb{F}^*$, and $x^2 - \alpha_1 x - \alpha_2, x^2 - \alpha_3 x - \alpha_4, x^3 - \beta_1 x^2 - \beta_2 x - \beta_3$, and $x^4 - \gamma_1 x^3 - \gamma_2 x^2 - \gamma_3 x - \gamma_4$ are irreducible polynomials in $\mathbb{F}[x]$.

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix} \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & \alpha_2 & 0 & 0 \\ 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 0 & \alpha_2 & 0 & 0 \\ 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix} \begin{pmatrix} 0 & \alpha_2 & 0 & 0 \\ 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & 0 & \alpha_4 \\ 0 & 0 & 1 & \alpha_3 \end{pmatrix} \begin{pmatrix} 0 & \alpha_2 & 1 & 0 \\ 1 & \alpha_1 & 0 & 1 \\ 0 & 0 & 0 & \alpha_2 \\ 0 & 0 & 1 & \alpha_1 \end{pmatrix} \begin{pmatrix} 0 & 0 & \beta_3 & 0 \\ 1 & 0 & \beta_2 & 0 \\ 0 & 1 & \beta_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & \gamma_4 \\ 1 & 0 & 0 & \gamma_3 \\ 0 & 1 & 0 & \gamma_2 \\ 0 & 0 & 1 & \gamma_1 \end{pmatrix}$$

We note also that the completeness of the list of representatives of the conjugacy classes of $\mathrm{GL}(3,2)$ given in Example 2.1.9 follows from Theorem 4.1.4 and the fact the the only irreducible polynomials of degree 2 or 3 over $\mathbb{F}_2$ are $x^2 + x + 1, x^3 + x + 1$, and $x^3 + x^2 + 1$.

Now suppose we are given a matrix A in the canonical form of Theorem 4.1.4. We will find it useful to have an explicit description of the centralizer of A, the subalgebra $\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(A)$ of $\mathrm{M}(n,\mathbb{F})$ consisting of all matrices which commute with A. In constructing this description, we will need some terminology.

4.1.6 Definition. Let $A \in \mathrm{M}(n,\mathbb{F})$. Then A is said to be *semisimple* if A is similar over $\overline{\mathbb{F}}$ to a diagonal matrix. A is *nilpotent* if $A^n = 0$.

Equivalently, A is semisimple if its minimal polynomial has no repeated roots, and A is nilpotent if it has characteristic polynomial x^n. A result found on page 95 of [7] can be adapted to our use as follows:

4.1.7 Proposition. Let $A \in \mathrm{M}(n, \overline{\mathbb{F}})$. Then there exist unique $A_S, A_N \in \mathrm{M}(n, \overline{\mathbb{F}})$ such that $A = A_S + A_N$, A_S is semisimple, A_N is nilpotent, and $A_S A_N = A_N A_S$. Further,

$$\mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(A) = \mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(A_S) \cap \mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(A_N).$$

The following proposition is easily checked.

4.1.8 Proposition. If $\mathrm{B}_\mu(A_1, \dots, A_k)$ is a block diagonal matrix, then

$$\mathrm{B}_\mu(A_1, \dots, A_k)_S = \mathrm{B}_\mu((A_1)_S, \dots, (A_k)_S)$$

and

$$\mathrm{B}_\mu(A_1, \dots, A_k)_N = \mathrm{B}_\mu((A_1)_N, \dots, (A_k)_N).$$

Now if C is a matrix in the canonical form of Theorem 4.1.4, it is easy to decompose C into semisimple and nilpotent parts. For recall that C is block diagonal, with each block having the form

$$C_l(f) = \begin{pmatrix} C(f) & I_d & & \\ & C(f) & \ddots & \\ & & \ddots & I_d \\ & & & C(f) \end{pmatrix}.$$

But $C_l(f) = C_l(f)_S + C_l(f)_N$, where

$$C_l(f)_S = \begin{pmatrix} C(f) & & & \\ & C(f) & & \\ & & \ddots & \\ & & & C(f) \end{pmatrix} \text{ and } C_l(f)_N = \begin{pmatrix} & I_d & & \\ & & \ddots & \\ & & & I_d \\ & & & \end{pmatrix}.$$

Each block $C(f)$ of $C_l(f)_S$ has characteristic polynomial f, which is irreducible over $\mathbb{F}$, and thus has distinct roots in $\overline{\mathbb{F}}$. Hence each $C(f)$ is semisimple, and so $C_l(f)_S$ is semisimple. All entries on or below the diagonal of $C_l(f)_N$ are 0, so $C_l(f)_N$ has characteristic polynomial x^n, and thus is nilpotent. It is easily checked that $C_l(f)_S C_l(f)_N = C_l(f)_N C_l(f)_S$ Hence by Proposition 4.1.8, C_S is block diagonal, with each block of the form $C_l(f)_S$ (which is itself block diagonal, with blocks $C(f)$), while C_N is block diagonal, with each block of the form $C_l(f)_N$. Note in particular that if $C \in \mathrm{M}(n, \mathbb{F})$, then C_S and C_N are also in $\mathrm{M}(n, \mathbb{F})$. Since any matrix $A \in \mathrm{M}(n, \mathbb{F})$ is similar over $\mathbb{F}$ to some such matrix C, the semisimple and nilpotent parts of A are also similar to those of C over $\mathbb{F}$, and thus are in $\mathrm{M}(n, \mathbb{F})$. Hence Proposition 4.1.7 holds with $\overline{\mathbb{F}}$ replaced by $\mathbb{F}$. With this decomposition in mind, we can prove the following lemma, which reduces the computation of the centralizer of C to the computation of the centralizer of each $C_\mu(f)$.

4.1.9 Lemma. If $f_1, \ldots, f_k$ are distinct irreducible polynomials, and $\mu_1, \ldots, \mu_k$ are partitions of $m_1, \ldots, m_k \in \mathbb{N}$, and if $n_i = m_i \deg f_i$ and $\nu = \{n_1, \ldots, n_k\}$ is the corresponding composition of some $n \in \mathbb{N}$, then the centralizer of

$$C = \mathrm{B}_\nu(C_{\mu_1}(f_1), \ldots, C_{\mu_k}(f_k))$$

is

$$\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C) = \mathrm{B}_\nu(\mathcal{C}_{\mathrm{M}(n_1,\mathbb{F})}(C_{\mu_1}(f_1)), \ldots, \mathcal{C}_{\mathrm{M}(n_k,\mathbb{F})}(C_{\mu_k}(f_k))),$$

the subalgebra of $\mathrm{B}(\nu, \mathbb{F})$ consisting of all block diagonal matrices with block i an element of the subalgebra $\mathcal{C}_{\mathrm{M}(n_i,\mathbb{F})}(C_{\mu_i}(f_i))$ of $\mathrm{M}(n_i, \mathbb{F})$.

Proof. A matrix $\mathrm{B}_\nu(A_1, \ldots, A_k) \in \mathrm{B}(\nu, \mathbb{F})$ commutes with C if and only if each A_i commutes with $C_{\mu_i}(f_i)$. Hence it suffices to show that $\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C) \subseteq \mathrm{B}(\nu, \mathbb{F})$. By Proposition 4.1.7, $\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C) \subseteq \mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_S)$, which in turn is contained in $\mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(C_S)$. Now

$$C_S = \mathrm{B}_\nu(C_{\mu_1}(f_1)_S, \ldots, C_{\mu_k}(f_k)_S),$$

where $C_{\mu_i}(f_i)_S = \mathrm{B}(C(f_i), \ldots, C(f_i))$ has m_i copies of $C(f_i)$ along the diagonal. Hence for each i there exists a matrix $X_i \in \mathrm{GL}(n_i, \overline{\mathbb{F}})$ such that $X_i C_{\mu_i}(f_i)_S X_i^{-1} = D_i$, where D_i is a diagonal matrix having roots of f_i along the diagonal. Thus if we set

$$X = \mathrm{B}_\nu(X_1, \ldots, X_k),$$

we have

$$X C_S X^{-1} = D,$$

where

$$D = \mathrm{B}_\nu(D_1, \ldots, D_k).$$

Conjugation by the invertible matrix X being an automorphism of $\mathrm{M}(n, \overline{\mathbb{F}})$, we see that $\mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(C_S) = X^{-1}\mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(D)X$. Now suppose $A \in \mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(D)$ has (x, y) entry $A_{x,y}$, and that D has diagonal entries $d_1, \ldots, d_n$. Then $DA = AD$ implies $(DA)_{x,y} = (AD)_{x,y}$, and so

$$d_x A_{x,y} = A_{x,y} d_y.$$

But if (x, y) is an index outside the blocks of C_S, d_x and d_y are roots of distinct irreducible polynomials, and so $d_x \neq d_y$, which implies that $A_{x,y} = 0$. Hence $A \in \mathrm{B}(\nu, \overline{\mathbb{F}})$, and so $\mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(D) \subseteq \mathrm{B}(\nu, \overline{\mathbb{F}})$. Thus since also $X \in \mathrm{B}(\nu, \overline{\mathbb{F}})$, we must have

$$X^{-1}\mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(D)X = \mathcal{C}_{\mathrm{M}(n,\overline{\mathbb{F}})}(C_S) \subseteq \mathrm{B}(\nu, \overline{\mathbb{F}}),$$

and so

$$\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C) \subseteq \mathrm{B}(\nu, \mathbb{F})$$

as desired. □

By the preceding lemma, we need only concern ourselves with determining $\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f))$, which, by Proposition 4.1.7, amounts to determining $\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f)_S)$ and $\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f)_N)$. We first consider $\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f)_S)$.

4.1.10 Lemma. Let f be an irreducible polynomial of degree d, μ a partition of $m \in \mathbb{N}$, and $n = md$. Then if $\mathcal{A} = \mathcal{C}_{\mathrm{M}(d,\mathbb{F})}(C(f))$ we have

$$\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f)_S) = \begin{pmatrix} \mathcal{A} & \dots & \mathcal{A} \\ \vdots & & \vdots \\ \mathcal{A} & \dots & \mathcal{A} \end{pmatrix},$$

the algebra of $n \times n$ matrices whose m^2 $d \times d$ blocks are all elements of $\mathcal{A}$.

Proof. Independent of the partition μ, we have

$$C_\mu(f)_S = \mathrm{B}(C(f), \dots, C(f)),$$

with m occurrences of $C(f)$ along the diagonal. Now suppose $A \in \mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f)_S)$, and split A into m^2 $d \times d$ blocks $A_{i,j}$, $1 \leq i, j \leq m$, so that

$$A = \begin{pmatrix} A_{1,1} & \dots & A_{1,m} \\ \vdots & & \vdots \\ A_{m,1} & \dots & A_{m,m} \end{pmatrix}.$$

Then $C_\mu(f)_S A$

$$= \begin{pmatrix} C(f) & & \\ & \ddots & \\ & & C(f) \end{pmatrix} \begin{pmatrix} A_{1,1} & \dots & A_{1,m} \\ \vdots & & \vdots \\ A_{m,1} & \dots & A_{m,m} \end{pmatrix}$$

$$= \begin{pmatrix} C(f)A_{1,1} & \dots & C(f)A_{1,m} \\ \vdots & & \vdots \\ C(f)A_{m,1} & \dots & C(f)A_{m,m} \end{pmatrix},$$

whereas

$$AC_\mu(f)_S = \begin{pmatrix} A_{1,1}C(f) & \dots & A_{1,m}C(f) \\ \vdots & & \vdots \\ A_{m,1}C(f) & \dots & A_{m,m}C(f) \end{pmatrix}.$$

Equating corresponding blocks of $C_\mu(f)_S A$ and $AC_\mu(f)_S$ yields

$$C(f)A_{i,j} = A_{i,j}C(f),$$

which implies that $A \in \mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f)_S)$ if and only if each block $A_{i,j}$ is contained in $\mathcal{A}$. □

In light of this lemma, we should next determine $\mathcal{C}_{\mathrm{M}(d,\mathbb{F})}(C(f))$.

4.1.11 Theorem. If $f \in \mathbb{F}[x]$ is an irreducible polynomial of degree d, then $\mathcal{C}_{\mathrm{M}(d,\mathbb{F})}(C(f)) = \langle C(f) \rangle$, the subalgebra of $\mathrm{M}(d,\mathbb{F})$ generated by $C(f)$.

Proof. Let λ be a root in $\overline{\mathbb{F}}$ of the irreducible polynomial f. Then the extension field $\mathbb{F}[\lambda]$ is a vector space over $\mathbb{F}$ of dimension d, with basis $\beta = \{1, \lambda, \lambda^2, \ldots, \lambda^{d-1}\}$. Multiplication by λ is an $\mathbb{F}$-linear map $\mathbb{F}[\lambda] \to \mathbb{F}[\lambda]$, and its matrix with respect to β is $C(f)$. Since the $\mathbb{F}$-algebra generated by λ is the field $\mathbb{F}[\lambda]$, $\langle C(f) \rangle$ and $\mathbb{F}[\lambda]$ are isomorphic as $\mathbb{F}$-algebras. For an element $x \in \mathbb{F}[\lambda]$, let M_x be the matrix in $\langle C(f) \rangle$ representing multiplication by x. Then for any $y \in \mathbb{F}[\lambda]$, if we view elements of $\mathbb{F}[\lambda]$ as column vectors with respect to the basis β, we have $M_x y = xy = yx = M_y x$. Now suppose we are given a matrix $A \in \mathcal{C}_{\mathrm{M}(d,\mathbb{F})}(C(f))$. Viewed as a matrix with respect to the basis β, A determines an $\mathbb{F}$-linear map $\mathbb{F}[\lambda] \to \mathbb{F}[\lambda]$. In particular, there is some element $a = A1 \in \mathbb{F}[\lambda]$. But A commutes with $C(f)$, and hence also with any element of the subalgebra generated by $C(f)$. Hence for any $x \in \mathbb{F}[\lambda]$ we have

$$Ax = AM_x 1 = M_x A1 = M_x a = M_a x,$$

which implies that $A = M_a \in \langle C(f) \rangle$, since any matrix in $\mathrm{M}(d,\mathbb{F})$ is completely determined by its action on $\mathbb{F}[\lambda]$. This proves that $\mathcal{C}_{\mathrm{M}(d,\mathbb{F})}(C(f)) \subseteq \langle C(f) \rangle$, and it is automatic that $\langle C(f) \rangle \subseteq \mathcal{C}_{\mathrm{M}(d,\mathbb{F})}(C(f))$, so we conclude $\mathcal{C}_{\mathrm{M}(d,\mathbb{F})}(C(f)) = \langle C(f) \rangle$. □

In practice, one should interpret Theorem 4.1.11 as allowing us to identify $C(f)$ with a root λ of f, its centralizer $\mathcal{A}$ with $\mathbb{F}[\lambda]$, and $md \times md$ matrices built from $d \times d$ blocks $A \in \mathcal{A}$ with elements of $\mathrm{M}(m, \mathbb{F}[\lambda])$. Under this identification one sees, for example, that Lemma 4.1.10 states that the centralizer in $\mathrm{M}(n,\mathbb{F})$ of the matrix corresponding to $\lambda I_m \in \mathrm{M}(m, \mathbb{F}[\lambda])$ is precisely the subalgebra of $\mathrm{M}(n,\mathbb{F})$ corresponding to $\mathrm{M}(m, \mathbb{F}[\lambda])$.

It now remains to determine $\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f)_N)$. So suppose $\deg f = d$ and that $\mu = \{a_1, \ldots, a_k\}$ is a partition such that $\nu = \{da_1, \ldots, da_k\}$ is a partition of n. For $l \in \mathbb{N}$, we let $J_{l,d}$ be the nilpotent part of $C_l(f)$. That is,

$$J_{l,d} = \begin{pmatrix} & I_d & & \\ & & \ddots & \\ & & & I_d \\ & & & \end{pmatrix}.$$

Then

$$C_\mu(f)_N = \mathrm{B}_\nu(J_{a_1,d}, \ldots, J_{a_k,d}).$$

Now suppose $A \in \mathrm{M}(n,\mathbb{F})$, and divide A into k^2 blocks $A_{i,j}$, with $1 \le i, j \le k$, and the block $A_{i,j}$ a $da_i \times da_j$ matrix. Equating corresponding blocks of $AC_\mu(f)_N$ and $C_\mu(f)_N A$ shows that $A \in \mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f)_N)$ if and only if

$$A_{i,j} J_{a_j,d} = J_{a_i,d} A_{i,j}$$

for all i, j. Now divide $A_{i,j}$ into $a_i a_j$ $d \times d$ blocks $B_{x,y}$, so that

$$A_{i,j} = \begin{pmatrix} B_{1,1} & B_{1,2} & \dots & B_{1,a_j} \\ B_{2,1} & B_{2,2} & \dots & B_{2,a_j} \\ \vdots & \vdots & & \vdots \\ B_{a_i,1} & B_{a_i,2} & \dots & B_{a_i,a_j} \end{pmatrix}.$$

Then

$$A_{i,j} J_{a_j,d} = \begin{pmatrix} 0 & B_{1,1} & \dots & B_{1,a_j-1} \\ 0 & B_{2,1} & \dots & B_{2,a_j-1} \\ \vdots & \vdots & & \vdots \\ 0 & B_{a_i,1} & \dots & B_{a_i,a_j-1} \end{pmatrix},$$

whereas

$$J_{a_i,d} A_{i,j} = \begin{pmatrix} B_{2,1} & B_{2,2} & \dots & B_{2,a_j} \\ \vdots & \vdots & & \vdots \\ B_{a_i,1} & B_{a_i,2} & \dots & B_{a_i,a_j} \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Equating corresponding blocks shows that $A_{i,j} J_{a_j,d} = J_{a_i,d} A_{i,j}$ if and only if for all $1 \leq x \leq a_i - 1$ and $2 \leq y \leq a_j$, we have $B_{x,y-1} = B_{x+1,y}$ and also $B_{x,y} = 0$ if $y = 1$ and $x \neq 1$ or $x = a_i$ and $y \neq a_j$. When applied repeatedly, the condition $B_{x,y-1} = B_{x+1,y}$ shows that the blocks of $A_{i,j}$ must be constant along each diagonal, while the second condition asserts that all but the first block in the first column is zero, and all but the last block in the last column is zero. Of course, if one block along a diagonal is zero, equality along diagonals forces the entire diagonal to be zero. Hence the only diagonals of $A_{i,j}$ which are non-zero are those which extend from the top row of $A_{i,j}$ to the last column. The number of these diagonal is the smaller of a_i and a_j. In summary, we have proven the following proposition.

4.1.12 Proposition. $\mathcal{C}_{\mathrm{M}(n,\mathbb{F})}(C_\mu(f)_N)$ consists of all block matrices with block sizes determined by ν and such that each block $A_{i,j}$ has the form

$$\begin{pmatrix} 0 & \dots & 0 & B_1 & B_2 & \dots & B_{a_i} \\ 0 & & & 0 & B_1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & B_2 \\ 0 & \dots & \dots & \dots & \dots & 0 & B_1 \end{pmatrix}$$

if $a_i \leq a_j$ and the form

$$\begin{pmatrix} B_1 & B_2 & \dots & B_{a_j} \\ 0 & B_1 & \ddots & \vdots \\ \vdots & 0 & \ddots & B_2 \\ \vdots & & \ddots & B_1 \\ \vdots & & & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

if $a_i \geq a_j$ for some $B_l \in \mathrm{M}(d, \mathbb{F})$.

Taken together, the results of this section show how to compute the centralizer of any element C of $\mathrm{M}(n, \mathbb{F})$ in canonical form. By Lemma 4.1.9 we need only compute the centralizers of the blocks $C_\mu(f)$ of C corresponding to the different irreducible factors f of the characteristic polynomial of C. For such an f, we first compute the centralizer of $C(f)$, which by Theorem 4.1.11 is the linear span of $\{I, C(f), C(f)^2, \dots, C(f)^{\deg f - 1}\}$. Then we split an arbitrary matrix of the appropriate size into blocks corresponding to the partition μ, and construct each block by placing arbitrary elements of the centralizer of $C(f)$ along its diagonals according to the rule of Proposition 4.1.12. We illustrate this with an example using a matrix from $\mathrm{M}(17, \mathbb{Q})$.

4.1.13 Example. Suppose we wish to find the centralizer over $\mathbb{Q}$ of the matrix

$$C = \mathrm{B}_{\{8,6,3\}}(C_{\{1,3\}}(x^2 - 2x - 2), C_{\{1,2,3\}}(x - 3), C_{\{1\}}(x^3 - 5)).$$

Thus

$$C = \begin{pmatrix}
 & 2 & & & & & & & & & & & & & & & \\
1 & 2 & & & & & & & & & & & & & & & \\
 & & & 2 & 1 & & & & & & & & & & & & \\
 & & 1 & 2 & & 1 & & & & & & & & & & & \\
 & & & & & 2 & 1 & & & & & & & & & & \\
 & & & & 1 & 2 & & 1 & & & & & & & & & \\
 & & & & & & & 2 & & & & & & & & & \\
 & & & & & & 1 & 2 & & & & & & & & & \\
 & & & & & & & & 3 & & & & & & & & \\
 & & & & & & & & & 3 & 1 & & & & & & \\
 & & & & & & & & & & 3 & & & & & & \\
 & & & & & & & & & & & 3 & 1 & & & & \\
 & & & & & & & & & & & & 3 & 1 & & & \\
 & & & & & & & & & & & & & 3 & & & \\
 & & & & & & & & & & & & & & & & 5 \\
 & & & & & & & & & & & & & & 1 & & \\
 & & & & & & & & & & & & & & & 1 &
\end{pmatrix}.$$

The centralizer of $C(x^2 - 2x - 2)$ is the $\mathbb{Q}$-span of I and $C(x^2 - 2x - 2)$, namely

$$\left\{ \left(\begin{smallmatrix} a & 2b \\ b & a+2b \end{smallmatrix} \right) \middle| \, a, b \in \mathbb{Q} \right\},$$

while the centralizer of $C(x^3 - 5)$ is the $\mathbb{Q}$-span of $I, C(x^3 - 5)$, and $C(x^3 - 5)^2$, namely

$$\left\{ \left(\begin{smallmatrix} a & 5c & 5b \\ b & a & 5c \\ c & b & a \end{smallmatrix} \right) \middle| \, a, b, c \in \mathbb{Q} \right\}.$$

Of course the centralizer of 3 is simply $\mathbb{Q}$. Hence putting numerous copies of these elements into a 17×17 matrix according to the rules specified in the preceding results shows that $\mathcal{C}_{\mathrm{M}(17,\mathbb{Q})}(C)$ consists of all matrices of the form

$$
\left(\begin{array}{cccccccccccccccccc}
a & 2b & & & & & c & 2d & & & & & & & & & \\
b & a+2b & & & & & d & c+2d & & & & & & & & & \\
e & 2f & g & 2h & i & 2j & k & 2l & & & & & & & & & \\
f & e+2f & h & g+2h & j & i+2j & l & k+2l & & & & & & & & & \\
& & & & g & 2h & i & 2j & & & & & & & & & \\
& & & & h & g+2h & j & i+2j & & & & & & & & & \\
& & & & & & g & 2h & & & & & & & & & \\
& & & & & & h & g+2h & & & & & & & & & \\
& & & & & & & & m & & n & & & o & & & \\
& & & & & & & & p & q & r & & s & t & & & \\
& & & & & & & & & & q & & & s & & & \\
& & & & & & & & t & u & v & w & x & y & & & \\
& & & & & & & & & & u & & w & x & & & \\
& & & & & & & & & & & & & w & & & \\
& & & & & & & & & & & & & & z & 5\beta & 5\alpha \\
& & & & & & & & & & & & & & \alpha & z & 5\beta \\
& & & & & & & & & & & & & & \beta & \alpha & z
\end{array}\right),
$$

as the letters used range over $\mathbb{Q}$.

4.2 Conjugacy Classes

The conjugacy classes of $\mathrm{Sp}(4,q)$ were determined by L. E. Dickson in [2] using methods similar to ours, so our description of the associated computations will only be representative of the complexities involved, not complete. First, though, we give a brief description of the form the list will take. Since our list of conjugacy classes is to be valid for any q, the representatives of the conjugacy classes must be given in terms of parameters from $\mathbb{F}_q$. There will be many different types of representatives (about 20, depending on the notation chosen), and for each type we will determine under what circumstances different values of the parameters give rise to the same classes and how many different classes of the given type there are. In general, the latter number will be a polynomial in q, as will the common size of the classes of a given type. A second issue which arises is the particular copy of $\mathrm{Sp}(4,q)$ within $\mathrm{GL}(4,q)$ which is used for the computations. For if J is any invertible matrix satisfying $J^T = -J$, we know that $\mathrm{Sp}(4,q)$ is isomorphic to the matrix group

$$\mathrm{Sp}(4,q)_J = \{A \in \mathrm{GL}(4,q) | A^T J A = J\},$$

and, in general, different matrices J yield distinct (though isomorphic) matrix groups. For the final list, we will take the version of $\mathrm{Sp}(4,q)$ for which J is the matrix $\widehat{B} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$, but since many conjugacy classes have much nicer representatives for more complicated choices of J, we will occasionally use other matrices for J in our computations. The method we use in the computations begins with Example 4.1.5: we know that any matrix in $\mathrm{Sp}(4,q)$ can, for the appropriate choice of J, be written in one of the canonical forms listed below.

$$
\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}
\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}
\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}
\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}
\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}
$$

$$
\begin{pmatrix} 0 & \alpha_2 & 0 & 0 \\ 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}
\begin{pmatrix} 0 & \alpha_2 & 0 & 0 \\ 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}
\begin{pmatrix} 0 & \alpha_2 & 0 & 0 \\ 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & 0 & \alpha_4 \\ 0 & 0 & 1 & \alpha_3 \end{pmatrix}
\begin{pmatrix} 0 & \alpha_2 & 1 & 0 \\ 1 & \alpha_1 & 0 & 1 \\ 0 & 0 & 0 & \alpha_2 \\ 0 & 0 & 1 & \alpha_1 \end{pmatrix}
\begin{pmatrix} 0 & 0 & \beta_3 & 0 \\ 1 & 0 & \beta_2 & 0 \\ 0 & 1 & \beta_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}
\begin{pmatrix} 0 & 0 & 0 & \gamma_4 \\ 1 & 0 & 0 & \gamma_3 \\ 0 & 1 & 0 & \gamma_2 \\ 0 & 0 & 1 & \gamma_1 \end{pmatrix}
$$

Thus for each canonical form A in this list, we determine, if possible, a matrix J and conditions on the parameters in A such that $A \in \mathrm{Sp}(4,q)_J$. We then determine what conditions on the parameters yield distinct representatives of the same conjugacy class and how many different conjugacy classes occur as the parameters range over the allowable values. We may need to distinguish numerous special cases corresponding to equality of certain parameters or values of ± 1 for the parameters. Finally, to compute the size of the conjugacy class of each A, we use the results of Section 4.1 to write down the general form X of an arbitrary element of $\mathcal{C}_{\mathrm{M}(4,q)}(A)$ and then apply the equation $X^T J X = J$ to determine the size of $\mathcal{C}_{\mathrm{Sp}(4,q)}(A) = \mathcal{C}_{\mathrm{M}(4,q)}(A) \cap \mathrm{Sp}(4,q)_J$. Note that we do not need to show explicitly that we have found all classes which correspond to canonical forms of a given type, since at the end of the computations we can compute the total number of elements in the various classes we have found and check that this number is equal to $|\mathrm{Sp}(4,q)|$, which by Proposition 2.2.16 is

$$(q-1)^2 q^4 (q+1)^2 (q^2+1).$$

Before we begin analyzing specific canonical forms, we state one useful fact which allows us to see immediately when many of the above canonical forms are contained in $\mathrm{Sp}(4,q)_{\widehat{B}}$.

4.2.1 Lemma. A matrix of the form $A = \mathrm{B}_{\{2,2\}}(A_1, A_2)$ satisfies $A^T \widehat{B} A = \widehat{B}$ if and only if $A_1, A_2 \in \mathrm{SL}(2,q)$.

Proof. Since A and $\widehat{B}$ are both contained in $\mathrm{B}(\{2,2\},q)$ (the subalgebra of $\mathrm{M}(4,q)$ consisting of all block diagonal matrices with two 2×2 blocks), $A^T \widehat{B} A = \widehat{B}$ if and only if $A_i^T M A_i = M$ for $i = 1,2$, where $M = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, and by Proposition 2.2.12 this is true if and only if $A_i \in \mathrm{SL}(2,q)$. □

A second useful fact concerns matrices which are diagonalizable over $\overline{\mathbb{F}_q}$, the algebraic closure of $\mathbb{F}_q$.

4.2.2 Lemma. Suppose $A \in \mathrm{Sp}(4,q)$ is similar to a diagonal matrix in $\mathrm{GL}(4,\overline{\mathbb{F}_q})$. Then there exist $\lambda, \mu \in \overline{\mathbb{F}_q}$ such that the 4 (not necessarily distinct) eigenvalues in $\overline{\mathbb{F}_q}$ of A are $\lambda, \lambda^{-1}, \mu$, and μ^{-1}.

Proof. Suppose A is conjugate to a diagonal matrix $D \in \mathrm{GL}(4,\overline{\mathbb{F}_q})$ via the matrix P. Since $\mathrm{Sp}(4,q) \leq \mathrm{Sp}(4,\overline{\mathbb{F}_q})$, and conjugation by P maps one copy of $\mathrm{Sp}(4,\overline{\mathbb{F}_q})$ within $\mathrm{GL}(4,\overline{\mathbb{F}_q})$ to another, there must exist $J \in \mathrm{GL}(4,\overline{\mathbb{F}_q})$ such that $J^T = -J$ and $D^T J D = J$. Suppose that $D = \begin{pmatrix} \alpha & & & \\ & \beta & & \\ & & \gamma & \\ & & & \delta \end{pmatrix}$ and that $J = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix}$. Then the equation $D^T J D = J$ implies that $a\alpha\beta = a$, $b\alpha\gamma = b$, $c\alpha\delta = c$, $d\beta\gamma = d$, $e\beta\delta = e$, and $f\gamma\delta = f$. Since J is invertible, at least one of a, b, and c must be non-zero, and so at least one of β, γ, and δ is α^{-1}. Since the matrix must have determinant 1 by Proposition 2.2.13, the two remaining eigenvalues must also be inverses of each other. □

We now turn to analyzing specific canonical forms. First we consider the diagonal matrices of the form

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}.$$

By Lemma 4.2.1, such a matrix is contained in $\mathrm{Sp}(4,q)_{\widehat{B}}$ if and only if there are $x, y \in \mathbb{F}_q^*$ such that $\lambda_1 = x, \lambda_2 = x^{-1}, \lambda_3 = y$, and $\lambda_4 = y^{-1}$. Thus we let

$$a_{x,y} = \begin{pmatrix} x & 0 & 0 & 0 \\ 0 & x^{-1} & 0 & 0 \\ 0 & 0 & y & 0 \\ 0 & 0 & 0 & y^{-1} \end{pmatrix}.$$

We need to determine under what circumstances two elements of this type are conjugate. If $C_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, then $C_1^{-1} a_{x,y} C_1 = a_{x^{-1},y}$, so $cl(a_{x,y}) = cl(a_{x^{-1},y})$. If $C_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$, then $C_2^{-1} a_{x,y} C_2 = a_{y,x}$, and in general,

$$C_2^{-1}\,\mathrm{B}_{\{2,2\}}(A_1, A_2) C_2 = \mathrm{B}_{\{2,2\}}(A_2, A_1).$$

Thus $cl(a_{x,y}) = cl(a_{y,x})$. By combining these two class equalities, we see that for any $x, y \in \mathbb{F}_q^*$ all elements of the forms a_{x^i,y^j} and a_{y^j,x^i} for $i, j \in \{1, -1\}$ represent the same conjugacy class in $\mathrm{Sp}(4,q)_{\widehat{B}}$. However, if $\{x_1, x_1^{-1}, y_1, y_1^{-1}\} \neq \{x_2, x_2^{-1}, y_2, y_2^{-1}\}$, a_{x_1,y_1} and a_{x_2,y_2} have different characteristic polynomials, and so are not conjugate.

We now distinguish several different types of elements of the form $a_{x,y}$. The elements $a_{1,1}$ and $a_{-1,-1}$ together constitute the centre of $\mathrm{Sp}(4,q)$, and so each is the unique representative of a distinct conjugacy class of size 1. If $x \notin \{-1, 0, 1\}$, then since $cl(a_{x,x}) = cl(a_{x^{-1},x^{-1}})$ there are $\frac{q-3}{2}$ distinct classes with a representative of the form $a_{x,x}$. Similarly, there are $q-3$ classes with representatives of the form $a_{x,\pm 1}$ for $x \notin \{-1, 0, 1\}$, and one additional class with representative $a_{1,-1}$. Finally, we have the most general class with representative of the form $a_{x,y}$, namely that for which $|\{x, x^{-1}, y, y^{-1}\}| = 4$. (Thus neither x nor y is ± 1, and x and y are neither equal nor inverses of each other.) In choosing x and y to produce an $a_{x,y}$ of this type, we can choose any element of $\mathbb{F}_q$ besides 0 and ± 1 for x, but y must also not be equal to $x^{\pm 1}$. Thus the number of $a_{x,y}$ of this type is $(q-3)(q-5)$. But by the conjugacy relations noted previously, each $a_{x,y}$ is conjugate to 7 other elements of the same form, and so there are $\frac{(q-3)(q-5)}{8}$ classes with representatives of this form.

We now determine the size of the conjugacy class of each of these elements. Since $a_{x,x}$ is not one of the canonical forms whose centralizer can be determined from Section 4.1, we first note that conjugation by the permutation matrix $P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ maps $\mathrm{Sp}(4,q)_{\widehat{B}}$ to $\mathrm{Sp}(4,q)_J$, where

$$J = P^T \widehat{B} P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

and maps $a_{x,x}$ to

$$a'_{x,x} = P^{-1} a_{x,x} P = \begin{pmatrix} x & 0 & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & 0 & x^{-1} & 0 \\ 0 & 0 & 0 & x^{-1} \end{pmatrix}.$$

By the results of Section 4.1, $\mathcal{C}_{\mathrm{M}(4,q)}(a'_{x,x}) = \mathrm{B}(\{2,2\},q)$, the set of all matrices of the form

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

for $A, B \in \mathrm{M}(2,q)$. But

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}^T \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} 0 & A^T B \\ -B^T A & 0 \end{pmatrix},$$

which equals $\left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$ if and only if $B = \left(A^T\right)^{-1}$. Hence

$$\mathcal{C}_{\mathrm{Sp}(4,q)_J}(a'_{x,x}) = \mathcal{C}_{\mathrm{M}(4,q)}(a'_{x,x}) \cap \mathrm{Sp}(4,q)_J = \left\{ \begin{pmatrix} A & 0 \\ 0 & \left(A^T\right)^{-1} \end{pmatrix} \middle| A \in \mathrm{GL}(2,q) \right\},$$

and so $\left|\mathcal{C}_{\mathrm{Sp}(4,q)_{\widehat{B}}}(a_{x,x})\right| = |\mathrm{GL}(2,q)| = (q-1)^2 q(q+1)$. Hence

$$|cl(a_{x,x})| = \frac{(q-1)^2 q^4 (q+1)^2 (q^2+1)}{(q-1)^2 q(q+1)} = q^3(q+1)(q^2+1).$$

The sizes of the classes of $a_{x,\pm 1}$, $a_{1,-1}$, and $a_{x,y}$ are determined similarly, except that now the computations are most easily carried out in $\mathrm{Sp}(4,q)_{\widehat{B}}$. For example, when x, y, x^{-1} and y^{-1} are all distinct, $\mathcal{C}_{\mathrm{M}(4,q)}(a_{x,y}) = \mathrm{B}(\{1,1,1,1\},q)$, the set of all diagonal matrices, and since, as noted before, a diagonal matrix is in $\mathrm{Sp}(4,q)_{\widehat{B}}$ if and only if its diagonal entries come in two inverse pairs, there are $(q-1)^2$ diagonal matrices in $\mathrm{Sp}(4,q)_{\widehat{B}}$. Thus $|cl(a_{x,y})| = q^4(q+1)^2(q^2+1)$. Table 4.1 lists, for each type of class discussed so far, the necessary restrictions on the defining parameters, the size of the class, and the number of classes of that type.

Representative	Parameters	Class Size	# of Classes
$a_{\pm 1,\pm 1}$	$-$	1	2
$a_{x,x}$	$x \notin \{-1,0,1\}$	$q^3(q+1)(q^2+1)$	$\frac{q-3}{2}$
$a_{x,\pm 1}$	$x \notin \{-1,0,1\}$	$q^3(q+1)(q^2+1)$	$q-3$
$a_{1,-1}$	$-$	$q^2(q^2+1)$	1
$a_{x,y}$	$\lvert\{x,x^{-1},y,y^{-1}\}\rvert = 4$	$q^4(q+1)^2(q^2+1)$	$\frac{(q-3)(q-5)}{8}$

Table 4.1: Class Sizes of $\mathrm{Sp}(4,q)$ Part 1.

We now consider the classes of matrices with canonical form

$$\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}.$$

By Lemma 4.2.1, such a matrix is in $\mathrm{Sp}(4,q)_{\widehat{B}}$ if and only $\lambda_1^2 = 1 \Rightarrow \lambda_1 = \pm 1$ and $\lambda_3 = \lambda_2^{-1}$. Thus we let

$$b_{\pm 1,x} = b^1_{\pm 1,x} = \begin{pmatrix} \pm 1 & 1 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x^{-1} \end{pmatrix}.$$

Now let ϵ be a generator of $\mathbb{F}_q^*$. (Recall that the multiplicative group of any finite field is cyclic.) Then all the elements of $\mathbb{F}_q^*$ which are squares form a subgroup of $\mathbb{F}_q^*$ of index 2, and so ϵ is not the square of any element of $\mathbb{F}_q^*$. As a result, it follows that the matrices $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & \epsilon \\ 0 & 1 \end{smallmatrix}\right)$ are not conjugate in $\mathrm{SL}(2,q)$. To see this, suppose they were conjugate via the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Then we must have

$$\begin{aligned} \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & \epsilon \\ 0 & 1 \end{smallmatrix}\right) &= \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \\ \Rightarrow \left(\begin{smallmatrix} a & a\epsilon+b \\ c & c\epsilon+d \end{smallmatrix}\right) &= \left(\begin{smallmatrix} a+c & b+d \\ c & d \end{smallmatrix}\right), \end{aligned}$$

which implies that $c = 0$ and $d = a\epsilon$. But then if $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2,q)$, we must have $ad - bc = 1 \Rightarrow a^2\epsilon = 1 \Rightarrow \epsilon = (a^{-1})^2$, contradicting our choice of ϵ. A similar argument shows that if we let

$$b^{\epsilon}_{\pm 1,x} = \begin{pmatrix} \pm 1 & \epsilon & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x^{-1} \end{pmatrix},$$

then $b^1_{\pm 1,x}$ and $b^{\epsilon}_{\pm 1,x}$ are not conjugate in $\mathrm{Sp}(4,q)_{\widehat{B}}$, despite having the same canonical form in $\mathrm{GL}(4,q)$. If $a \in \{1,\epsilon\}$, then $cl(b^a_{\pm 1,x}) = cl(b^a_{\pm 1,x^{-1}})$ but with this exception, distinct values of a and x yield representatives of distinct conjugacy classes. Hence there are 4 classes with representatives of the form[1] $b^a_{\pm 1,\pm 1}$, 4 of the form $b^a_{\pm 1,\mp 1}$ and $2(q-3)$ of the form $b^a_{\pm 1,x}$ for $x \notin \{-1,0,1\}$.

We will show how to compute $\left|cl(b^a_{\pm 1,\pm 1})\right|$, the computations for $b^a_{\pm 1,\mp 1}$ and $b^a_{\pm 1,x}$ being similar and easier. By the results of Section 4.1, we see that $\mathcal{C}_{\mathrm{M}(4,q)}(b^1_{\pm 1,\pm 1})$ consists of all matrices of the form

$$\begin{pmatrix} a & b & i & j \\ 0 & a & 0 & 0 \\ 0 & g & c & d \\ 0 & h & e & f \end{pmatrix},$$

as the various parameters used range over $\mathbb{F}_q$. Now $b^1_{\pm 1,\pm 1}$ is conjugate to $b^{\epsilon}_{\pm 1,\pm 1}$ in $\mathrm{GL}(4,q)$ via the diagonal matrix $D = \mathrm{B}(1,\epsilon,1,1)$, so conjugation by D maps $\mathcal{C}_{\mathrm{M}(4,q)}(b^1_{\pm 1,\pm 1})$ to $\mathcal{C}_{\mathrm{M}(4,q)}(b^{\epsilon}_{\pm 1,\pm 1})$. But

$$D^{-1}\begin{pmatrix} a & b & i & j \\ 0 & a & 0 & 0 \\ 0 & g & c & d \\ 0 & h & e & f \end{pmatrix} D = \begin{pmatrix} a & b\epsilon & i & j \\ 0 & a & 0 & 0 \\ 0 & g\epsilon & c & d \\ 0 & h\epsilon & e & f \end{pmatrix},$$

which is again in $\mathcal{C}_{\mathrm{M}(4,q)}(b^1_{\pm 1,\pm 1})$, and hence $\mathcal{C}_{\mathrm{M}(4,q)}(b^1_{\pm 1,\pm 1}) = \mathcal{C}_{\mathrm{M}(4,q)}(b^{\epsilon}_{\pm 1,\pm 1})$.

[1] We adopt the convention that if the symbol ± 1 is used twice in an equation it means either 1 in both places or -1 in both places, and similarly, $\mp 1 = -(\pm 1)$.

Thus $\left|\mathcal{C}_{\mathrm{Sp}(4,q)_{\widehat{B}}}(b^a_{\pm 1,\pm 1})\right|$ is simply the number of matrices of the form

$$\begin{pmatrix} a & b & i & j \\ 0 & a & 0 & 0 \\ 0 & g & c & d \\ 0 & h & e & f \end{pmatrix}$$

which are contained in $\mathrm{Sp}(4,q)_{\widehat{B}}$. If B is the symplectic form corresponding to $\widehat{B}$, then any matrix in $\mathrm{M}(4,q)$ with columns c_1, c_2, c_3, c_4 is contained in $\mathrm{Sp}(4,q)_{\widehat{B}}$ if and only if $B(c_1,c_2) = B(c_3,c_4) = 1$ and $B(c_1,c_3) = B(c_2,c_3) = B(c_1,c_4) = B(c_2,c_4) = 0$. Applying these equations to the matrix $\begin{pmatrix} a & b & i & j \\ 0 & a & 0 & 0 \\ 0 & g & c & d \\ 0 & h & e & f \end{pmatrix}$, we see that for it to be contained in $\mathrm{Sp}(4,q)_{\widehat{B}}$ we must have

$$a^2 = 1, cf - ed = 1, -ia + ge - ch = 0, \text{ and } -ja + gf - dh = 0.$$

Hence we must have $a = \pm 1$, $\left(\begin{smallmatrix} c & d \\ e & f \end{smallmatrix}\right) \in \mathrm{SL}(2,q)$, $i = a^{-1}(ge - ch)$, and $j = a^{-1}(gf - dh)$, but we can choose b, g, and h freely from $\mathbb{F}_q$. Hence

$$\left|\mathcal{C}_{\mathrm{Sp}(4,q)_{\widehat{B}}}(b^a_{\pm 1,\pm 1})\right| = 2\left|\mathbb{F}_q\right|^3 \left|\mathrm{SL}(2,q)\right| = 2(q-1)q^4(q+1),$$

and so

$$\left|cl(b^a_{\pm 1,\pm 1})\right| = \frac{(q-1)(q+1)(q^2+1)}{2}.$$

Computing the remaining class sizes similarly, we get Table 4.2.

Representative	Parameters	Class Size	# of Classes
$b^a_{\pm 1,\pm 1}$	$a \in \{1, \epsilon\}$	$\frac{(q-1)(q+1)(q^2+1)}{2}$	4
$b^a_{\pm 1,\mp 1}$	$a \in \{1, \epsilon\}$	$\frac{(q-1)q^2(q+1)(q^2+1)}{2}$	4
$b^a_{\pm 1,x}$	$a \in \{1, \epsilon\},\ x \notin \{-1, 0, 1\}$	$\frac{(q-1)q^3(q+1)^2(q^2+1)}{2}$	$2(q-3)$

Table 4.2: Class Sizes of $\mathrm{Sp}(4,q)$ Part 2.

We now consider classes of matrices with canonical form

$$\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}.$$

If such a matrix is contained in $\mathrm{Sp}(4,q)_{\widehat{B}}$, we must have $\lambda_1 = \pm 1$ and $\lambda_2 = \pm 1$. As in the previous case, we can replace a 1 above the diagonal with a generator ϵ of $\mathbb{F}_q^*$ to obtain a representative of a different class. Thus if we let

$$c^a_{\pm 1,\pm 1} = \begin{pmatrix} \pm 1 & a & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & 1 \\ 0 & 0 & 0 & \pm 1 \end{pmatrix} \text{ and } c^{a,b}_{1,-1} = \begin{pmatrix} 1 & a & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & b \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

where $a, b \in \{1, \epsilon\}$, we can show that there are 4 distinct classes with representatives of the form $c^a_{\pm 1,\pm 1}$ and 4 with representatives of the form $c^{a,b}_{1,-1}$.

To determine whether a matrix of the canonical form under consideration with $\lambda_i \neq \pm 1$ can be contained in $\mathrm{Sp}(4,q)_J$ for some other form matrix J, we note that if J is any matrix for a symplectic form, we can write

$$J = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix}$$

for values of the parameters such that J is non-singular. The system of equations corresponding to

$$\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}^T \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix}$$

has, among others, a family of solutions of the form

$$\lambda_1 = x, \lambda_2 = x^{-1}, a = b = e = f = 0, c = x, d = -x^{-1}$$

for any $x \in \mathbb{F}_q^*$, so if we set

$$J_x = \begin{pmatrix} 0 & 0 & 0 & x \\ 0 & 0 & -x^{-1} & 0 \\ 0 & x^{-1} & 0 & 0 \\ -x & 0 & 0 & 0 \end{pmatrix} \text{ and } c'_{x,x^{-1}} = \begin{pmatrix} x & 1 & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & 0 & x^{-1} & 1 \\ 0 & 0 & 0 & x^{-1} \end{pmatrix}$$

we see that $c'_{x,x^{-1}} \in \mathrm{Sp}(4,q)_{J_x}$. Under the matrix $\begin{pmatrix} x^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & x & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$, $c'_{x,x^{-1}}$ is conjugate to

$$c_{x,x^{-1}} = \begin{pmatrix} x & 0 & 0 & x \\ 0 & x^{-1} & 0 & 0 \\ 0 & x^{-1} & x^{-1} & 0 \\ 0 & 0 & 0 & x \end{pmatrix} \in \mathrm{Sp}(4,q)_{\widehat{B}}.$$

Thus since $c_{x,x^{-1}}$ and $c_{x^{-1},x}$ are conjugate, there are $\frac{q-3}{2}$ classes with representatives of the form $c_{x,x^{-1}}$ for $x \notin \{-1,0,1\}$.

The sizes of the classes of $c^{a,b}_{\pm 1,\mp 1}$ and $c_{x,x^{-1}}$ are relatively easily determined, although for the latter one should work in $\mathrm{Sp}(4,q)_{J_x}$, but determining $\left|cl(c^a_{\pm 1,\pm 1})\right|$ is sufficiently difficult to warrant further discussion. We will need the following lemma.

4.2.3 Lemma. Let $k, l \in \mathbb{F}_q^*$. Then

$$\left|\{(x,y) \in \mathbb{F}_q^2 \,\middle|\, x^2 - ky^2 = l\}\right| = \begin{cases} q-1 & \text{if } k \in \left(\mathbb{F}_q^*\right)^2 \\ q+1 & \text{if } k \notin \left(\mathbb{F}_q^*\right)^2 \end{cases},$$

where $\left(\mathbb{F}_q^*\right)^2$ is the subgroup of squares in $\mathbb{F}_q^*$.

Proof. If k is a square in $\mathbb{F}_q^*$, then $x^2 - ky^2 = l \Rightarrow (x + \sqrt{k}y)(x - \sqrt{k}y) = l$, so if we set $\lambda = x + \sqrt{k}y$, we have $\frac{l}{\lambda} = x - \sqrt{k}y$, and so

$$x = \frac{1}{2}\left(\lambda + \frac{l}{\lambda}\right) \text{ and } y = \frac{1}{2\sqrt{k}}\left(\lambda - \frac{l}{\lambda}\right).$$

Conversely, if we define x and y by the above formulae for any $\lambda \in \mathbb{F}_q^*$, a simple computation shows that $x^2 - ky^2 = l$. Thus there is a bijection between solutions of $x^2 - ky^2 = l$ and $\mathbb{F}_q^*$, and so the number of solutions is $q - 1$.

Now suppose that k is not a square in $\mathbb{F}_q^*$. Then $\mathbb{F}_q[\sqrt{k}]$ is a degree 2 extension of $\mathbb{F}_q$, and thus can be identified with $\mathbb{F}_{q^2}$. Now recall that in general, if L is a finite extension of the field K of degree d, there is a homomorphism $N_{L/K} : L^* \to K^*$ such that the value of $N_{L/K}$ on any element of L is the determinant of any matrix for the element in $\mathrm{GL}(d, K)$. Further, when L and K are finite fields, $N_{L/K}$ is surjective. (See, for example, [5] page 115.) Thus there is an epimorphism $N : \mathbb{F}_{q^2}^* \to \mathbb{F}_q^*$ given by $N(x + \sqrt{k}y) = x^2 - ky^2$, and so the set of solutions of $x^2 - ky^2 = l$ is in bijection with a coset of the kernel of N. Since N is an epimorphism,

$$|\ker N| = \frac{\left|\mathbb{F}_{q^2}^*\right|}{\left|\mathbb{F}_q^*\right|} = \frac{q^2 - 1}{q - 1} = q + 1,$$

as desired. □

We will now show how to compute $\left|cl(c^{\epsilon}_{\pm 1,\pm 1})\right|$; the computations for $c^1_{\pm 1,\pm 1}$ follow identically by replacing ϵ with 1. Since $c^{\epsilon}_{\pm 1,\pm 1}$ is not in the canonical form of Section 4.1, we conjugate it by the matrix $D = \begin{pmatrix} 1&0&0&0\\ 0&\epsilon&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}$ to obtain $c'_{\pm 1,\pm 1} = \begin{pmatrix} \pm 1&1&0&0\\ 0&\pm 1&0&0\\ 0&0&\pm 1&1\\ 0&0&0&\pm 1 \end{pmatrix} \in \mathrm{Sp}(4, q)_J$, where

$$J = D^T \widehat{B} D = \begin{pmatrix} 0&\epsilon&0&0\\ -\epsilon&0&0&0\\ 0&0&0&1\\ 0&0&-1&0 \end{pmatrix}.$$

By the results of Section 4.1, $\mathcal{C}_{\mathrm{M}(4,q)}(c'_{\pm 1,\pm 1})$ consists of all matrices of the form

$$A = \begin{pmatrix} a&b&c&d\\ 0&a&0&c\\ e&f&g&h\\ 0&e&0&g \end{pmatrix}.$$

Applying the condition $A^T J A = J$ yields the equations

$$e^2 + \epsilon a^2 = \epsilon, g^2 + \epsilon c^2 = 1, eg + ac\epsilon = 0, \text{ and } fg - eh + bc\epsilon - ad\epsilon = 0.$$

Now ϵ is non-square in $\mathbb{F}_q^*$, so $-\epsilon$ is non-square if and only if -1 is a square in $\mathbb{F}_q^*$. But -1 is square or non-square according as $q \equiv 1 \mod 4$ or $q \equiv 3 \mod 4$. Thus by Lemma 4.2.3, the number of $(e, a) \in \mathbb{F}_q^2$ with $e^2 + \epsilon a^2 = \epsilon$ is $q + 1$ if $q \equiv 1 \mod 4$ and $q - 1$ if $q \equiv 3 \mod 4$.

Now suppose we have fixed a pair (e, a) with $e^2 + \epsilon a^2 = \epsilon$. If $e = 0$, the equations

$$g^2 + \epsilon c^2 = 1 \text{ and } eg + ac\epsilon = 0$$

imply $c = 0$ and $g = \pm 1$, while if $e \neq 0$, they imply that $c = \pm e\epsilon^{-1}$ and $g = \mp a$. Thus in all cases there are two possibilities for (g, c). Once we have chosen a, e, f, and g, at least one of a and e must be non-zero, so we can select 3 of b, d, f, and h arbitrarily from $\mathbb{F}_q$ and use the equation

$$fg - eh + bc\epsilon - ad\epsilon = 0$$

to solve for the fourth. Multiplying the number of choices at each step, we see that

$$\left|\mathcal{C}_{\mathrm{Sp}(4,q)_J}(c^{\epsilon}_{\pm1,\pm1})\right| = \begin{cases} 2q^3(q+1) & \text{if } q \equiv 1 \mod 4 \\ 2q^3(q-1) & \text{if } q \equiv 3 \mod 4 \end{cases},$$

and from here, the class sizes are easily determined. Table 4.3 records the usual data.

Representative	Parameters	Class Size	# of Classes
$c_{\pm1,\pm1}$	$-$	$\frac{(q-1)q(q+1)^2(q^2+1)}{2}$ if $q\equiv1 \mod 4$ $\frac{(q-1)^2q(q+1)(q^2+1)}{2}$ if $q\equiv3 \mod 4$	2
$c^{\epsilon}_{\pm1,\pm1}$	$-$	$\frac{(q-1)^2q(q+1)(q^2+1)}{2}$ if $q\equiv1 \mod 4$ $\frac{(q-1)q(q+1)^2(q^2+1)}{2}$ if $q\equiv3 \mod 4$	2
$c^{a,b}_{1,-1}$	$a, b \in \{1, \epsilon\}$	$\frac{(q-1)^2q^2(q+1)^2(q^2+1)}{4}$	4
$c_{x,x^{-1}}$	$x \notin \{-1, 0, 1\}$	$(q-1)q^3(q+1)^2(q^2+1)$	$\frac{q-3}{2}$

Table 4.3: Class Sizes of $\mathrm{Sp}(4, q)$ Part 3.

If we attempt to use the technique which produced $c'_{x,x^{-1}}$ to find a matrix J such that

$$\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \in \mathrm{Sp}(4, q)_J$$

for some λ_1, λ_2, we find that the resulting equations have no solutions for which J is non-singular. Hence no elements of $\mathrm{Sp}(4, q)$ have this canonical form.

If we use this technique for matrices with canonical form

$$\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix},$$

we find that for any $a \in \mathbb{F}_q^*$, if $J^a_{\pm1} = \begin{pmatrix} 0 & 0 & 0 & a \\ 0 & 0 & -a & \pm a \\ 0 & a & 0 & 0 \\ -a & \mp a & 0 & 0 \end{pmatrix}$, then

$$d'_{\pm1} = \begin{pmatrix} \pm1 & 1 & 0 & 0 \\ 0 & \pm1 & 1 & 0 \\ 0 & 0 & \pm1 & 1 \\ 0 & 0 & 0 & \pm1 \end{pmatrix} \in \mathrm{Sp}(4, q)_{J^a_{\pm1}}.$$

If we set $D = \begin{pmatrix} 1 & 0 & 0 & \mp a^{-1} \\ 0 & 0 & 0 & a^{-1} \\ 0 & 0 & 1 & 0 \\ 0 & a^{-1} & 0 & 0 \end{pmatrix}$, then $D^T J^a_{\pm1} D = \widehat{B}$, and so conjugation by D maps $d'_{\pm1}$ to an element of $\mathrm{Sp}(4, q)_{\widehat{B}}$, namely

$$d^a_{\pm1} = \begin{pmatrix} \pm1 & 0 & \pm1 & a^{-1} \\ 0 & \pm1 & 0 & 0 \\ 0 & a^{-1} & \pm1 & 0 \\ 0 & 0 & a & \pm1 \end{pmatrix}.$$

If we write down the general form of a matrix by which $d^1_{\pm1}$ is conjugate to $d^\epsilon_{\pm1}$, the equations for membership of this matrix in $\mathrm{Sp}(4,q)_{\widehat{B}}$ have no solution, so $d^1_{\pm1}$ and $d^\epsilon_{\pm1}$ are not conjugate. Thus there are 4 distinct classes with representatives of the form $d^a_{\pm1}$ for $a \in \{1,\epsilon\}$. By working in $\mathrm{Sp}(4,q)_{J^a_{\pm1}}$, we can readily compute the size of these classes, yielding Table 4.4.

Representative	Parameters	Class Size	# of Classes
$d^a_{\pm1}$	$a \in \{1,\epsilon\}$	$\frac{(q-1)^2q^2(q+1)^2(q^2+1)}{2}$	4

Table 4.4: Class Sizes of Sp(4, q) Part 4.

We now consider the classes of matrices with canonical forms

$$\begin{pmatrix} 0 & \alpha_2 & 0 & 0 \\ 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & \alpha_2 & 0 & 0 \\ 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$$

for $t^2-\alpha_1 t-\alpha_2$ irreducible in $\mathbb{F}_q[t]$. By Lemma 4.2.1, if a matrix of either type is contained in $\mathrm{Sp}(4,q)_{\widehat{B}}$ we must have $\alpha_2 = -1$. For the first, we must have $\lambda_2 = \lambda_1^{-1}$, while for the second we must have $\lambda_1 = \pm 1$. Thus for $t^2 - xt + 1$ irreducible in $\mathbb{F}_q[t]$, $y \in \mathbb{F}_q^*$, and $a \in \{1,\epsilon\}$ we define

$$e_{x,y} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & x & 0 & 0 \\ 0 & 0 & y & 0 \\ 0 & 0 & 0 & y^{-1} \end{pmatrix} \text{ and } f^a_{x,\pm1} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & x & 0 & 0 \\ 0 & 0 & \pm1 & a \\ 0 & 0 & 0 & \pm1 \end{pmatrix}.$$

Since different irreducible polynomials t^2-xt+1 give rise to representatives of different classes, we need to know how many irreducible polynomials of this type are contained in $\mathbb{F}_q[t]$. If ϵ is the usual non-square element of $\mathbb{F}_q^*$, then the roots of any irreducible polynomials of degree 2 in $\mathbb{F}_q[t]$ are contained in $\mathbb{F}_q[\sqrt{\epsilon}]$. Further, if $\xi = u + \sqrt{\epsilon}v$ is a root of such a polynomial for $u, v \in \mathbb{F}_q$, the other root of the polynomial is the conjugate of ξ, $\overline{\xi} = u - \sqrt{\epsilon}v$. If ξ and $\overline{\xi}$ of this form are roots of t^2-xt+1, we must have $\xi\overline{\xi} = 1 \Rightarrow u^2-\epsilon v^2 = 1 \Rightarrow N(\xi) = 1$, where $N : \mathbb{F}_{q^2}^* \to \mathbb{F}_q^*$ is the homomorphism used in the proof of Lemma 4.2.3. As noted in that proof, there are $q+1$ elements ξ in $\mathbb{F}_{q^2}^*$ such that $N(\xi) = 1$, and if one excludes $\xi = \pm 1$, these elements are precisely the roots in $\mathbb{F}_{q^2}^*$ of the irreducible polynomials of the form $t^2 - xt + 1$. Since there are two roots corresponding to each such polynomial, there must be $\frac{q-1}{2}$ irreducible polynomials of this type.

With this in mind, we see that there are $q-1$ classes with representatives of the form $e_{x,\pm1}$. Since $e_{x,y}$ and $e_{x,y^{-1}}$ are conjugate, there are $\frac{(q-1)(q-3)}{4}$ classes with representatives of the form $e_{x,y}$ for $y \notin \{-1,0,1\}$, and since $f^1_{x,\pm1}$ and $f^\epsilon_{x,\pm1}$ are not conjugate in $\mathrm{Sp}(4,q)_{\widehat{B}}$, there are $2(q-1)$ classes with representatives of the form $f^a_{x,\pm1}$.

Determining the sizes of these classes is relatively simple. For example, the results of Section 4.1 imply that $\mathcal{C}_{\mathrm{M}(4,q)}(e_{x,y})$ consists of all matrices of the form $\begin{pmatrix} a & -b & 0 & 0 \\ b & a+bx & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix}$. According to Lemma 4.2.1, such a matrix is contained in

$\mathrm{Sp}(4,q)_{\widehat{B}}$ if and only if $d = c^{-1}$ and $a^2+abx+b^2 = 1$. But if ξ is a root in $\mathbb{F}_{q^2}$ of t^2-xt+1, then for any $a,b \in \mathbb{F}_q$, $N(a+b\xi) = a^2+abx+b^2$, where N is the homomorphism of Lemma 4.2.3, so pairs $(a,b) \in \mathbb{F}_q^2$ such that $a^2+abx+b^2 = 1$ are in bijection with $\ker N$. Hence there are $q+1$ choices for a and b, and $q-1$ choices for c, so that $\left|\mathcal{C}_{\mathrm{Sp}(4,q)_{\widehat{B}}}(e_{x,y})\right| = (q-1)(q+1)$, and from this the class size follows. Table 4.5 records the usual data; in addition to the listed restrictions on the parameters, we require that t^2-xt+1 be irreducible in $\mathbb{F}_q[t]$.

Representative	Parameters	Class Size	# of Classes
$e_{x,\pm 1}$	$-$	$(q-1)q^3(q^2+1)$	$q-1$
$e_{x,y}$	$y \notin \{-1,0,1\}$	$(q-1)q^4(q+1)(q^2+1)$	$\frac{(q-1)(q-3)}{4}$
$f^a_{x,\pm 1}$	$a \in \{1,\epsilon\}$	$\frac{(q-1)^2q^3(q+1)(q^2+1)}{2}$	$2(q-1)$

Table 4.5: Class Sizes of Sp(4, q) Part 5.

We now consider classes with representatives of the form

$$\begin{pmatrix} 0 & \alpha_2 & 0 & 0 \\ 1 & \alpha_1 & 0 & 0 \\ 0 & 0 & 0 & \alpha_4 \\ 0 & 0 & 1 & \alpha_3 \end{pmatrix}.$$

Such a matrix is contained in $\mathrm{Sp}(4,q)_{\widehat{B}}$ if and only if $\alpha_2 = \alpha_4 = -1$. Thus for t^2-xt+1 and t^2-yt+1 irreducible in $\mathbb{F}_q[t]$ we define

$$g_{x,y} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & x & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & y \end{pmatrix}.$$

There are $\frac{q-1}{2}$ classes with representatives of the form $g_{x,x}$. Since $cl(g_{x,y}) = cl(g_{y,x})$, there are $\frac{1}{2}\frac{q-1}{2}\left(\frac{q-1}{2}-1\right) = \frac{(q-1)(q-3)}{8}$ classes with representatives of the form $g_{x,y}$ for $x \neq y$.

To obtain one more type of class with representatives of this form, let $J = \left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$. Recall from the computations for $a_{x,x}$ that a matrix of the form $\left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)$ is contained in $\mathrm{Sp}(4,q)_J$ if and only if $B = \left(A^T\right)^{-1}$. Thus if $A = \left(\begin{smallmatrix} 0 & y \\ 1 & x \end{smallmatrix}\right)$ for t^2-xt-y irreducible in $\mathbb{F}_q[t]$, we should set $B = \left(\begin{smallmatrix} -xy^{-1} & y^{-1} \\ 1 & 0 \end{smallmatrix}\right)$ to obtain the element

$$h'_{x,y} = \begin{pmatrix} 0 & y & 0 & 0 \\ 1 & x & 0 & 0 \\ 0 & 0 & -xy^{-1} & y^{-1} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

of $\mathrm{Sp}(4,q)_J$. The matrix $h'_{x,y}$ is conjugate in $\mathrm{GL}(4,q)$ via a permutation matrix to

$$h_{x,y} = \begin{pmatrix} 0 & 0 & y & 0 \\ 0 & -xy^{-1} & 0 & y^{-1} \\ 1 & 0 & x & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \in \mathrm{Sp}(4,q)_{\widehat{B}}.$$

Provided $y \neq -1$, $h_{x,y}$ is not conjugate to any $g_{x,y}$. Thus since $cl(h_{x,y}) = cl(h_{-xy^{-1},y^{-1}})$, the number of conjugacy classes with representatives of this type is half the number of monic irreducible quadratics $t^2 - xt - y$ with $y \neq -1$. Since $\mathbb{F}_{q^2}$ contains $q^2 - q$ elements of degree 2 over $\mathbb{F}_q$, there are $\frac{q^2-q}{2}$ monic irreducible quadratics in $\mathbb{F}_q[t]$. Hence since $\frac{q-1}{2}$ of these have $y = -1$, there are $\frac{q^2-q}{2} - \frac{q-1}{2} = \frac{(q-1)^2}{2}$ irreducible quadratics with $y \neq -1$, and thus $\frac{(q-1)^2}{4}$ classes with representatives of the form $h_{x,y}$.

Determining $|cl(g_{x,y})|$ is straightforward. Determining $|cl(h_{x,y})|$ is relatively straightforward, but is best done through $h'_{x,y}$. (Care must be taken in applying the results of Section 4.1, since $h'_{x,y}$ is not quite in the canonical form of that section.) By the results of Section 4.1, $\mathcal{C}_{\mathrm{M}(4,q)}(g_{x,x})$ consists of all matrices of the form

$$\begin{pmatrix} a & -b & c & -d \\ b & a+bx & d & c+dx \\ e & -f & g & -h \\ f & e+fx & h & g+hx \end{pmatrix}.$$

Applying the condition that this element is contained in $\mathrm{Sp}(4,q)_{\widehat{B}}$ yields equations which are equivalent to the statement that if ξ and $\overline{\xi}$ are the roots in $\mathbb{F}_{q^2}$ of $t^2 - xt + 1$, then

$$\begin{pmatrix} a+b\overline{\xi} & e+f\overline{\xi} \\ c+d\overline{\xi} & g+h\overline{\xi} \end{pmatrix} \begin{pmatrix} a+b\xi & c+d\xi \\ e+f\xi & g+h\xi \end{pmatrix} = I,$$

and so $\mathcal{C}_{\mathrm{Sp}(4,q)}(g_{x,x})$ is isomorphic to $\mathrm{U}(2,q^2)$, the group of unitary matrices over $\mathbb{F}_{q^2}$. Hence $\left|\mathcal{C}_{\mathrm{Sp}(4,q)}(g_{x,x})\right| = |\mathrm{U}(2,q^2)| = (q-1)q(q+1)^2$. (For a treatment of the basic properties of the unitary groups, including their orders, see [6].) Table 4.6 records the usual information; we assume the parameters are such that the corresponding quadratics are irreducible.

Representative	Parameters	Class Size	# of Classes
$g_{x,x}$	–	$(q-1)q^3(q^2+1)$	$\frac{q-1}{2}$
$g_{x,y}$	–	$(q-1)^2q^4(q^2+1)$	$\frac{(q-1)(q-3)}{8}$
$h_{x,y}$	$y \neq -1$	$(q-1)q^4(q+1)(q^2+1)$	$\frac{(q-1)^2}{4}$

Table 4.6: Class Sizes of Sp$(4,q)$ Part 6.

If we attempt to solve the equations on the various parameters corresponding to membership of

$$\begin{pmatrix} 0 & \alpha_2 & 1 & 0 \\ 1 & \alpha_1 & 0 & 1 \\ 0 & 0 & 0 & \alpha_2 \\ 0 & 0 & 1 & \alpha_1 \end{pmatrix}$$

in $\mathrm{Sp}(4,q)_J$, where $J = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix}$ and $t^2 - \alpha_1 t - \alpha_2$ is irreducible, we find that we must have $\alpha_2 = -1, a = 0, e = b = \frac{cx}{2}$, and $d = c\left(\frac{x^2}{2} - 1\right)$, where $x = \alpha_1$. (To arrive at these equations it is helpful to note that since $\mathrm{Sp}(V)$ is contained in $\mathrm{SL}(V)$ for any vector space V, we must have $\alpha_2 = \pm 1$. If $\alpha_2 = 1$, the equations imply $\alpha_1 = 0$, which contradicts the irreducibility of $t^2 - \alpha_1 t - \alpha_2$.

Thus $\alpha_2 = -1$, and from here the remaining equations follow easily.) To keep J relatively simple, we set $c = 2$ and $f = 0$, so that $J = \begin{pmatrix} 0 & 0 & x & 2 \\ 0 & 0 & x^2-2 & x \\ -x & 2-x^2 & 0 & 0 \\ -2 & -x & 0 & 0 \end{pmatrix}$. Then if we set $k'_x = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & x & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & x \end{pmatrix}$ we have that $k'_x \in \mathrm{Sp}(4,q)_J$ for $t^2 - xt + 1$ irreducible. If we set $D = \begin{pmatrix} 1 & 0 & -x & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & \frac{1}{x^2-4} \\ 0 & \frac{1}{2} & 0 & \frac{-x}{2x^2-8} \end{pmatrix}$, then $D^T J D = \widehat{B}$, and so if we set

$$k_x = D^{-1} k'_x D = \begin{pmatrix} \frac{x}{2} & \frac{x}{4} & \frac{x^2}{2}-2 & -\frac{1}{4} \\ 0 & \frac{x}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & \frac{x}{2} & \frac{-x}{4x^2-16} \\ 0 & 2-\frac{x^2}{2} & 0 & \frac{x}{2} \end{pmatrix},$$

we see that $k_x \in \mathrm{Sp}(4,q)_{\widehat{B}}$ for t^2-xt+1 irreducible. (Note that since t^2-xt+1 is irreducible, $x \neq \pm 2$, and thus all denominators in the previous expressions are non-zero.) The number of classes with representatives of this form is equal to the number of irreducible quadratics of the form t^2-xt+1, which, as previously calculated, is $\frac{q-1}{2}$. Now by the results of Section 4.1, $\mathcal{C}_{\mathrm{Sp}(4,q)_J}(k'_x)$ consists of all matrices of the form

$$A = \begin{pmatrix} a & -b & c & -d \\ b & a+bx & d & c+dx \\ 0 & 0 & a & -b \\ 0 & 0 & b & a+bx \end{pmatrix}.$$

Setting $A^T J A = J$ yields a seemingly complicated system of equations which reduces to $a^2 + abx + b^2 = 1$ and $ac + (ax+b)d = 0$. Thus by previous results there are $q+1$ choices for (a, b), and since for each of these either $a \neq 0$ or $ax + b \neq 0$, we can solve for one of c and d in terms of the other, so that there are q choices for (c, d). Hence $\left|\mathcal{C}_{\mathrm{Sp}(4,q)_{\widehat{B}}}(k_x)\right| = q(q+1)$, which yields Table 4.7. (As usual, we require that $t^2 - xt + 1$ be irreducible in $\mathbb{F}_q[t]$.)

Representative	Parameters	Class Size	# of Classes
k_x	–	$(q-1)^2q^3(q+1)(q^2+1)$	$\frac{q-1}{2}$

Table 4.7: Class Sizes of Sp(4, q) Part 7.

A matrix of the form $\begin{pmatrix} 0 & 0 & \beta_3 & 0 \\ 1 & 0 & \beta_2 & 0 \\ 0 & 1 & \beta_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$ for $t^3 - \beta_1 t^2 - \beta_2 t - \beta_3$ irreducible cannot be contained in $\mathrm{Sp}(4,q)_J$ for any form matrix J. To see this, note that such a matrix is diagonalizable in $\mathrm{GL}(4, q^3)$, and therefore if it were contained in $\mathrm{Sp}(4,q)_J$ for any form matrix J, Lemma 4.2.2 would imply that its eigenvalues formed two pairs, with each pair consisting of a number and its inverse. But we know that the eigenvalues of the matrix in question are λ_1, an element of $\mathbb{F}_q$, and the three roots of $t^3 - \beta_1 t^2 - \beta_2 t - \beta_3$, which have degree 3 over $\mathbb{F}_q$, and so no such pairing of the eigenvalues is possible.

Finally, we consider matrices of the form

$$\begin{pmatrix} 0 & 0 & 0 & \gamma_4 \\ 1 & 0 & 0 & \gamma_3 \\ 0 & 1 & 0 & \gamma_2 \\ 0 & 0 & 1 & \gamma_1 \end{pmatrix}$$

for $m(t) = t^4 - \gamma_1 t^3 - \gamma_2 t^2 - \gamma_3 t - \gamma_4$ irreducible. Since such a matrix is diagonalizable in $\mathrm{GL}(4, q^4)$, Lemma 4.2.2 implies that its eigenvalues, the four distinct roots of $m(t)$, come in inverse pairs. Now suppose that ξ is a root of $m(t)$. Since the Galois group of $\mathbb{F}_{q^4}$ over $\mathbb{F}_q$ is cyclic of order 4, being generated by the automorphism σ_q such that for $\alpha \in \mathbb{F}_{q^4}$, $\sigma_q(\alpha) = \alpha^q$ (see, for example, [3] page 452), the other 3 roots of $m(t)$ must be ξ^q, ξ^{q^2}, and ξ^{q^3}. Since the eigenvalues of the matrix come in inverse pairs, one of these 3 roots must be ξ^{-1}. But if $\xi^{q+1} = 1$ or $\xi^{q^3+1} = 1$, it follows that $\xi^{q^2-1} = 1$, which implies that $\xi \in \mathbb{F}_{q^2}$. This cannot be the case if $m(t)$ is irreducible, so we must have $\xi^{-1} = \xi^{q^2}$, or in other words, $\xi^{q^2+1} = 1$. Now since $m(t)$ has ξ, ξ^{-1}, ξ^q, and ξ^{-q} as roots,

$$m(t) = (t - \xi)(t - \xi^{-1})(t - \xi^q)(t - \xi^{-q}).$$

Multiplying this expression out shows that $\gamma_1 = \gamma_3$ and $\gamma_4 = -1$, so that for certain $x, y \in \mathbb{F}_q$,

$$m(t) = t^4 - xt^3 - yt^2 - xt + 1.$$

Thus we set

$$m'_{x,y} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & x \end{pmatrix}.$$

If we attempt to solve the equation $(m'_{x,y})^T J m'_{x,y} = J$ for the unknown form matrix J, we find that $m'_{x,y} \in \mathrm{Sp}(4, q)_J$ for

$$J = \begin{pmatrix} 0 & 0 & 1 & x \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ -x & -1 & 0 & 0 \end{pmatrix}.$$

$m'_{x,y}$ is conjugate in $\mathrm{GL}(4, q)$ to the matrix

$$m_{x,y} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & x & 1 & y \\ 1 & 0 & 0 & x \\ 0 & 1 & 0 & 0 \end{pmatrix} \in \mathrm{Sp}(4, q)_{\widehat{B}}.$$

To determine the number of classes with representatives of the form $m_{x,y}$, we note first that we get one such class for each irreducible quartic in $\mathbb{F}_q[t]$ whose roots ξ satisfy $\xi^{q^2+1} = 1$. Since each of these quartics has 4 distinct roots, the number of classes is one-fourth the number of elements $\xi \in \mathbb{F}_{q^4}$ which are of degree 4 over $\mathbb{F}_q$ and satisfy $\xi^{q^2+1} = 1$. Since $\mathbb{F}_{q^4}^*$ is cyclic, the elements satisfying $\xi^{q^2+1} = 1$ form a subgroup of order $q^2 + 1$; but not all of these elements are of degree 4 over $\mathbb{F}_q$. More specifically, $\mathbb{F}_{q^2}$ is the unique maximal proper subfield of $\mathbb{F}_{q^4}$ containing $\mathbb{F}_q$, so if $\xi^{q^2+1} = 1$, but ξ is not of degree 4 over $\mathbb{F}_q$, we must have $\xi \in \mathbb{F}_{q^2}$. But this implies that $\xi^{q^2-1} = 1$, which when combined with $\xi^{q^2+1} = 1$ shows that $\xi = \pm 1$. Hence there are $q^2 - 1$ degree

4 elements in the subgroup, and so there are $\frac{q^2-1}{4}$ classes with representatives of the form $m_{x,y}$.

To determine the size of these classes, first choose an $m'_{x_0,y_0} \in \mathrm{Sp}(4,q)_J$ such that the roots of $t^4 - x_0t^3 - y_0t^2 - x_0t + 1$ generate the subgroup of $\mathbb{F}^*_{q^4}$ of order q^2+1. Then m'_{x_0,y_0} itself has order q^2+1, and so the cyclic subgroup H it generates contains representatives of all the classes which have representatives of the form $m_{x,y}$ (with respect to $\widehat{B}$). Further, since all of the elements of H except $\pm I$ correspond to elements of $\mathbb{F}_{q^4}$ of degree 4 over $\mathbb{F}_q$, they all generate the same subalgebra $\langle m'_{x_0,y_0}\rangle$ of $\mathrm{M}(4,q)$, one isomorphic to $\mathbb{F}_{q^4}$. By Theorem 4.1.11, $\langle m'_{x_0,y_0}\rangle$ is self-centralizing in $\mathrm{M}(4,q)$, and hence any element of H except $\pm I$ has as its centralizer in $\mathrm{Sp}(4,q)_J$ the subgroup $\langle m'_{x_0,y_0}\rangle \cap \mathrm{Sp}(4,q)_J$. This subgroup obviously contains H. It cannot properly contain H, for if it did, it would necessarily contain elements with eigenvalues in $\mathbb{F}_{q^4}$ of degree 4 over $\mathbb{F}_q$ and of order not dividing q^2+1. But as seen earlier, no element of $\mathrm{Sp}(4,q)$ can have eigenvalues of this sort. Hence $\left|\mathcal{C}_{\mathrm{Sp}(4,q)_{\widehat{B}}}(m_{x,y})\right| = |H| = q^2+1$. Table 4.8 records the usual data; we require, of course, that $t^4 - xt^3 - yt^2 - xt + 1$ be irreducible in $\mathbb{F}_q[t]$.

Representative	Parameters	Class Size	# of Classes
$m_{x,y}$	–	$(q-1)^2q^4(q+1)^2$	$\frac{q^2-1}{4}$

Table 4.8: Class Sizes of $\mathrm{Sp}(4,q)$ Part 8.

We claim that the preceding list of conjugacy classes is complete. To prove this, one can simply check that sum over all the preceding tables of the product of the 'Class Size' and '# of Classes' entries is, indeed, $|\mathrm{Sp}(4,q)| = (q-1)^2q^4(q+1)^2(q^2+1)$.

Tables 4.9 and 4.10 summarize the list of conjugacy classes of $\mathrm{Sp}(4,q)$ derived in this section. The matrix for each class representative is given with respect to $\widehat{B}$. The 'Repetitions' column gives the number of representatives of each class of the given type which occur as the parameters range over the allowable values.

Representative	Matrix in $\mathrm{Sp}(4,q)_{\widehat{B}}$	Parameters	Repetitions	Class Size	# of Classes
$a_{\pm 1,\pm 1}$	$\begin{pmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & \pm 1 \end{pmatrix}$	-	1	1	2
$a_{x,x}$	$\begin{pmatrix} x & 0 & 0 & 0 \\ 0 & x^{-1} & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x^{-1} \end{pmatrix}$	$x \notin \{-1,0,1\}$	2	$q^3(q+1)(q^2+1)$	$\frac{q-3}{2}$
$a_{x,\pm 1}$	$\begin{pmatrix} x & 0 & 0 & 0 \\ 0 & x^{-1} & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & \pm 1 \end{pmatrix}$	$x \notin \{-1,0,1\}$	2	$q^3(q+1)(q^2+1)$	$q-3$
$a_{1,-1}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$	-	1	$q^2(q^2+1)$	1
$a_{x,y}$	$\begin{pmatrix} x & 0 & 0 & 0 \\ 0 & x^{-1} & 0 & 0 \\ 0 & 0 & y & 0 \\ 0 & 0 & 0 & y^{-1} \end{pmatrix}$	$\lvert\{x,x^{-1},y,y^{-1}\}\rvert = 4$	8	$q^4(q+1)^2(q^2+1)$	$\frac{(q-3)(q-5)}{8}$
$b^a_{\pm 1,\pm 1}$	$\begin{pmatrix} \pm 1 & a & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & \pm 1 \end{pmatrix}$	$a \in \{1,\epsilon\}$	1	$\frac{(q-1)(q+1)(q^2+1)}{2}$	4
$b^a_{\pm 1,\mp 1}$	$\begin{pmatrix} \pm 1 & a & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \mp 1 & 0 \\ 0 & 0 & 0 & \mp 1 \end{pmatrix}$	$a \in \{1,\epsilon\}$	1	$\frac{(q-1)q^2(q+1)(q^2+1)}{2}$	4
$b^a_{\pm 1,x}$	$\begin{pmatrix} \pm 1 & a & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x^{-1} \end{pmatrix}$	$a \in \{1,\epsilon\},\ x \notin \{-1,0,1\}$	2	$\frac{(q-1)q^3(q+1)^2(q^2+1)}{2}$	$2(q-3)$
$c_{\pm 1,\pm 1}$	$\begin{pmatrix} \pm 1 & 1 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & 1 \\ 0 & 0 & 0 & \pm 1 \end{pmatrix}$	-	1	$\frac{(q-1)q(q+1)^2(q^2+1)}{2}$ if $q \equiv 1 \mod 4$ $\frac{(q-1)^2q(q+1)(q^2+1)}{2}$ if $q \equiv 3 \mod 4$	2
$c^{\epsilon}_{\pm 1,\pm 1}$	$\begin{pmatrix} \pm 1 & \epsilon & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & 1 \\ 0 & 0 & 0 & \pm 1 \end{pmatrix}$	-	1	$\frac{(q-1)^2q(q+1)(q^2+1)}{2}$ if $q \equiv 1 \mod 4$ $\frac{(q-1)q(q+1)^2(q^2+1)}{2}$ if $q \equiv 3 \mod 4$	2
$c^{a,b}_{1,-1}$	$\begin{pmatrix} 1 & a & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & b \\ 0 & 0 & 0 & -1 \end{pmatrix}$	$a,b \in \{1,\epsilon\}$	1	$\frac{(q-1)^2q^2(q+1)^2(q^2+1)}{4}$	4

Table 4.9: Conjugacy Classes of $\mathrm{Sp}(4,q)$ for q odd Part 1.

Representative	Matrix in $\mathrm{Sp}(4,q)_{\widehat{B}}$	Parameters	Repetitions	Class Size	# of Classes
$c_{x,x^{-1}}$	$\begin{pmatrix} x & 0 & 0 & x \\ 0 & x^{-1} & 0 & 0 \\ 0 & x^{-1} & x^{-1} & 0 \\ 0 & 0 & 0 & x \end{pmatrix}$	$x \notin \{-1,0,1\}$	2	$(q-1)q^3(q+1)^2(q^2+1)$	$\frac{q-3}{2}$
$d^a_{\pm 1}$	$\begin{pmatrix} \pm 1 & 0 & \pm 1 & a^{-1} \\ 0 & \pm 1 & 0 & 0 \\ 0 & a^{-1} & \pm 1 & 0 \\ 0 & 0 & a & \pm 1 \end{pmatrix}$	$a \in \{1,\epsilon\}$	1	$\frac{(q-1)^2q^2(q+1)^2(q^2+1)}{2}$	4
$e_{x,\pm 1}$	$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & x & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & \pm 1 \end{pmatrix}$	t^2-xt+1 irreducible	1	$(q-1)q^3(q^2+1)$	$q-1$
$e_{x,y}$	$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & x & 0 & 0 \\ 0 & 0 & y & 0 \\ 0 & 0 & 0 & y^{-1} \end{pmatrix}$	t^2-xt+1 irreducible $y \notin \{-1,0,1\}$	2	$(q-1)q^4(q+1)(q^2+1)$	$\frac{(q-1)(q-3)}{4}$
$f^a_{x,\pm 1}$	$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & x & 0 & 0 \\ 0 & 0 & \pm 1 & a \\ 0 & 0 & 0 & \pm 1 \end{pmatrix}$	t^2-xt+1 irreducible $a \in \{1,\epsilon\}$	1	$\frac{(q-1)^2q^3(q+1)(q^2+1)}{2}$	$2(q-1)$
$g_{x,x}$	$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & x & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & x \end{pmatrix}$	t^2-xt+1 irreducible	1	$(q-1)q^3(q^2+1)$	$\frac{q-1}{2}$
$g_{x,y}$	$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & x & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & y \end{pmatrix}$	t^2-xt+1, t^2-yt+1 irreducible, $x \neq y$	2	$(q-1)^2q^4(q^2+1)$	$\frac{(q-1)(q-3)}{8}$
$h_{x,y}$	$\begin{pmatrix} 0 & 0 & y & 0 \\ 0 & -xy^{-1} & 0 & y^{-1} \\ 1 & 0 & x & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$	t^2-xt-y irreducible $y \neq -1$	2	$(q-1)q^4(q+1)(q^2+1)$	$\frac{(q-1)^2}{4}$
k_x	$\begin{pmatrix} \frac{x}{2} & \frac{x}{4} & \frac{x^2}{2}-2 & -\frac{1}{4} \\ 0 & \frac{x}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & \frac{x}{2} & \frac{-x}{4x^2-16} \\ 0 & 2-\frac{x^2}{2} & 0 & \frac{x}{2} \end{pmatrix}$	t^2-xt+1 irreducible	1	$(q-1)^2q^3(q+1)(q^2+1)$	$\frac{q-1}{2}$
$m_{x,y}$	$\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & x & 1 & y \\ 1 & 0 & 0 & x \\ 0 & 1 & 0 & 0 \end{pmatrix}$	$t^4-xt^3-yt^2-xt+1$ irreducible	1	$(q-1)^2q^4(q+1)^2$	$\frac{q^2-1}{4}$

Table 4.10: Conjugacy Classes of $\mathrm{Sp}(4,q)$ for q odd Part 2.

4.3 Flag Character Values

Now that we have a list of the conjugacy classes of $\mathrm{Sp}(4,q)$, we can compute the values of the characters of the permutation representations corresponding to the orbits of the action of $\mathrm{Sp}(4,q)$ on flags. If $\mathcal{O}(\mathcal{I}_\lambda)$ is the orbit corresponding to the incidence matrix $\mathcal{I}_\lambda$, recall that we denote the permutation representation of $\mathrm{Sp}(4,q)$ on $\mathcal{O}(\mathcal{I}_\lambda)$ by $\Phi(\mathcal{I}_\lambda)$. We denote the character of $\Phi(\mathcal{I}_\lambda)$ by $\chi(\mathcal{I}_\lambda)$. If $g \in \mathrm{Sp}(4,q)$, then by Lemma 2.1.13, $\chi(\mathcal{I}_\lambda)(g)$ is the number of flags in $\mathcal{O}(\mathcal{I}_\lambda)$ which are fixed by g. Since a flag is fixed by g if and only if all of its subspaces are fixed, it will useful for purposes of these computations to know which subspaces of $\mathbb{F}_q^4$ are fixed by each of the class representatives discussed in the previous section. Any subspace can be represented by a matrix whose columns form a basis for the subspace. Hence via row reduction, we see that any 1 or 2-dimensional subspace (line or plane) can be represented uniquely by one of the following matrices.

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix} \begin{pmatrix}\alpha\\1\\0\\0\end{pmatrix} \begin{pmatrix}\alpha\\\beta\\1\\0\end{pmatrix} \begin{pmatrix}\alpha\\\beta\\\gamma\\1\end{pmatrix} \begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix} \begin{pmatrix}1&0\\0&\alpha\\0&1\\0&0\end{pmatrix} \begin{pmatrix}\alpha&\beta\\1&0\\0&1\\0&0\end{pmatrix} \begin{pmatrix}1&0\\0&\alpha\\0&\beta\\0&1\end{pmatrix} \begin{pmatrix}\alpha&\beta\\1&0\\0&\gamma\\0&1\end{pmatrix} \begin{pmatrix}\alpha&\beta\\\gamma&\delta\\1&0\\0&1\end{pmatrix}$$

A 3-dimensional subspace is fixed if and only if its 1-dimensional orthogonal complement is fixed, so it is unnecessary to list fixed 3-dimensional subspaces. Tables 4.11 and 4.12 list, for each class representative g, the lines and planes which are fixed by g. In all cases, the parameters used may vary over $\mathbb{F}_q$. Below each plane, we list the number of fixed lines it contains.

In reference to the construction of Tables 4.11 and 4.12, we note that determining the lines fixed by any g simply amounts to determining the eigenspaces of g according to the usual methods and listing all the lines they contain. Determining the planes fixed by g is somewhat more difficult; usually this can be done by considering the eigenspace structure of g and using the fact that the minimal and characteristic polynomials of the restriction of g to a fixed plane must divide the minimal and characteristic polynomials of g. If necessary, the fixed planes can be computed by multiplying the six different types of 4×2 matrices corresponding to the planes by g, and determining under what conditions on the parameters the resulting matrix represents the same plane as the original matrix.

g	Lines fixed by g	Planes fixed by g
$a_{\pm1,\pm1}$	All	All
$a_{x,x}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}\alpha\\0\\1\\0\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\begin{pmatrix}0\\\alpha\\0\\1\end{pmatrix}$	$\underset{q+1}{\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}}$ $\underset{q+1}{\begin{pmatrix}0&0\\1&0\\0&0\\0&1\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&\alpha\\0&0\\0&1\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}0&\alpha\\1&0\\0&1\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}\alpha&0\\0&\beta\\1&0\\0&1\end{pmatrix}}$
$a_{x,\pm1}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\begin{pmatrix}0\\0\\\alpha\\1\end{pmatrix}$	$\underset{q+1}{\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&\alpha\\0&1\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}0&0\\1&0\\0&1\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}0&0\\1&0\\0&\alpha\\0&1\end{pmatrix}}$
$a_{1,-1}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}\alpha\\1\\0\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\begin{pmatrix}0\\0\\\alpha\\1\end{pmatrix}$	$\underset{q+1}{\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&\alpha\\0&1\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}\alpha&0\\1&0\\0&1\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}\alpha&0\\1&0\\0&\beta\\0&1\end{pmatrix}}$ $\underset{q+1}{\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}}$
$a_{x,y}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$	$\underset{2}{\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&0\\0&1\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}0&0\\1&0\\0&1\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}0&0\\1&0\\0&0\\0&1\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}}$
$b^a_{\pm1,\pm1}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}\alpha\\0\\1\\0\end{pmatrix}\begin{pmatrix}\alpha\\0\\\beta\\1\end{pmatrix}$	$\underset{1}{\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}}$ $\underset{\substack{1 \text{ if } \alpha\neq0\\ q+1 \text{ if } \alpha=0}}{\begin{pmatrix}1&0\\0&\alpha\\0&1\\0&0\end{pmatrix}}$ $\underset{\substack{1 \text{ if } \alpha\neq0\\ q+1 \text{ if } \alpha=0}}{\begin{pmatrix}1&0\\0&\alpha\\0&\beta\\0&1\end{pmatrix}}$ $\underset{q+1}{\begin{pmatrix}\alpha&\beta\\0&0\\1&0\\0&1\end{pmatrix}}$
$b^a_{\pm1,\mp1}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\begin{pmatrix}0\\0\\\alpha\\1\end{pmatrix}$	$\underset{1}{\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&\alpha\\0&1\end{pmatrix}}$ $\underset{q+1}{\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}}$
$b^a_{\pm1,x}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$	$\underset{1}{\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&0\\0&1\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}}$
$c^a_{\pm1,\pm1}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}\alpha\\0\\1\\0\end{pmatrix}$	$\underset{1}{\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}}$ $\underset{\substack{1 \text{ if } \alpha\neq0\\ q+1 \text{ if } \alpha=0}}{\begin{pmatrix}1&0\\0&\alpha\\0&1\\0&0\end{pmatrix}}$ $\underset{1}{\begin{pmatrix}a\alpha&\beta\\0&\alpha\\1&0\\0&1\end{pmatrix}}$
$c^{a,b}_{1,-1}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$	$\underset{1}{\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}}$ $\underset{2}{\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}}$ $\underset{1}{\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}}$

Table 4.11: Lines and Planes Fixed by Elements of $\mathrm{Sp}(4, q)$ Part 1.

g	Lines fixed by g	Planes fixed by g
$c_{x,x^{-1}}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ $\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}$ 1 $\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}$ 2 $\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}$ 1
$d^a_{\pm1}$	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\\0&0\\0&1\end{pmatrix}$ 1
$e_{x,\pm1}$	$\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$ $\begin{pmatrix}0\\0\\\alpha\\1\end{pmatrix}$	$\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}$ 0 $\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}$ $q+1$
$e_{x,y}$	$\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$ $\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$	$\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}$ 0 $\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}$ 2
$f^a_{x,\pm1}$	$\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}$ 0 $\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}$ 1
$g_{x,x}$	–	$\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}$ 0 $\begin{pmatrix}\alpha&-\beta\\\beta&\alpha+\beta x\\1&0\\0&1\end{pmatrix}$ 0
$g_{x,y}$	–	$\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}$ 0 $\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}$ 0
$h_{x,y}$	–	$\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}$ 0 $\begin{pmatrix}0&0\\1&0\\0&0\\0&1\end{pmatrix}$ 0
k_x	–	$\begin{pmatrix}1&0\\0&0\\0&1\\0&0\end{pmatrix}$ 0
$m_{x,y}$	–	–

Table 4.12: Lines and Planes Fixed by Elements of $\mathrm{Sp}(4,q)$ Part 2.

With Tables 4.11 and 4.12 at hand, we now discuss the computation of the values of the various flag characters. The computation of $\chi\left((1)_{\{3,1\}}\right)$ is particularly easy, since for any $g \in \mathrm{Sp}(4,q)$, $\chi\left((1)_{\{3,1\}}\right)(g)$ is simply the number of lines fixed by g, and this can be read off the tables. For example, since $b^a_{\pm1,\pm1}$ fixes 1 line of type $(1,0,0,0)^T$, q lines of type $(\alpha,1,0,0)^T$, and q^2 lines of type $(\alpha,0,\beta,1)^T$, $\chi\left((1)_{\{3,1\}}\right)(b^a_{\pm1,\pm1}) = q^2+q+1$.

The computation of $\chi\left((2)_{\{2,2\}}\right)$ and $\chi\left((0)_{\{2,2\}}\right)$ is similar, except that we must distinguish between two different types of planes. Flags of types $(2)_{\{2,2\}}$ and $(0)_{\{2,2\}}$ correspond to planes with radicals of dimensions 2 and 0 respectively. The dimension of the radical of a plane can be determined by computing the value of B on the two basis vectors in a matrix representing the plane. If B takes the value 0 on these basis vectors, the radical of the plane is equal to the plane, and thus has dimension 2, while if B is non-zero on these basis vectors, the radical of the plane has dimension 0. (In computing B, we must recall that in Tables 4.11 and 4.12, we are using the version of $\mathrm{Sp}(4,q)$ for which $J = \widehat{B}$.) For example, in the case of $a_{x,x}$, the fixed planes represented by

$$\begin{pmatrix} 1&0\\0&0\\0&1\\0&0 \end{pmatrix}, \begin{pmatrix} 0&0\\1&0\\0&0\\0&1 \end{pmatrix}, \begin{pmatrix} 1&0\\0&0\\0&0\\0&1 \end{pmatrix}, \begin{pmatrix} 0&0\\1&0\\0&1\\0&0 \end{pmatrix}, \text{ and } \begin{pmatrix} \alpha&0\\0&-\alpha^{-1}\\1&0\\0&1 \end{pmatrix} \quad \alpha \neq 0$$

are of type $(2)_{\{2,2\}}$, while all the remaining fixed planes are of type $(0)_{\{2,2\}}$. Hence $\chi\left((2)_{\{2,2\}}\right)(a_{x,x}) = q+3$, while $\chi\left((0)_{\{2,2\}}\right)(a_{x,x}) = (q^2+2q+3)-(q+3) = q^2+q$. In computing these characters, a slight complication occurs for the elements $c^a_{\pm1,\pm1}$. These elements fix all the planes represented by matrices of the form

$$\begin{pmatrix} a\alpha&\beta\\0&\alpha\\1&0\\0&1 \end{pmatrix}.$$

The value of B on the columns of this matrix is $a\alpha^2+1$, which is 0 if and only if $\alpha^2 = -a^{-1}$. Hence if $-a^{-1}$ is a square, there are two values of α for which such a plane is of type $(2)_{\{2,2\}}$, while if $-a^{-1}$ is not a square, all such planes are of type $(0)_{\{2,2\}}$. Since a is either 1 or ϵ, a non-square element of $\mathbb{F}_q^*$, this amounts to determining whether -1 is a square, which depends on whether q is congruent to 1 or 3 modulo 4. Taking into account the other types of fixed planes as well, we get the character values for $c^a_{\pm1,\pm1}$ listed in Table 4.13.

a	$\chi\left((2)_{\{2,2\}}\right)(c^a_{\pm1,\pm1})$	$\chi\left((0)_{\{2,2\}}\right)(c^a_{\pm1,\pm1})$
1	$2q+1$ if $q\equiv1 \mod 4$ 1 if $q\equiv3 \mod 4$	q^2-q if $q\equiv1 \mod 4$ q^2+q if $q\equiv3 \mod 4$
ϵ	1 if $q\equiv1 \mod 4$ $2q+1$ if $q\equiv3 \mod 4$	q^2+q if $q\equiv1 \mod 4$ q^2-q if $q\equiv3 \mod 4$

Table 4.13: Character Values Depending on the Value of $q \mod 4$.

The only other complication occurs for the elements of the form $g_{x,x}$. A fixed plane of the form

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha+\beta x \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is of type $(2)_{\{2,2\}}$ if and only if $\alpha^2 + \alpha\beta x + \beta^2 = -1$. Since this is equivalent to $N(\alpha + \beta\xi) = -1$, where $N : \mathbb{F}_{q^2}^* \to \mathbb{F}_q^*$ is the usual homomorphism, there are $q + 1$ planes of this form which are of type $(2)_{\{2,2\}}$.

Once it has been determined which planes are of type $(2)_{\{2,2\}}$ and which are of type $(0)_{\{2,2\}}$, the computation of the characters $\chi\left(\left(\begin{smallmatrix} 2 & 1 \\ 2 & 1 \end{smallmatrix}\right)_{\{2,1,1\}}\right)$ and $\chi\left(\left(\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right)_{\{2,1,1\}}\right)$ is straightforward. For a flag of type $\left(\begin{smallmatrix} 2 & 1 \\ 2 & 1 \end{smallmatrix}\right)_{\{2,1,1\}}$ consists of a plane of type $(2)_{\{2,2\}}$ containing a line, while a flag of type $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right)_{\{2,1,1\}}$ consists of a plane of type $(0)_{\{2,2\}}$ containing a line. Hence for any $g \in \mathrm{Sp}(4, q)$, $\chi\left(\left(\begin{smallmatrix} 2 & 1 \\ 2 & 1 \end{smallmatrix}\right)_{\{2,1,1\}}\right)(g)$ is the sum over all the planes of type $(2)_{\{2,2\}}$ fixed by g of the number of lines which are contained in the plane and fixed by g. Since the number of fixed lines contained in each fixed plane is recorded in Tables 4.11 and 4.12, this sum is easily computed. For example, in the case of $a_{x,x}$, the fixed planes

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

of type $(2)_{\{2,2\}}$ each contain $q + 1$ fixed lines, while the other fixed planes of type $(2)_{\{2,2\}}$, namely

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ and } \begin{pmatrix} \alpha & 0 \\ 0 & -\alpha^{-1} \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \alpha \neq 0,$$

each contain 2 fixed lines. Thus $\chi\left(\left(\begin{smallmatrix} 2 & 1 \\ 2 & 1 \end{smallmatrix}\right)_{\{2,1,1\}}\right)(a_{x,x}) = (q+1) + (q+1) + 2 + 2 + 2(q-1) = 4q + 4$. The computation of $\chi\left(\left(\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right)_{\{2,1,1\}}\right)$ is similar.

In computing the values of $\chi\left(\left(\begin{smallmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{smallmatrix}\right)_{\{1,1,1,1\}}\right) = \chi\left(\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)_{\{1,2,1\}}\right)$, it is useful to translate information about 3-dimensional subspaces to information about their 1-dimensional orthogonal complements. By definition, a flag of type $\left(\begin{smallmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{smallmatrix}\right)_{\{1,1,1,1\}}$ consists of a 3-dimensional subspace containing a plane of type $(2)_{\{2,2\}}$ containing a line which is not the orthogonal complement of the 3-dimensional subspace. But since a plane of type $(2)_{\{2,2\}}$ is its own orthogonal complement, a 3-dimensional subspace contains a plane of type $(2)_{\{2,2\}}$ if and only if the plane contains the orthogonal complement of the 3-dimensional subspace. Thus a flag of type $\left(\begin{smallmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{smallmatrix}\right)_{\{1,1,1,1\}}$ corresponds naturally to a plane of type $(2)_{\{2,2\}}$ containing an ordered pair of distinct lines. Hence for $g \in \mathrm{Sp}(4, q)$, we can compute $\chi\left(\left(\begin{smallmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{smallmatrix}\right)_{\{1,1,1,1\}}\right)(g)$ by taking the sum over all the planes of type $(2)_{\{2,2\}}$ fixed by g of $l(l - 1)$, where l is the number of fixed lines contained in the plane. For example, referring to the example for $\chi\left(\left(\begin{smallmatrix}2&1\\2&1\end{smallmatrix}\right)_{\{2,1,1\}}\right)$, in which the number of fixed lines contained in each fixed plane was noted for $a_{x,x}$, we see that $\chi\left(\left(\begin{smallmatrix}1&1&0\\2&2&1\\2&2&1\end{smallmatrix}\right)_{\{1,1,1,1\}}\right)(a_{x,x}) = (q+1)q+(q+1)q+2+2+2(q-1) = 2q^2+4q+2$. The computation of $\chi\left(\left(\begin{smallmatrix}1&0&0\\1&0&0\\2&1&1\end{smallmatrix}\right)_{\{1,1,1,1\}}\right)$ is similar, but slightly more complicated. By definition, a flag of type $\left(\begin{smallmatrix}1&0&0\\1&0&0\\2&1&1\end{smallmatrix}\right)_{\{1,1,1,1\}}$ consists of a 3-dimensional subspace containing a plane of type $(0)_{\{2,2\}}$ containing a line, but on taking the orthogonal complement of the 3-dimensional subspace, we can translate this to a plane of type $(0)_{\{2,2\}}$ containing a line together with a line contained in the orthogonal complement of the plane. Thus to compute $\chi\left(\left(\begin{smallmatrix}1&0&0\\1&0&0\\2&1&1\end{smallmatrix}\right)_{\{1,1,1,1\}}\right)(g)$, we take the sum over all planes of type $(0)_{\{2,2\}}$ fixed by g of the number of fixed lines in the plane times the number of fixed lines in the orthogonal complement of the plane. For example, $a_{x,\pm1}$ fixes two planes of type $(0)_{\{2,2\}}$, namely

$$\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix} \text{ and } \begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix},$$

and these planes are orthogonal complements of each other. Since the first contains $q+1$ fixed lines, while the second contains 2 fixed lines,

$$\chi\left(\left(\begin{smallmatrix}1&0&0\\1&0&0\\2&1&1\end{smallmatrix}\right)_{\{1,1,1,1\}}\right)(a_{x,\pm1}) = (q+1)2+2(q+1) = 4q+4.$$

Tables 4.14 and 4.15 give the values of all the flag characters of $\mathrm{Sp}(4,q)$ on the various conjugacy classes. When two numbers are given in a box, the top number corresponds to $q \equiv 1 \mod 4$, while the bottom number corresponds to $q \equiv 3 \mod 4$.

From these tables, we can see that although the characters $\chi\left((1)_{\{3,1\}}\right)$ and $\chi\left((2)_{\{2,2\}}\right)$ have the same degree, they take different values on most conjugacy classes, and therefore, as was claimed in the previous chapter, the corresponding representations are not isomorphic.

Character	$a_{\pm 1,\pm 1}$	$a_{x,x}$	$a_{x,\pm 1}$	$a_{1,-1}$	$a_{x,y}$	$b^a_{\pm 1,\pm 1}$	$b^a_{\pm 1,\mp 1}$	$b^a_{\pm 1,x}$
$\chi\left(\emptyset_{\{4\}}\right)$	1	1	1	1	1	1	1	1
$\chi\left((1)_{\{3,1\}}\right)$	q^3+q^2+q+1	$2q+2$	$q+3$	$2q+2$	4	q^2+q+1	$q+2$	3
$\chi\left((2)_{\{2,2\}}\right)$	q^3+q^2+q+1	$q+3$	$2q+2$	q^2+2q+1	4	$q+1$	$q+1$	2
$\chi\left((0)_{\{2,2\}}\right)$	q^4+q^2	q^2+q	2	2	2	$2q^2$	2	2
$\chi\left(\left(\begin{smallmatrix}2&1\\2&1\end{smallmatrix}\right)_{\{2,1,1\}}\right)$	$q^4+2q^3+2q^2+2q+1$	$4q+4$	$4q+4$	$2q^2+4q+2$	8	q^2+2q+1	$2q+2$	4
$\chi\left(\left(\begin{smallmatrix}0&0\\1&1\end{smallmatrix}\right)_{\{2,1,1\}}\right)$	$q^5+q^4+q^3+q^2$	$2q^2+2q$	$q+3$	$2q+2$	4	q^3+2q^2	$q+2$	3
$\chi\left(\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)_{\{1,2,1\}}\right)$	$q^5+2q^4+2q^3+2q^2+q$	$2q^2+4q+2$	$4q+4$	$2q^2+4q+2$	8	q^3+2q^2+q	$2q+2$	4
$\chi\left(\left(\begin{smallmatrix}1&0&0\\1&0&0\\2&1&1\end{smallmatrix}\right)_{\{1,1,1,1\}}\right)$	$q^6+2q^5+2q^4+2q^3+q^2$	$4q^2+4q$	$4q+4$	$2q^2+4q+2$	8	$2q^3+2q^2$	$2q+2$	4

Table 4.14: Values of Flag Characters of $\mathrm{Sp}(4,q)$ for q odd Part 1.

Character	$c_{\pm1,\pm1}$	$c^{\epsilon}_{\pm1,\pm1}$	$c^{a,b}_{1,-1}$	$c_{x,x^{-1}}$	$d^{a}_{\pm1}$	$e_{x,\pm1}$	$e_{x,y}$	$f^{a}_{x,\pm1}$	$g_{x,x}$	$g_{x,y}$	$h_{x,y}$	k_x	$m_{x,y}$
$\chi\left(\emptyset_{\{4\}}\right)$	1	1	1	1	1	1	1	1	1	1	1	1	1
$\chi\left((1)_{\{3,1\}}\right)$	$q+1$	$q+1$	2	2	1	$q+1$	2	1	0	0	0	0	0
$\chi\left((2)_{\{2,2\}}\right)$	$\begin{smallmatrix}2q+1\\1\end{smallmatrix}$	$\begin{smallmatrix}1\\2q+1\end{smallmatrix}$	1	3	1	0	0	0	$q+1$	0	2	1	0
$\chi\left((0)_{\{2,2\}}\right)$	$\begin{smallmatrix}q^2-q\\q^2+q\end{smallmatrix}$	$\begin{smallmatrix}q^2+q\\q^2-q\end{smallmatrix}$	2	0	0	2	2	2	q^2-q	2	0	0	0
$\chi\left(\left(\begin{smallmatrix}2&1\\2&1\end{smallmatrix}\right)_{\{2,1,1\}}\right)$	$\begin{smallmatrix}3q+1\\q+1\end{smallmatrix}$	$\begin{smallmatrix}q+1\\3q+1\end{smallmatrix}$	2	4	1	0	0	0	0	0	0	0	0
$\chi\left(\left(\begin{smallmatrix}0&0\\1&1\end{smallmatrix}\right)_{\{2,1,1\}}\right)$	$\begin{smallmatrix}q^2-q\\q^2+q\end{smallmatrix}$	$\begin{smallmatrix}q^2+q\\q^2-q\end{smallmatrix}$	2	0	0	$q+1$	2	1	0	0	0	0	0
$\chi\left(\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)_{\{1,2,1\}}\right)$	q^2+q	q^2+q	2	2	0	0	0	0	0	0	0	0	0
$\chi\left(\left(\begin{smallmatrix}1&0&0\\1&0&0\\2&1&1\end{smallmatrix}\right)_{\{1,1,1,1\}}\right)$	$\begin{smallmatrix}q^2-q\\q^2+q\end{smallmatrix}$	$\begin{smallmatrix}q^2+q\\q^2-q\end{smallmatrix}$	2	0	0	0	0	0	0	0	0	0	0

Table 4.15: Values of Flag Characters of $\mathrm{Sp}(4,q)$ for q odd Part 2.

4.4 An Irreducible Character of $\mathrm{Sp}(4, q)$

In this section we will show how to extract an irreducible character of $\mathrm{Sp}(4, q)$ from the characters computed in the last section. Since we know the values of the various flag characters on the classes of $\mathrm{Sp}(4, q)$, as well as the size of those classes and the number of each type of class, all as functions of q, we can compute the inner product of any two flag characters as a function of q. This is best done by means of a computer algebra system. Table 4.16 gives the inner product of any two of these characters. For brevity, our indexing uses only the incidence matrix to which the character corresponds.

	$\emptyset$	(1)	(2)	(0)	$\left(\begin{smallmatrix}2&1\\2&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\1&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0&0\\1&0&0\\2&1&1\end{smallmatrix}\right)$
$\emptyset$	1	1	1	1	1	1	1	1
(1)	1	3	2	3	4	5	6	8
(2)	1	2	3	2	4	3	5	6
(0)	1	3	2	$q+2$	4	$q+4$	$q+5$	$2q+6$
$\left(\begin{smallmatrix}2&1\\2&1\end{smallmatrix}\right)$	1	4	4	4	8	8	12	18
$\left(\begin{smallmatrix}0&0\\1&1\end{smallmatrix}\right)$	1	5	3	$q+4$	8	$2q+9$	$2q+12$	$6q+16$
$\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)$	1	6	5	$q+5$	12	$2q+12$	$2q+18$	$6q+26$
$\left(\begin{smallmatrix}1&0&0\\1&0&0\\2&1&1\end{smallmatrix}\right)$	1	8	6	$2q+6$	18	$6q+16$	$6q+26$	$2q^2+16q+36$

Table 4.16: Inner Products of Flag Characters of $\mathrm{Sp}(4, q)$.

Examining the diagonal entries of this table, we see that for most of the flag characters χ, $\langle\chi, \chi\rangle$ is quite large, and often depends on q. This means that most of these characters have either a large number of irreducible constituents, or a few irreducible constituents repeated numerous times. In particular, $\langle\chi, \chi\rangle$ is never equal to 2, so we cannot obtain any irreducible characters simply by subtracting the trivial character from a flag character. Thus it is unlikely that we will be able to obtain a significant number of irreducible characters of $\mathrm{Sp}(4, q)$ simply by taking linear combinations of flag characters, as we could in the case of $\mathrm{GL}(n, q)$. Nevertheless, we can obtain one irreducible character by this process. To see how, let us focus on the characters χ for which $\langle\chi, \chi\rangle$ does not depend on q, namely $\rho_1 = \chi\left(\emptyset_{\{4\}}\right)$, $\rho_2 = \chi\left((1)_{\{3,1\}}\right)$, $\rho_3 = \chi\left((2)_{\{2,2\}}\right)$, and $\rho_4 = \chi\left(\left(\begin{smallmatrix}2&1\\2&1\end{smallmatrix}\right)_{\{2,1,1\}}\right)$.

By definition, $\rho_1 = \chi_1$, the trivial character of $\mathrm{Sp}(4, q)$. Since $\langle\rho_2, \rho_2\rangle = 3$, and since $3 = 1^2 + 1^2 + 1^2$ is the only way to express 3 as a sum of squares, ρ_2 is the sum of 3 distinct irreducible characters. Since $\langle\chi_1, \rho_2\rangle = 1$, one of these is χ_1; we label the other 2 as χ_2 and χ_3. Thus

$$\rho_2 = \chi_1 + \chi_2 + \chi_3.$$

Likewise, ρ_3 is the sum of 3 distinct irreducible characters, one of which is χ_1.

Since $\langle\rho_2, \rho_3\rangle = 2$, one of the other 2 irreducible constituents of ρ_3 is also a constituent of ρ_2. Without loss of generality, we can assume it is χ_2. Thus there is an additional irreducible character χ_4 such that

$$\rho_3 = \chi_1 + \chi_2 + \chi_4.$$

Since $\langle\rho_4, \rho_4\rangle = 8$, an irreducible constituent of ρ_4 can appear at most twice. Since $\langle\chi_1, \rho_4\rangle = 1$, χ_1 appears only once, but since $\langle\rho_2, \rho_4\rangle = \langle\rho_3, \rho_4\rangle = 4$, at least one appears twice. The only decomposition of 8 as a sum of squares which involves both 1 and 2 is $8 = 1^2 + 1^2 + 1^2 + 1^2 + 2^2$, so ρ_4 is the sum of 5 distinct irreducible characters, of which exactly one appears twice. The character which appears twice must be a constituent of both ρ_2 and ρ_3 to produce the listed inner products, so it must be χ_2. In order to produce the listed inner products, χ_3 and χ_4 must also be constituents of ρ_4. Thus there is one additional irreducible character χ_5 such that

$$\rho_4 = \chi_1 + 2\chi_2 + \chi_3 + \chi_4 + \chi_5.$$

Based on these equations, we see that

$$\chi_5 = \rho_4 - \rho_3 - \rho_2 + \rho_1.$$

Since the values of the ρ_i are already known, we can easily compute the values of χ_5, as shown in Table 4.17.

g	$a_{\pm1,\pm1}$	$a_{x,x}$	$a_{x,\pm1}$	$a_{1,-1}$	$a_{x,y}$	$b^a_{\pm1,\pm1}$	$b^a_{\pm1,\mp1}$	$b^a_{\pm1,x}$	$c^a_{\pm1,\pm1}$	$c^{a,b}_{1,-1}$
$\chi_5(g)$	q^4	q	q	q^2	1	0	0	0	0	0
g	$c_{x,x^{-1}}$	$d^a_{\pm1}$	$e_{x,\pm1}$	$e_{x,y}$	$f^a_{x,\pm1}$	$g_{x,x}$	$g_{x,y}$	$h_{x,y}$	k_x	$m_{x,y}$
$\chi_5(g)$	0	0	$-q$	-1	0	$-q$	1	-1	0	1

Table 4.17: Values of the Character χ_5.

Using the values given in the table, it is straightforward to check that $\langle\chi_5, \chi_5\rangle = 1$, so that χ_5 is, indeed, an irreducible character of $\mathrm{Sp}(4, q)$. By examining its values, we see that χ_5 is the Steinberg character of $\mathrm{Sp}(4, q)$. (For a discussion of this character in a much more general context, see Chapter 6 of [1].) We note that the computation of an irreducible character of $\mathrm{Sp}(4, q)$ is nothing new. In fact, Srinivasan computed all the irreducible characters of $\mathrm{Sp}(4, q)$ in [10]. Nevertheless, the results of this section do show that flag characters provide a relatively elementary method of calculating one of these irreducible characters.

Bibliography

[1] Roger W. Carter, *Finite Groups of Lie Type*, John Wiley and Sons, Chichester, 1985.

[2] Leonard E. Dickson, *Canonical forms of quaternary abelian substitutions in an arbitrary Galois field*, Transactions of the American Mathematical Society, **2** (1901), 103–138.

[3] John B. Fraleigh, *A First Course in Abstract Algebra*, Addison Wesley, 2003.

[4] J. A. Green, *The characters of the finite general linear groups*, Transactions of the American Mathematical Society, **80** (1955), 402–447.

[5] Larry C. Grove, *Algebra*, Dover Publications, Inc., Mineola, NY, 1983.

[6] Larry C. Grove, *Classical Groups and Geometric Algebra*, volume 39 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, 2002.

[7] James E. Humphreys, *Linear Algebraic Groups*, volume 21 of *Graduate Texts in Mathematics*, Springer-Verlag, New York, NY, 1975.

[8] I. Martin Isaacs, *Character Theory of Finite Groups*, Dover Publications, Inc., New York, NY, 1976.

[9] G. D. James, *Representations of General Linear Groups*, volume 94 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, 1984.

[10] Bhama Srinivasan, *The Characters of the Finite Symplectic Group Sp(4,q)*, Transactions of the American Mathematical Society, **131** (1968), 488–525.

[11] Robert Steinberg, *A geometric approach to the representations of the full linear group over a Galois field*, Transactions of the American Mathematical Society, **71** (1951), 274–282.

[12] Robert Steinberg, *The representations of GL(3,q), GL(4,q), PGL(3,q), and PGL(4,q)*, Canadian Journal of Mathematics, **3** (1951), 225–235.

[13] Zhe-xian Wan, *Geometry of Classical Groups over Finite Fields*, Studentlitteratur, Lund, 1993.