D. E. Taylor

The Geometry of the Classical Groups

Corrections: April 2021

School of Mathematics and Statistics
The University of Sydney
Australia 2006

To Tristan and Ursula

Preface

This book began as a set of lecture notes for a postgraduate course at the University of Sydney in 1982. The aim was to start with vector spaces, introduce sesquilinear forms, and then study the classical groups (i.e., the special linear, symplectic, unitary and orthogonal groups) along the lines of Artin (1957)†.

In 1983 the manuscript was extensively rewritten while on leave at the University of Oregon and the University of New South Wales. During the rewriting more emphasis was placed on the "buildings" of the groups and their corresponding BN-pairs. However, the buildings and other related geometries are constructed from the vector spaces and sesquilinear forms in keeping with the original approach to the classical groups. In the early parts of the book vector spaces over arbitrary division rings are considered. But after Chapter 7 it is assumed that the vector spaces and groups are defined over fields and that the bilinear forms have isotropic points. In the latter half of the book more attention is given to groups defined over finite fields. The classification of the finite simple groups was completed in 1980 and as a consequence all but 26 of the non-abelian finite simple groups are now known to be either an alternating group or a group of Lie type. If the dimension is large enough, a simple group of Lie type is a classical group. Thus the classical groups play a central rôle in the study of all finite groups. A wealth of information about the finite case can be found in Kleidman and Liebeck (1990).

There is a uniform approach to these groups which begins with complex semisimple Lie algebras. This was discovered by Chevalley (1955) and developed by Steinberg (1959), Tits (1959) and Hertzig (1961). However, in order to make this book more accessible to students with a background in linear algebra and a little group theory, but no Lie theory, I have not used Chevalley's approach. Similarly, even though I have included a chapter on buildings, I have avoided basing the book on the abstract theory of buildings and algebraic groups. The intention has always been to build up the examples from the underlying vector spaces. Moreover, excellent accounts of the other approaches can be found in Brown (1989), Carter (1972), Ronan (1989) and Tits (1974); the present book can be regarded as an introduction to their works via linear algebra.

† References to the bibliography are given in the form "Name (date)".

The first draft of Chapters 1 through 8 was produced in 1982. I am indebted to Cathy Kicinski for typing it, and for typing part of the bibliography. In 1983 I typed the second draft, including the present Chapter 10, on an Apple II computer. Not long after that the files were transferred to a system where Knuth's typesetting program TEX was available. Since then the computers have changed but the typesetting program has remained TEX.

Most of the final draft was produced between January and June 1990 while on leave at the Université du Québec à Montréal. I am especially grateful to Pierre Leroux, André Joyal, François Bergeron and their colleagues at LACIM (Laboratoire de combinatoire et d'informatique mathématique) for a warm and hospitable environment and the encouragement to complete this project. It was there that I decided to take a uniform approach to the orthogonal groups. Originally there were two chapters: one for fields of characteristic 2 and another for all other characteristics. In reviewing these chapters it became apparent that many of the proofs for the groups over fields of characteristic 2 applied generally. On the other hand, if one were willing to exclude the characteristic 2 case it was almost always possible to obtain much shorter proofs. I have opted for uniformity at the expense of brevity.

The last chapter was written at the University of Sydney but it is based on material written in 1978 at the University of Eindhoven and published as an Appendix to Higman (1978).

The book has been greatly improved by many suggestions from the participants in the various courses I have given at Sydney on this material and by the remarks of Bill Kantor and Gary Seitz during my stay at the University of Oregon. I am most grateful to Stephen Glasby, Bob Howlett and Mathew Nicol for their numerous excellent suggestions on improving the content and the style. Special thanks is due to Leanne Rylands for her support during the project and for her careful reading of the many drafts.

Finally, my apologies to Dr Heldermann for the long delay in producing the final copy and much gratitude for his patience.

March 1991 D.E.T.

Contents

1

Groups Acting on Sets

This chapter provides a quick review of the basic notions of permutation group theory needed in the rest of the book. The main result is Iwasawa's criterion for the simplicity of a primitive group. In some chapters more specialized results about permutation groups acting on graphs and designs are needed and these are developed when required. Not everything is proved, but most of the unproved assertions appear again as exercises at the end of the chapter.

Let G be a group which *acts* on a non-empty set Ω. That is, for each element $g \in G$, there is a bijection $\alpha \mapsto g(\alpha)$ of Ω such that

$$\begin{aligned} 1(\alpha) &= \alpha && \text{for all } \alpha \in \Omega, \text{ and} \\ (gh)(\alpha) &= g(h(\alpha)) && \text{for all } \alpha \in \Omega \text{ and all } g, h \in G. \end{aligned}$$

In general, the groups considered in this book consist of linear transformations and usually a linear transformation is written to the *left* of the vector (or subspace) on which it operates. This is the reason for using the notation $g(\alpha)$ rather than α^g as in Wielandt (1964).

For $\Delta \subseteq \Omega$ and $g \in G$, define $g(\Delta) := \{\, g(\alpha) \mid \alpha \in \Delta \,\}$ and call the subgroup $G_\Delta := \{\, g \in G \mid g(\Delta) = \Delta \,\}$ the *stabilizer* of Δ in G. The *pointwise stabilizer* of Δ is the subgroup

$$G(\Delta) := \{\, g \in G \mid g(\alpha) = \alpha \quad \text{for all } \alpha \in \Delta \,\}.$$

If $\Delta = \{\alpha, \ldots, \omega\}$, we write $G_{\alpha, \ldots, \omega}$ instead of $G(\Delta)$.

For $\Delta \subseteq \Omega$ and $g \in G$, the assignment

$$\Delta \mapsto g(\Delta)$$

defines an action of G on the set of all subsets of Ω. The group $G(\Omega)$ is called the *kernel* of the action of G on Ω. If $G(\Omega) = 1$, G is said to act *faithfully* on Ω or it is said to be a *group of permutations* of Ω. Each element $g \in G_\Delta$ induces a permutation g^Δ of Δ. The assignment $g \mapsto g^\Delta$ is a homomorphism with kernel $G(\Delta)$. Thus $G^\Delta := G_\Delta / G(\Delta)$ is (up to isomorphism) the group of permutations that G_Δ induces on Δ.

A subset Δ of Ω is said to be *G-invariant* if $G_\Delta = G$. A minimal (non-empty) G-invariant subset of Ω is called an *orbit*. It is readily seen that Ω is the disjoint union of its orbits and that the orbit which contains $\alpha \in \Omega$ is $\mathrm{orb}_G(\alpha) := \{\, g(\alpha) \mid g \in G \,\}$.

The mapping $g(\alpha) \mapsto gG_\alpha$ is a well-defined bijection between the orbit $\mathrm{orb}_G(\alpha)$ and the set $G : G_\alpha$ of left cosets of G_α in G. In particular, if $\mathrm{orb}_G(\alpha)$ is finite, we have

$$|\,\mathrm{orb}_G(\alpha)| = |G : G_\alpha|.$$

(We use $|X|$ to denote the *size* (or *order*) of the finite set X.)

When G has only one orbit (namely Ω), G is said to be *transitive* (or to act *transitively* on Ω). If G is transitive and $G_\alpha = 1$ for all $\alpha \in \Omega$, we say that G acts *regularly* on Ω.

Suppose that G acts on Ω and that k is a natural number. Then G acts on the set Ω^k of k-tuples of elements of Ω. The action is defined by

$$g(\alpha_1, \alpha_2, \ldots, \alpha_k) := (g(\alpha_1), g(\alpha_2), \ldots, g(\alpha_k)).$$

Let $\Omega^{[k]}$ be the subset of Ω^k consisting of those k-tuples with *distinct* entries from Ω. If G is transitive on $\Omega^{[k]}$ and $k \leq |\Omega|$, then G is said to act *k-transitively* on Ω.

A subset Δ of $\Omega \times \Omega$ is a relation between Ω and itself. The *converse* of the relation Δ is defined to be

$$\Delta\breve{} := \{\, (\alpha, \beta) \mid (\beta, \alpha) \in \Delta \,\}.$$

For $\alpha \in \Omega$ we define

$$\Delta(\alpha) := \{\, \beta \mid (\alpha, \beta) \in \Delta \,\}.$$

If G acts transitively on Ω, then the assignment $\Delta \mapsto \Delta(\alpha)$ induces a bijection between the orbits of G acting on $\Omega \times \Omega$ and the orbits of G_α acting on Ω. The number of orbits of G on $\Omega \times \Omega$ is called the *permutation rank* of G.

A relation $\Delta \subseteq \Omega \times \Omega$ can be represented by the directed graph on Ω in which α is joined to β whenever $(\alpha, \beta) \in \Delta$. If Δ is G-invariant, then G acts as a group of automorphisms of this graph. If $\Delta = \Delta\breve{}$ (i.e., if Δ is *symmetric*), we shall replace each pair of directed edges between two vertices by a single undirected edge.

Suppose that G acts on Ω. The diagonal $I := \{\, (\alpha, \alpha) \mid \alpha \in \Omega \,\}$, and $\Omega \times \Omega$ are G-invariant equivalence relations on Ω. If G is transitive and if I and $\Omega \times \Omega$ are the *only* G-invariant equivalence relations on Ω, G is said to be *primitive*; otherwise G is said to be *imprimitive*.

If G is imprimitive and E is a G-invariant equivalence relation that is neither the diagonal I nor the whole set $\Omega \times \Omega$, the equivalence classes of E are called *blocks of imprimitivity*. Thus a block of imprimitivity is a subset B of Ω such that $|B| > 1$, $B \neq \Omega$ and for all $g \in G$, $g(B) = B$ or $g(B) \cap B = \emptyset$.

1.1 Theorem. *Suppose that G acts transitively on Ω. If $|\Omega| > 1$ and $\alpha \in \Omega$, then G is primitive if and only if G_α is a maximal subgroup of G.*

Proof. Suppose that B is a block of imprimitivity that contains α. Then $G_\alpha \leq G_B \leq G$. Since $|B| > 1$, $G_\alpha \neq G_B$ and since $B \neq \Omega$, $G_B \neq G$. Hence G_α is not maximal.

Conversely, suppose that H is a proper subgroup of G which contains G_α as a proper subgroup. Then $\mathrm{orb}_H(\alpha)$ is a block of imprimitivity for G. (Exercise 1.11.) □

At this point it is possible to give the simplicity criterion of Iwasawa (1941). Recall that a group G is *simple* if 1 and G are the only normal subgroups of G. Also, the *commutator subgroup* (or *derived group*) of a group G is the subgroup

$$G' := \langle\, ghg^{-1}h^{-1} \mid g, h \in G \,\rangle.$$

The element $[g, h] := ghg^{-1}h^{-1}$ is the *commutator* of g and h. Note that if N is a normal subgroup of G, then G/N is abelian if and only if $G' \leq N$.

1.2 Theorem. *Suppose that G acts primitively on the set Ω and that $G = \langle\, gAg^{-1} \mid g \in G \,\rangle$, where A is an abelian normal subgroup of G_α for some $\alpha \in \Omega$. Then*

(i) *if N is a normal subgroup of G, either $N \leq G(\Omega)$ or $G' \leq N$;*

(ii) *if $G = G'$, then $G/G(\Omega)$ is a simple group.*

Proof. (i) Suppose that $N \not\leq G(\Omega)$. Then $N \not\leq G_\alpha$ (see Exercise 1.3 (ii)) and therefore $G = NG_\alpha$ by the maximality of G_α. If $g \in G$, then $g = nh$, where $n \in N$ and $h \in G_\alpha$. Hence $gAg^{-1} = nhAh^{-1}n^{-1} = nAn^{-1} \leq NA$. It follows that $G = NA$ and therefore $G/N = NA/N \simeq A/A \cap N$ is an abelian group. Hence $G' \leq N$ and (i) is proved.

(ii) This follows from (i) applied to $G/G(\Omega)$. □

Given a group G we set

$$G^{(0)} := G, \quad G^{(m)} := G^{(m-1)'}$$

and call $G^{(0)}$, $G^{(1)}$, $G^{(2)}$, ... the *derived series* of G. The group G is *soluble* if $G^{(m)} = 1$ for some m. A theorem similar to the one just proved holds when the subgroup A is assumed to be soluble rather than abelian. The details are left as an exercise.

EXERCISES

1.1 Suppose that G acts on Ω. If $\Delta \subseteq \Omega$, show that G_Δ is a subgroup of G and that $G(\Delta)$ is a normal subgroup of G_Δ.

1.2 If G acts on Ω, show that Ω is the disjoint union of its orbits and that the orbit containing $\alpha \in \Omega$ is $\mathrm{orb}_G(\alpha)$.

1.3 Suppose that G acts on Ω.

(i) Show that $g(\alpha) \mapsto gG_\alpha$ is a well-defined bijection between the orbit $\mathrm{orb}_G(\alpha)$ and the set of left cosets of G_α.

(ii) If $\beta = g(\alpha)$, show that $G_\beta = gG_\alpha g^{-1}$.

1.4 Suppose that G acts on Ω and that N is a normal subgroup of G which acts regularly on Ω. Fix $\alpha \in \Omega$ and show that $n(\alpha) \mapsto n$ defines a bijection between Ω and N. Show that under this bijection the action of G_α on Ω corresponds to the action of G_α on N by conjugation.

1.5 Suppose that G acts faithfully on Ω and that the only G-invariant equivalence relations are $\Omega \times \Omega$ and the diagonal. If G is not transitive, show that $|\Omega| = 2$ and $G = 1$.

1.6 Suppose that G acts transitively on Ω and that H is a subgroup of G. Show that H is transitive on Ω if and only if $G = G_\alpha H$ for any $\alpha \in \Omega$. Taking Ω to be the set of Sylow p-subgroups of a normal subgroup K of G, deduce that $G = KN_G(P)$ for any $P \in \Omega$.

1.7 Show that a non-trivial normal subgroup of a faithful primitive group is transitive.

1.8 If G acts transitively on Ω, show that the assignment $\Delta \mapsto \Delta(\alpha)$ induces a one-to-one correspondence between the orbits of G on $\Omega \times \Omega$ and the orbits of G_α on Ω.

1.9 Suppose that H is a subgroup of a group G and let Ω be the set of cosets gH, $g \in G$. Let G act on Ω by multiplication on the left. (Thus $x \in G$ sends gH to xgH.) Show that the orbits of H on Ω correspond to the double cosets of H in G. That is, each double coset HgH is a union of the elements in an orbit of H.

1.10 Suppose that G acts transitively on Ω and that Δ is a non-diagonal orbit of G on $\Omega \times \Omega$. Show that if the graph associated with Δ is connected and $(\alpha, g(\alpha)) \in \Delta$ for some $g \in G$, then $G = \langle G_\alpha, g \rangle$. Deduce that if the graphs associated with each of the non-diagonal orbits of G on $\Omega \times \Omega$ are connected, then G is primitive. Show that when G is finite the converse holds.

1.11 Fill in the details of the proof of Theorem 1.1. In particular, show that $\mathrm{orb}_H(\alpha)$ is a block of imprimitivity for G.

1.12 Suppose that G acts transitively on Ω and that E is an equivalence relation on $\{1, 2, \ldots, k\}$. Let Ω_E be the set of k-tuples $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ such that $\alpha_i = \alpha_j$ if and only if $(i, j) \in E$. Show that

(*i*) Ω_E is G-invariant.

(*ii*) Ω^k is the disjoint union of all the Ω_E.

(*iii*) G is k-transitive if and only if all the Ω_E are orbits.

1.13 Show that a 2-transitive group is primitive.

1.14 Show that G is k-transitive on Ω if and only if G_α is $(k-1)$-transitive on $\Omega - \{\alpha\}$ and G is transitive on Ω.

1.15 Suppose that G acts faithfully and primitively on Ω and that S is a soluble normal subgroup of G_α such that $G = \langle\, gSg^{-1} \mid g \in G \,\rangle$. If $G = G'$, show that G is simple.

1.16 Suppose that G acts transitively on a set Ω of size n and that G_α has orbits of length 1, k and l on Ω. If G is imprimitive, show that either $k+1$ or $l+1$ divides n.

1.17 Suppose that G acts faithfully on Ω and that H is a subgroup of G. Show that

(*i*) the centralizer $C_G(H) = \{\, g \in G \mid gh = hg \text{ for all } h \in H \,\}$ of H in G acts on the set

$$\{\, \alpha \in \Omega \mid h(\alpha) = \alpha \text{ for all } h \in H \,\},$$

(*ii*) if $C_G(H)$ is transitive, then $H_\alpha = 1$ for all $\alpha \in \Omega$, and

(*iii*) if H is transitive and abelian, then H is regular and $C_G(H) = H$.

2

Affine Geometry

In later chapters we shall restrict ourselves to the study of the structure of groups of linear transformations of vector spaces defined over *fields*. But commutativity (or the lack of it) does not affect the material in this chapter. So until further notice we deal with vector spaces defined over *division rings*. The main result is an affine version of the 'Fundamental Theorem of Projective Geometry'. This will be used in the next chapter to obtain the projective version.

Let V be a finite-dimensional left vector space over a division ring K. An *affine space* based on V is a set $\mathcal{A}$ on which V acts regularly. The elements of $\mathcal{A}$ will be called *points*. The *dimension* of $\mathcal{A}$ is defined to be the vector space dimension $\dim_K V$ of V.

Since V is an additive group we use an additive notation to describe the action of V on $\mathcal{A}$. That is, for each point P of $\mathcal{A}$ and each vector v of V, there is a point $P + v$ in $\mathcal{A}$ such that

(*i*) $P + 0 = P$,

(*ii*) $P + (u + v) = (P + u) + v$, and

(*iii*) given P, Q in $\mathcal{A}$, there is a unique vector $\overrightarrow{PQ}$ in V such that $P + \overrightarrow{PQ} = Q$.

Given points $P_1, P_2, \ldots, P_r$ in $\mathcal{A}$ and elements $a_1, a_2, \ldots, a_r$ in K such that $a_1 + \cdots + a_r = 1$, the point $O + a_1\overrightarrow{OP_1} + \cdots + a_r\overrightarrow{OP_r}$ is independent of the choice of the point O and therefore we may unambiguously refer to it as $a_1P_1 + \cdots + a_rP_r$ (Exercise 2.2).

An *affine subspace* of $\mathcal{A}$ is a set $\mathcal{B}$ of points of the form $O + w$, where O is a point and w runs through a subspace W of V. It is not hard to see that $\mathcal{B}$ is an affine space in its own right based on the vector space W. Thus the dimension of $\mathcal{B}$ is $\dim_K W$. We call W the *direction* of $\mathcal{B}$. If $P_1, P_2, \ldots, P_r$ are points of $\mathcal{B}$, a straightforward calculation shows that $a_1P_1 + a_2P_2 + \cdots + a_rP_r$ is in $\mathcal{B}$ for all choices of $a_1, a_2, \ldots, a_r$ such that $a_1 + a_2 + \cdots + a_r = 1$. Indeed, the smallest affine subspace of $\mathcal{A}$ that contains the given points $P_1, P_2, \ldots, P_r$ consists of the points of the form $a_1P_1 + a_2P_2 + \cdots + a_rP_r$, where $a_1 + a_2 + \cdots + a_r = 1$.

A *frame* of the subspace $\mathcal{B}$ is a set of points $\{P_1, P_2, \ldots, P_r\}$ such that every point of $\mathcal{B}$ can be written uniquely in the form $a_1P_1 + a_2P_2 + \cdots + a_rP_r$, where $a_1 + a_2 + \cdots + a_r = 1$. We leave it as an exercise to show that $\mathcal{B}$ has a frame and that the number of points in any frame of $\mathcal{B}$ is $1 + \dim \mathcal{B}$.

The affine subspaces of dimension 1 are called the *lines* of $\mathcal{A}$ and those of dimension 2 are called *planes*. A set of points is said to be *collinear* if they lie on a line. If P and Q are distinct points, the line through them is written PQ. It consists of the points of the form $aP + (1-a)Q$.

If P and Q belong to a subspace $\mathcal{B}$, then all the points of PQ belong to $\mathcal{B}$. Conversely, if $|K| \neq 2$ and if $\mathcal{B}$ is a non-empty subset of $\mathcal{A}$ with the property that whenever P and Q are distinct points of $\mathcal{B}$ the entire line PQ is contained in $\mathcal{B}$, it can be shown that $\mathcal{B}$ is an affine subspace (Exercise 2.5).

Semilinear Transformations

For the remainder of the chapter we study maps between affine spaces. Suppose that $\mathcal{A}_1$ and $\mathcal{A}_2$ are affine spaces based on the vector spaces V_1 and V_2 over the division rings K_1 and K_2. A *σ-semilinear transformation* from V_1 to V_2 is a pair (f, σ) where

$$\begin{aligned} f : V_1 &\to V_2, \\ \sigma : K_1 &\to K_2 \text{ is an isomorphism of division rings,} \\ f(u+v) &= f(u) + f(v) \text{ and} \\ f(av) &= \sigma(a)f(v) \text{ for all } u, v \in V_1 \text{ and all } a \in K_1. \end{aligned}$$

Often we omit reference to σ and simply refer to 'the semilinear transformation f'. We call σ the isomorphism *associated* with f.

For each choice of points $O_1 \in \mathcal{A}_1$ and $O_2 \in \mathcal{A}_2$, the semilinear transformation f induces a transformation $\varphi : \mathcal{A}_1 \to \mathcal{A}_2$ defined by

$$\varphi(O_1 + v) := O_2 + f(v).$$

It follows directly from this definition that

$$\varphi(a_1P_1 + \cdots + a_rP_r) = \sigma(a_1)\varphi(P_1) + \cdots + \sigma(a_r)\varphi(P_r).$$

In addition, φ takes collinear points to collinear points and maps each affine subspace of $\mathcal{A}_1$ onto an affine subspace of $\mathcal{A}_2$.

These considerations motivate the following definition. If $\mathcal{A}_1$ and $\mathcal{A}_2$ are affine spaces of the *same* dimension, an *affine isomorphism* is a bijection $\varphi : \mathcal{A}_1 \to \mathcal{A}_2$ that sends each affine subspace of $\mathcal{A}_1$ onto an affine subspace of $\mathcal{A}_2$. (It is tempting to define affine transformations between affine spaces of arbitrary dimensions but then one must be careful to avoid certain pathological examples. See Exercise 2.14.)

2.1 Lemma. *Suppose that $\mathcal{A}_1$ and $\mathcal{A}_2$ are affine spaces of the same dimension over division rings K_1 and K_2, $|K_1| \neq 2$. Suppose that $\varphi : \mathcal{A}_1 \to \mathcal{A}_2$ is a bijection such that whenever P, Q and R are collinear points of $\mathcal{A}_1$, then $\varphi(P)$, $\varphi(Q)$ and $\varphi(R)$ are collinear points of $\mathcal{A}_2$. Then φ is an affine isomorphism and $\dim \varphi(\mathcal{B}) = \dim \mathcal{B}$ for every affine subspace $\mathcal{B}$ of $\mathcal{A}_1$.*

Proof. Given a subspace $\mathcal{B}$, choose subspaces

$$\mathcal{B}_0 \subset \mathcal{B}_1 \subset \ldots \subset \mathcal{B}_n = \mathcal{A}_1$$

so that $\dim \mathcal{B}_i = i$ $(i = 0, 1, \ldots, n)$ and $\mathcal{B} = \mathcal{B}_k$ for some k. Let $\mathcal{C}_i$ be the smallest subspace of $\mathcal{A}_2$ that contains $\varphi(\mathcal{B}_i)$. Since $|K_1| \neq 2$, we can choose distinct points P, Q of $\mathcal{B}_{i+1} \backslash \mathcal{B}_i$ such that PQ contains a point of $\mathcal{B}_i$. Then every point of $\mathcal{B}_{i+1}$ is either on a line through P and a point of $\mathcal{B}_i$ or on a line through Q and a point of $\mathcal{B}_i$. It follows that $\dim \mathcal{C}_{i+1} \leq 1 + \dim \mathcal{C}_i$ and hence $\dim \mathcal{C}_i \leq i$ for all i. But $\dim \mathcal{C}_n = n$ and consequently $\dim \mathcal{C}_i = i$ for all i. If $\varphi(\mathcal{B})$ were not a subspace we could choose $\mathcal{B}_{k+1}$ so that $\varphi(\mathcal{B}_{k+1}) \subseteq \mathcal{C}_k$, contradicting $\dim \mathcal{C}_{k+1} = k + 1$. Thus $\varphi(\mathcal{B})$ is a subspace of $\mathcal{A}_2$ of the same dimension as $\mathcal{B}$. □

This lemma implies that the inverse of an affine isomorphism is also an affine isomorphism. (The case of affine spaces over the field of two elements needs to be considered separately.)

In preparation for the main theorem we need to consider the effect of an affine isomorphism on parallel lines. Two lines are *parallel* if they coincide or if they lie in the same plane and have no point in common. Thus an affine isomorphism takes parallel lines to parallel lines.

A 4-tuple of points (P, Q, R, S), no three of which are collinear, is called a *parallelogram* if PQ is parallel to RS and QR is parallel to SP.

2.2 Lemma. *If (P, Q, R, S) is a parallelogram, then $\overrightarrow{PQ} + \overrightarrow{RS} = 0$ and $\overrightarrow{QR} + \overrightarrow{SP} = 0$.*

Proof. Exercise 2.9. □

The next result is an affine version of the 'Fundamental Theorem of Projective Geometry'.

2.3 Theorem. *Suppose that $\mathcal{A}_1$ and $\mathcal{A}_2$ are affine spaces of dimension $n \geq 2$. If $\varphi : \mathcal{A}_1 \to \mathcal{A}_2$ is an affine isomorphism, there is a unique (invertible) semilinear transformation $\varphi_* : V_1 \to V_2$ such that $\varphi(P + v) = \varphi(P) + \varphi_*(v)$ for all $P \in \mathcal{A}_1$ and $v \in V_1$.*

Proof. Choose $P \in \mathcal{A}_1$ and for $v \in V_1$ define $\varphi_*(v)$ to be the (unique) element of V_2 such that $\varphi(P+v) = \varphi(P) + \varphi_*(v)$. We shall show that $\varphi_*(v)$ does not depend on the choice of P. In any case, if $v = 0$, then $\varphi_*(v) = 0$, and so we may suppose $v \neq 0$ from now on.

Suppose that Q is a point not on the line through P and $P + v$. Then $(P, Q, Q+v, P+v)$ is a parallelogram and so $(\varphi(P), \varphi(Q), \varphi(Q+v), \varphi(P+v))$ is a parallelogram also. It follows from Lemma 2.2 that

$$\varphi(Q + v) = \varphi(Q) + \varphi_*(v).$$

If R is any point collinear with P and $P + v$, the same argument with Q in place of P shows that

$$\varphi(R + v) = \varphi(R) + \varphi_*(v).$$

Thus $\varphi_*(v)$ is well-defined.

For $u, v \in V_1$ we have $\varphi(P)+\varphi_*(u)+\varphi_*(v) = \varphi(P+u)+\varphi_*(v) = \varphi(P+u+v)$ and hence $\varphi_*(u+v) = \varphi_*(u)+\varphi_*(v)$. To complete the proof that $\varphi_* : V_1 \to V_2$ is semilinear we must construct an isomorphism $\sigma : K_1 \to K_2$.

For $a \in K_1$, the points $P, P + v$ and $P + av$ are collinear and therefore the same is true of $\varphi(P)$, $\varphi(P) + \varphi_*(v)$ and $\varphi(P) + \varphi_*(av)$. It follows that

$$\varphi_*(av) = \sigma_v(a)\varphi_*(v) \text{ for some } \sigma_v(a) \in K_2.$$

Suppose that u and v are linearly independent. Then $P, P+u$ and $P+v$ are not collinear and therefore neither are $\varphi(P)$, $\varphi(P)+\varphi_*(u)$ and $\varphi(P)+\varphi_*(v)$. It follows that $\varphi_*(u)$ and $\varphi_*(v)$ are linearly independent. Consequently we have $\sigma_{u+v}(a)\varphi_*(u + v) = \varphi_*(au + av)$ and hence

$$\sigma_{u+v}(a)(\varphi_*(u) + \varphi_*(v)) = \sigma_u(a)\varphi_*(u) + \sigma_v(a)\varphi_*(v).$$

Therefore $\sigma_u(a) = \sigma_{u+v}(a) = \sigma_v(a)$. If w is a multiple of v, then $\sigma_w(a) = \sigma_u(a)$ and hence $\sigma_w(a) = \sigma_v(a)$. Thus $\sigma_v(a)$ does not depend on v and we may write $\varphi_*(av) = \sigma(a)\varphi_*(v)$. It is very easy to check that $\sigma : K_1 \to K_2$ is the required isomorphism. This completes the proof. □

The semilinear transformation φ_* is the *derivative* of φ.

The Affine Group

Let us return to the affine space $\mathcal{A}$ based on the vector space V over the division ring K. The collection of all affine isomorphisms from $\mathcal{A}$ to itself forms the *affine group* $\mathrm{Aff}(\mathcal{A})$. Since there is only one affine space of a given dimension (up to isomorphism), we may also denote the affine group by $\mathrm{Aff}(n, K)$, where n is the dimension of $\mathcal{A}$.

For each $v \in V$, the function $\tau_v : \mathcal{A} \to \mathcal{A}$ defined by

$$\tau_v(P) = P + v$$

is an affine isomorphism called *translation by* v. For $\varphi \in \mathrm{Aff}(\mathcal{A})$ we have

$$\varphi \tau_v \varphi^{-1} = \tau_{\varphi_*(v)}$$

and so the group $T(V)$ of all translations is a normal subgroup of $\mathrm{Aff}(\mathcal{A})$. Moreover, the assignment $\varphi \mapsto \varphi_*$ is a homomorphism from $\mathrm{Aff}(\mathcal{A})$ onto the group $\Gamma L(V)$ of all invertible semilinear transformations of V. The kernel of this homomorphism is $T(V)$.

The group $\mathrm{Aff}(\mathcal{A})$ acts on $\mathcal{A}$ and $T(V)$ is a regular normal subgroup. The stabilizer of a point $O \in \mathcal{A}$ is isomorphic to $\Gamma L(V)$ and therefore we may write $\mathrm{Aff}(\mathcal{A}) \simeq T(V).\Gamma L(V)$. More specifically, if $\varphi \in \mathrm{Aff}(\mathcal{A})$, let $w = \overrightarrow{O\varphi(O)}$ so that $\varphi(P) = \tau_w(O) + \varphi_*(\overrightarrow{OP})$. Represent φ by the pair (τ_w, φ_*) and define multiplication of pairs by

$$(\tau_v, \psi_*)(\tau_w, \varphi_*) = (\tau_v \tau_{\psi_*(w)}, \psi_* \varphi_*).$$

Thus $\mathrm{Aff}(\mathcal{A})$ is the semidirect product of $T(V)$ and $\Gamma L(V)$.

EXERCISES

2.1 If O, P and Q are points and v is a vector, show that $P + v = Q$ if and only if $\overrightarrow{OP} + v = \overrightarrow{OQ}$.

2.2 If $a_1 + a_2 + \cdots + a_r = 1$, show that the point $O + a_1\overrightarrow{OP_1} + a_2\overrightarrow{OP_2} + \cdots + a_r\overrightarrow{OP_r}$ is independent of O.

2.3 Show that the smallest affine subspace of an affine space $\mathcal{A}$ that contains the points $P_1, \ldots, P_r$ of $\mathcal{A}$ consists of the points $a_1P_1 + \cdots + a_rP_r$, where $a_1 + a_2 + \cdots + a_r = 1$.

Call this the subspace *spanned* by $P_1, P_2, \ldots, P_r$.

2.4 Show that $P_1, P_2, \ldots, P_r$ is a frame if and only if, for all i, P_i is not in the subspace spanned by the remaining points.

2.5 Let $\mathcal{B}$ be a non-empty subset of an affine space over a field of at least three elements. If every line that contains at least two points of $\mathcal{B}$ lies entirely within $\mathcal{B}$ show that $\mathcal{B}$ is an affine subspace.

2.6 If $\varphi : \mathcal{A}_1 \to \mathcal{A}_2$ is an affine isomorphism induced by a semilinear transformation, show that

$$\varphi(a_1P_1 + \cdots + a_rP_r) = \sigma(a_1)\varphi(P_1) + \cdots + \sigma(a_r)\varphi(P_r),$$

where σ is the associated isomorphism of division rings.

2.7 Show that two lines are parallel if and only if they have the same direction.

2.8 If ℓ is a line and P is a point, show that there is a unique line through P parallel to ℓ.

2.9 Prove Lemma 2.2.

2.10 Show that an affine isomorphism takes parallel lines to parallel lines.

2.11 Suppose that $f : V_1 \to V_2$ is a semilinear transformation and that W is a subspace of V_1. Show that $f(W)$ is a subspace of V_2.

2.12 Fill in the details of Theorem 2.3. In particular, show that σ is an isomorphism of division rings.

2.13 Let K be a division ring and let $\mathcal{A}$ be the set of all n-tuples $(a_1, \ldots, a_n)$ with entries from K. Make this set into a left vector space V over K by defining addition and multiplication by scalars coordinatewise. Let V act on $\mathcal{A}$ by addition and show that $\mathcal{A}$ becomes an affine space. If $\mathcal{B}$ is any affine space of dimension n over K, show that there is an affine isomorphism from $\mathcal{B}$ to $\mathcal{A}$.

2.14 Consider the affine space of pairs (a, b) of real numbers. Let φ be the map which sends (a, b) to $(a^2+b, 0)$. Show that φ takes collinear points to collinear points but φ is not induced by a semilinear transformation.

2.15 Show that the affine group $\mathrm{Aff}(\mathcal{A})$ acts 2-transitively on the points of $\mathcal{A}$ but in general it does not act 3-transitively.

2.16 Suppose that V is a finite dimensional vector space over a *field* $\mathbb{F}$ and that W is a subspace of V. Let $\mathcal{A}$ be the set of subspaces U such that $V = U \oplus W$. Show that $\mathcal{A}$ is an affine space based on the vector space $\mathrm{Hom}_{\mathbb{F}}(V/W, W)$ of linear transformations $V/W \to W$, where the action of $f \in \mathrm{Hom}_{\mathbb{F}}(V/W, W)$ on U is defined by

$$U + f := \{\, u + f(u + W) \mid u \in U \,\}.$$

2.17 If V is a vector space of dimension n over the finite field $\mathbb{F}_q$ of q elements and if U is a subspace of dimension k, show that the number of subspaces W of dimension l such that $U \cap W = \{0\}$ is

$$q^{kl}\,\frac{(q^{n-k}-1)(q^{n-k-1}-1)\cdots(q^{n-k-l+1}-1)}{(q^l-1)(q^{l-1}-1)\cdots(q-1)}.$$

(See Exercise 3.6 for an interpretation of the coefficient of q^{kl}.)

2.18 If U_1 and U_2 are subspaces of the vector space V and $\dim U_1 = \dim U_2$, show that U_1 and U_2 have a common complement in V.

3

Projective Geometry

Let V be a finite-dimensional left vector space over a division ring K. The set of subspaces of V of dimension k is known as the *Grassmannian* $G_k(V)$. The partially ordered set of *all* subspaces of V is the *projective geometry* $\mathcal{P}(V)$. If U and W are subspaces, we use $U + W$ or $\langle U, W\rangle$ to denote the subspace *spanned* by U and W. This is the smallest subspace of V that contains both U and W. Similarly, if $v_1, v_2, \ldots, v_r \in V$ we use $\langle v_1, v_2, \ldots, v_r\rangle$ to denote the subspace spanned by these vectors.

The elements of $G_1(V)$ are called *points*, the elements of $G_2(V)$ are called *lines*, and so on. If $\dim_K V = n$, the elements of $G_{n-1}(V)$ are called *hyperplanes.* In conformity with this notation the (projective) *dimension* of an element of $\mathcal{P}(V)$ is defined to be one less than its vector space dimension. Thus points have dimension 0, lines have dimension 1, planes have dimension 2, and so on. If U and W are subspaces, Grassmann's relation

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

remains true when dim is interpreted as the projective dimension.

In what follows it will be convenient to identify a subspace with the set of points that belong to it. Under this identification the subspace $\{0\}$ of V corresponds to the empty set and has dimension -1. The lattice $\mathcal{P}(V)$ can be thought of as a lattice of subsets (corresponding to subspaces) of the set of points. Its structure is determined by $G_1(V)$ and $G_2(V)$.

If V_1 and V_2 are left vector spaces of the same dimension over division rings K_1 and K_2 respectively, then any bijective semilinear transformation $f : V_1 \to V_2$ induces bijections

$$G_k(f) : G_k(V_1) \to G_k(V_2)$$

and

$$\mathcal{P}(f) : \mathcal{P}(V_1) \to \mathcal{P}(V_2).$$

The map $G_1(f)$ takes collinear points to collinear points. Conversely, if

$$\varphi : G_1(V_1) \to G_1(V_2)$$

is a bijection that takes collinear points to collinear points, the proof of Lemma 2.1 carries over to this situation (without any restriction on the size of the field this time) and shows that φ extends to a (unique) bijection

$$\varphi : \mathcal{P}(V_1) \to \mathcal{P}(V_2)$$

which preserves inclusion. Such a map is called a *collineation*.

The aim of this chapter is to show that, when the projective dimension is at least two, every collineation has the form $\mathcal{P}(f)$ for some semilinear transformation $f : V_1 \to V_2$. We do this by establishing a connection between projective and affine geometries which allows us to apply Theorem 2.3.

Indeed, if H is a hyperplane of the projective geometry $\mathcal{P}(V)$, the set $\mathcal{A} := G_1(V) \setminus G_1(H)$ can be given the structure of an affine space based on H. To see this, choose a vector $u \in V \setminus H$. Then every point of $\mathcal{A}$ can be written as $\langle u + h \rangle$, where $h \in H$ is uniquely determined by the point. The action of $w \in H$ on $\langle u + h \rangle$ is defined by

$$\langle u + h \rangle + w := \langle u + h + w \rangle.$$

The (affine) lines of $\mathcal{A}$ are the sets $G_1(W) \setminus \{W \cap H\}$, where W is a line of $\mathcal{P}(V)$ that is not contained in H. (Exercise 3.2.)

The next result is the Fundamental Theorem of Projective Geometry.

3.1 Theorem. *Suppose that V_1 and V_2 are left vector spaces of dimension n over division rings K_1 and K_2, and that $n \geq 3$. If $\varphi : G_1(V_1) \to G_1(V_2)$ is a bijection that takes collinear points of V_1 to collinear points of V_2, then*

(*i*) *$\varphi = G_1(f)$, where $f : V_1 \to V_2$ is a semilinear transformation with associated isomorphism $\sigma : K_1 \to K_2$, and*

(*ii*) *if $\varphi = G_1(f')$, where $f' : V_1 \to V_2$ is a semilinear transformation with associated isomorphism $\sigma' : K_1 \to K_2$, then there exists $b \in K_2$ such that for all $v \in V_1$ and for all $a \in K_1$ we have $f'(v) = bf(v)$ and $\sigma'(a) = b\sigma(a)b^{-1}$.*

Proof. (*i*) As proved in Lemma 2.1 of Chapter 2 for the affine case, φ takes subspaces to subspaces. Therefore it may be regarded as a collineation from $\mathcal{P}(V_1)$ to $\mathcal{P}(V_2)$. Choose a hyperplane H_1 of $\mathcal{P}(V_1)$, a vector $u_1 \in V_1 \setminus H_1$, set $H_2 = \varphi(H_1)$ and choose u_2 to span $\varphi(\langle u_1 \rangle)$. Give the sets

$$\mathcal{A}_1 = G_1(V_1) \setminus G_1(H_1) \qquad \text{and} \qquad \mathcal{A}_2 = G_1(V_2) \setminus G_1(H_2)$$

the structure of affine spaces based on H_1 and H_2 respectively (as indicated above). Then φ induces an affine isomorphism $\varphi : \mathcal{A}_1 \to \mathcal{A}_2$ and by Theorem 2.3 there is a semilinear transformation $\varphi_* : H_1 \to H_2$ such that for all $h \in$

H_1, $\varphi(\langle u_1 + h\rangle) = \varphi(\langle u_1\rangle) + \varphi_*(h)$. Let $\sigma : K_1 \to K_2$ be the isomorphism associated with φ_* and define $f : V_1 \to V_2$ by $f(au_1 + h) := \sigma(a)u_2 + \varphi_*(h)$ for all $a \in K_1$, $h \in H_1$. By construction, $G_1(f)$ coincides with φ on $\mathcal{A}_1$. Also, if $\langle h\rangle$ is a point of H_1, then

$$\varphi(\langle h\rangle) = \varphi(\langle u_1, h\rangle) \cap H_2 = \langle \varphi_*(h)\rangle = \langle f(h)\rangle.$$

Hence $\varphi = G_1(f)$ as required.

(*ii*) The map $g := f'f^{-1}$ is a semilinear transformation from V_2 to itself with associated isomorphism $\tau := \sigma'\sigma^{-1}$ and g induces the identity collineation on $G_1(V_2)$. For $v \in V_2$ we have $g(v) = b_v v$, where $b_v \in K_2$. If u and v are linearly independent, then

$$b_{u+v}(u + v) = g(u + v) = g(u) + g(v) = b_u u + b_v v$$

and therefore $b_u = b_{u+v} = b_v$. If w is any vector in V_2, then either u, v or else v, w is a linearly independent pair. In any case it follows that the value of b_w does not depend on w and we may dispense with the subscripts on b. Thus $f' = bf$ for some $b \in K_2$.

Finally, we have

$$\sigma'(a)bf(v) = b\sigma(a)f(v)$$

and therefore $\sigma'(a) = b\sigma(a)b^{-1}$ for all $a \in K_1$. □

This result will be used in Chapter 7 to determine the possible types of reflexive bilinear forms on a vector space. But now we use it to describe the group of all collineations of the projective geometry $\mathcal{P}(V)$. We assume that $\dim_K V \geq 3$.

Every element f of the group $\mathit{\Gamma L}(V)$ of semilinear transformations of V induces a collineation $\mathcal{P}(f)$ of $\mathcal{P}(V)$ and, by the theorem, every collineation of $\mathcal{P}(V)$ has this form. Also, the elements of $\mathit{\Gamma L}(V)$ that fix every point of $\mathcal{P}(V)$ form the subgroup

$$Z(V) := \{\, a\mathbf{1} \mid a \in K \,\},$$

where $\mathbf{1}$ denotes the identity transformation of V. Thus the group $P\mathit{\Gamma L}(V)$ of collineations of $\mathcal{P}(V)$ is (isomorphic to) $\mathit{\Gamma L}(V)/Z(V)$.

Axioms for Projective Geometry

Instead of defining a projective geometry as the set of subspaces of some vector space it is possible to start with a set (whose elements are called "points") together with a set of distinguished subsets ("lines") which satisfy the following axioms:

I If A and B are distinct points, there is exactly one line AB which contains both A and B.

II Each line contains at least three points.

III There is at least one line and not all points are on that line.

IV If A, B, C, D and E are five points such that B, C and D are collinear, A, C and E are collinear but A, B and C are not collinear, then there is a point F collinear with D and E and with A and B.

A set of points is said to be a *subspace* if whenever it contains distinct points A and B it contains the entire line AB. A chain of distinct non-empty subspaces

$$\mathcal{B}_0 \subset \mathcal{B}_1 \subset \mathcal{B}_2 \subset \ldots \subset \mathcal{B}_n$$

is a *flag* of length n. The geometry is *finite-dimensional* if there is an upper bound on the lengths of its flags. The *dimension* of a finite-dimensional geometry is the length of a maximal flag. If this dimension is at least three it can be shown that the geometry has the form $\mathcal{P}(V)$ for some finite-dimensional vector space over a division ring. The details can be found in Baer (1952), Chapter VII. Other discussions of the axioms occur in Whitehead (1906), Veblen and Young (1918) and Dembowski (1968). A particular case of this result will be proved in Chapter 6.

EXERCISES

3.1 If H is hyperplane of V and $\mathcal{A} = G_1(V) \setminus G_1(H)$, define an action of H on $\mathcal{A}$ by

$$\langle u + h \rangle + w := \langle u + h + w \rangle,$$

where $u \in V \setminus H$ and h, $w \in H$. Show that $\mathcal{A}$ becomes an affine space based on H.

3.2 Let $\mathcal{A}$ be the affine space of 3.1. Show that its lines are the sets $G_1(W) \setminus \{W \cap H\}$, where $W \in G_2(V) \setminus G_2(H)$.

3.3 If in 3.1 we replace u by another vector $u' \in V \setminus H$ we obtain another affine space $\mathcal{A}'$ also based on H. Give an explicit isomorphism between $\mathcal{A}$ and $\mathcal{A}'$.

3.4 Show that the group $P\Gamma L(V)$ acts faithfully and transitively on $G_k(V)$ for $0 < k \leq \dim_K V - 1$.

3.5 Let Δ be the subset of $G_k(V) \times G_k(V)$ consisting of the pairs (U_1, U_2) such that $\dim_K(U_1 \cap U_2) = k-1$. Show that Δ is an orbit of $P\mathit{\Gamma}L(V)$.

3.6 If K is the finite field of q elements and $\dim_K V = n$, show that

$$|G_k(V)| = \prod_{i=0}^{k-1} (q^n - q^i)/(q^k - q^i).$$

3.7 If H is a hyperplane of V, show that $P\mathit{\Gamma}L(V)_H$ is the affine group $T(H).\mathit{\Gamma}L(H)$.

3.8 Show that $P\mathit{\Gamma}L(V)$ acts 2-transitively on $G_1(V)$.

4

The General and Special Linear Groups

Throughout this chapter we restrict our attention to groups associated with vector spaces over *fields*. Let V be a vector space of dimension m over a field $\mathbb{F}$ and assume that $m \geq 2$. The group of all invertible linear transformations from V to itself is denoted by $GL(V)$ and called the *general linear group* of V. The group $GL(V)$ is a normal subgroup of $\mathit{\Gamma L}(V)$, the group of all invertible *semilinear* transformations of V. The quotient group is isomorphic to the group of automorphisms of $\mathbb{F}$ (Exercise 4.1).

If $e_1, e_2, \ldots, e_m$ is a basis for V and $f \in GL(V)$, then

$$f(e_j) = \sum_{i=1}^{m} a_{ij} e_i,$$

for some $a_{ij} \in \mathbb{F}$.

The assignment $f \mapsto (a_{ij})$ is an isomorphism from $GL(V)$ to the group $GL(m, \mathbb{F})$ of non-singular $m \times m$ matrices over the field $\mathbb{F}$ (Exercise 4.2).

The Dual Space

The set V^* of linear functionals $\varphi : V \to \mathbb{F}$ is also a vector space over $\mathbb{F}$ provided we define $\varphi_1 + \varphi_2$ and $a\,\varphi$ by

$$(\varphi_1 + \varphi_2)(v) := \varphi_1(v) + \varphi_2(v)$$

and

$$(a\varphi)(v) := a\varphi(v).$$

We call V^* the *dual space* of V. The basis of V^* *dual* to $e_1, e_2, \ldots, e_m$ is $\omega_1, \omega_2, \ldots, \omega_m$, where $\omega_i(a_1 e_1 + \cdots + a_m e_m) = a_i$. If $\varphi \in V^*$, then $\varphi = \varphi(e_1)\omega_1 + \cdots + \varphi(e_m)\omega_m$ and, in particular, $\dim V^* = m$. Define the action of $GL(V)$ on V^* as follows. If $f \in GL(V)$ and $\varphi \in V^*$, set

$$f(\varphi) := \varphi\, f^{-1}.$$

It turns out that the matrix of the linear transformation f (acting on V^*) with respect to $\omega_1, \omega_2, \ldots, \omega_m$ is the *transposed inverse* of the matrix of f with respect to $e_1, e_2, \ldots, e_m$ (Exercise 4.2). We also find that $\ker f(\varphi) = f(\ker \varphi)$.

The Groups $SL(V)$ and $PSL(V)$

The determinant map from $GL(V)$ to the multiplicative group $\mathbb{F}^\times$ of non-zero elements of $\mathbb{F}$ is a homomorphism onto $\mathbb{F}^\times$. The kernel of this map is the group $SL(V)$ of linear transformations of V of determinant 1. We call $SL(V)$ the *special linear group* of V.

The groups $GL(V)$ and $SL(V)$ are subgroups of $\Gamma L(V)$ and as such they act on the projective geometry $\mathcal{P}(V)$. The groups of collineations they induce on $\mathcal{P}(V)$ are denoted by $PGL(V)$ and $PSL(V)$ and are called the *projective general linear group* of V and the *projective special linear group* of V, respectively. In Chapter 3 we showed that the subgroup of $\Gamma L(V)$ that fixes every point of $\mathcal{P}(V)$ is the group $Z(V)$ of scalar transformations of V. The group $Z(V)$ is a normal subgroup of $GL(V)$ and we have $PGL(V) \simeq GL(V)/Z(V)$. We also have

$$SL(V) \cap Z(V) = \{ a\mathbf{1} \mid a^m = 1 \}$$

and

$$PSL(V) \simeq SL(V)/(SL(V) \cap Z(V)).$$

Order Formulae

If $\mathbb{F}$ is the finite field $\mathbb{F}_q$ of q elements, we denote the groups $GL(m, \mathbb{F})$, $SL(m, \mathbb{F}), \ldots$ etc. by the symbols $GL(m, q)$, $SL(m, q), \ldots$ etc. To determine the orders of these groups we first observe that $GL(V)$ acts regularly on the set of ordered bases of V. When $\mathbb{F} = \mathbb{F}_q$ the number of ordered bases is $\prod_{i=0}^{m-1}(q^m - q^i)$ (Exercise 4.5) and hence

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

Since $|\mathbb{F}_q^\times| = q - 1$ we have

$$|PGL(m, q)| = |SL(m, q)| = q^{m(m-1)/2} \prod_{i=2}^{m} (q^i - 1)$$

and

$$|PSL(m, q)| = d^{-1} q^{m(m-1)/2} \prod_{i=2}^{m} (q^i - 1),$$

where $d := (m, q - 1)$ is the greatest common divisor of m and $q - 1$.

The Action of $PSL(V)$ on $\mathcal{P}(V)$

4.1 Theorem. *The group $PSL(V)$ acts doubly transitively on the points of $\mathcal{P}(V)$.*

Proof. Let $\langle u_1\rangle$, $\langle u_2\rangle$, $\langle v_1\rangle$ and $\langle v_2\rangle$ be four points of $\mathcal{P}(V)$ such that $\langle u_1\rangle \neq \langle u_2\rangle$ and $\langle v_1\rangle \neq \langle v_2\rangle$. Extend u_1, u_2 to a basis $u_1, u_2, u_3, \ldots, u_m$ for V and similarly extend v_1, v_2 to a basis $v_1, v_2, v_3, \ldots, v_m$. Define $f \in SL(V)$ by $f(u_i) := v_i$ for $i = 2, 3, \ldots, m$ and $f(u_1) := av_1$, where a is chosen so that $\det(f) = 1$. Then $\mathcal{P}(f)$ takes $\langle u_1\rangle$ to $\langle v_1\rangle$ and $\langle u_2\rangle$ to $\langle v_2\rangle$. □

One of the aims of this chapter is to prove that, with two exceptions, $PSL(V)$ is a simple group. To do this we make use of Iwasawa's simplicity criterion which was proved in Chapter 1. The fact that $PSL(V)$ is a primitive group is an immediate consequence of the theorem just proved (*cf.* Exercise 1.13). The other conditions that we need will be obtained by studying the elements of $SL(V)$ known as *transvections*.

Transvections

Consider a linear transformation $t \in GL(V)$ that fixes every element of a hyperplane H of V. Let $v \in V \setminus H$ and let φ be the linear functional defined by

$$\varphi(av + h) := a \quad \text{for all } a \in \mathbb{F} \text{ and } h \in H.$$

Then $H = \ker\varphi$ and if $w \in V$, we have $w - \varphi(w)v \in H$. Consequently $t(w - \varphi(w)v) = w - \varphi(w)v$ and it follows that

$$t(w) = w + \varphi(w)u, \quad \text{where } u := t(v) - v.$$

Conversely, if $\varphi \in V^*$ and if u is a vector such that $\varphi(u) \neq -1$, then the map defined by

$$t_{\varphi,u}(w) := w + \varphi(w)u$$

is an element of $GL(V)$ which fixes every vector in $\ker\varphi$. If $\varphi(u) = 0$, we say that $t_{\varphi,u}$ is a *transvection*. If $\varphi(u) \neq 0, -1$, we say that $t_{\varphi,u}$ is a *dilatation*. (If $\varphi = 0$ or $u = 0$, then $t_{\varphi,u} = \mathbf{1}$.)

When considered as transformations of $\mathcal{P}(V)$, transvections are known as *elations*. In matrix theory, the elementary row operations are effected by multiplying by transvections, dilatations and permutation matrices. This is the basis of the proof of Theorem 4.3 below.

4.2 Theorem.

(i) $t_{\varphi,au} = t_{a\varphi,u}$.

(ii) $t_{\varphi_1+\varphi_2,u} = t_{\varphi_1,u}\, t_{\varphi_2,u}$, *whenever* $\varphi_1(u) = 0$.

(iii) $t_{\varphi,u_1+u_2} = t_{\varphi,u_1}\, t_{\varphi,u_2}$, *whenever* $\varphi(u_2) = 0$.

(iv) $f\, t_{\varphi,u} f^{-1} = t_{f(\varphi),f(u)}$ *for all* $f \in GL(V)$.

Proof. Exercise 4.7. □

We shall see that, provided the basis is suitably chosen, the matrix form of a transvection is extremely simple. Suppose that $e_1, e_2, \ldots, e_m$ is a basis for V and that $\omega_1, \omega_2, \ldots, \omega_m$ is the dual basis. If φ and u are non-zero elements of V^* and V such that $\varphi(u) = 0$, there is an element $f \in GL(V)$ such that $f(\varphi) = \omega_m$ and $f(u) = e_1$. Then by Theorem 4.2 (*iv*) we have

$$f t_{\varphi,u} f^{-1} = t_{\omega_m,e_1}.$$

The matrix of t_{ω_m,e_1} with respect to $e_1, e_2, \ldots, e_m$ is

$$\begin{pmatrix} 1 & 0 & \ldots & 1 \\ 0 & 1 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

and hence $\det(t_{\omega_m,e_1}) = 1$. This shows that in $GL(V)$ all transvections are conjugate to t_{ω_m,e_1} and consequently they all belong to $SL(V)$. Moreover, if $m \geq 3$, the element f can be chosen in $SL(V)$ and therefore, in this case, $SL(V)$ has only one conjugacy class of transvections. (If $m = 2$, the conjugacy classes of transvections are in one-to-one correspondence with the cosets of $\{\, a^2 \mid a \in \mathbb{F}^\times \,\}$ in $\mathbb{F}^\times$.)

Now let $P = \langle u \rangle$ be a point of $\mathcal{P}(V)$ and consider the set

$$X_P := \{\, t_{\varphi,u} \mid \varphi \in V^*,\ \varphi(u) = 0 \,\}.$$

4.3 Theorem. *(i)* X_P *is an abelian normal subgroup of* $GL(V)_P$.

(ii) $SL(V) = \langle\, f X_P f^{-1} \mid f \in SL(V) \,\rangle$.

Proof. It follows from Theorem 4.2 (*ii*) that X_P is an abelian group and from Theorem 4.2 (*iv*) that X_P is normal in $GL(V)_P$. The argument immediately preceding this theorem shows that every transvection is conjugate in $SL(V)$ to one of the form $t_{a\varphi,u}$ for some *fixed* $\varphi \in V^*$, $u \in V$. In particular, the subgroup

$$\langle\, f X_P f^{-1} \mid f \in SL(V) \,\rangle$$

contains all the transvections of $SL(V)$. It remains to show that every element of $SL(V)$ is a product of transvections. We do this by relating transvections to the elementary row operations of matrix theory.

Choose a basis $e_1, e_2, \ldots, e_m$ for V and let $\omega_1, \omega_2, \ldots, \omega_m$ be the dual basis. If $i \neq j$ and $f \in GL(V)$, the matrix of $t_{\omega_j, ae_i} f$ with respect to $e_1, e_2, \ldots, e_m$ is obtained from the matrix of f by adding a times the jth row to the ith row. This is because

$$t_{\omega_j, ae_i} = \mathbf{1} + ae_{ij},$$

where

$$e_{ij}(v) := \omega_j(v)e_i$$

The matrix of e_{ij} has 1 in the (i, j)th place and 0 everywhere else. Thus multiplication by t_{ω_j, ae_i} (on the left) effects an elementary row operation on the matrix of f. But it is clear that by using elementary row operations of this type the matrix of f can be reduced to the diagonal matrix

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & d \end{pmatrix}$$

where $d := \det(f)$. In particular, every element of $SL(V)$ can be expressed as a product of transvections. □

The Simplicity of $PSL(V)$

4.4 Theorem. *If $m \geq 3$ or if $|\mathbb{F}| \geq 4$, then $SL(m, \mathbb{F})' = SL(m, \mathbb{F})$.*

Proof. Suppose that $m \geq 3$ and choose a pair of transvections $t_{\varphi_1, u}$ and $t_{\varphi_2, u}$ with φ_1 and φ_2 linearly independent. We know that for some $f \in SL(V)$ we have $t_{\varphi_1, u} = f\, t_{\varphi_2, u} f^{-1}$ and therefore

$$t_{\varphi_1 - \varphi_2, u} = f\, t_{\varphi_2, u} f^{-1}\, t_{\varphi_2, u}^{-1} \in SL(V)'.$$

Since $SL(V)$ is generated by its single conjugacy class of transvections we have $SL(V)' = SL(V)$.

Now suppose that $m = 2$ and $|\mathbb{F}| \geq 4$. Choose a basis e_1, e_2 with dual basis ω_1, ω_2 and define $f \in SL(V)$ by $f(e_1) := ae_1$ and $f(e_2) := a^{-1}e_2$, where $a^2 \neq 1$. Then

$$f\, t_{\omega_2, be_1} f^{-1} = t_{\omega_2, a^2 be_1}$$

and consequently $SL(V)'$ contains the transvection

$$f\,t_{\omega_2,be_1}f^{-1}t^{-1}_{\omega_2,be_1} = t_{\omega_2,(a^2-1)be_1}.$$

It follows that $SL(V)'$ contains all the transvections and hence $SL(V)' = SL(V)$ as required. □

4.5 Theorem. *The groups $PSL(m,\mathbb{F})$ are simple except for $PSL(2,\mathbb{F}_2)$ and $PSL(2,\mathbb{F}_3)$.*

Proof. This follows from Theorem 1.2 (Iwasawa's criterion) together with Theorems 4.1, 4.3 and 4.4. The groups $PSL(2,\mathbb{F}_2)$ and $PSL(2,\mathbb{F}_3)$ are genuine exceptions to this theorem This will appear as a consequence of the more extensive calculations with $PSL(2,\mathbb{F}_q)$ in the next section. □

The Groups $PSL(2,q)$

Let $\Omega := \{1,2,\ldots,n\}$ and recall that the *alternating group* A_n consists of the permutations of Ω that can be written as a product of an *even* number of transpositions. (A *transposition* is a permutation that interchanges a pair of elements of Ω and fixes the rest.) The group A_n is the unique subgroup of index 2 in the *symmetric group* S_n of all permutations of Ω. For certain small values of m, n and q the groups A_n and $PSL(m,q)$ are isomorphic.

Consider the group $PSL(2,q)$. It acts on the projective line $\mathcal{P}(V)$, where $V := \langle e_1, e_2\rangle$ is a two-dimensional vector space over $\mathbb{F}_q$. Identify the point $\langle ae_1 + be_2\rangle$, $b \neq 0$ with $ab^{-1} \in \mathbb{F}_q$ and set $\infty := \langle e_1\rangle$. The projective line $\mathcal{P}(V)$ may be regarded as the set $\mathcal{P} := \mathbb{F}_q \cup \{\infty\}$ and $PSL(2,q)$ may be identified with the group of *linear fractional transformations*

$$z \mapsto \frac{az+b}{cz+d},$$

where $ad - bc$ is a square in $\mathbb{F}^\times$.

Note that

$$PSL(2,q)_\infty = \{\, z \mapsto a^2z + b \mid a \in \mathbb{F}^\times, b \in \mathbb{F}\,\}$$

and that

$$PSL(2,q)_{\infty,0} = \{\, z \mapsto a^2z \mid a \in \mathbb{F}^\times\,\}$$

is a cyclic group of order $q-1$ if q is even or of order $(q-1)/2$ if q is odd.

The group

$$X_\infty = \{\, z \mapsto z + b \mid b \in \mathbb{F}\,\}$$

is a Sylow subgroup of $PSL(2,q)$ and isomorphic to the additive group of $\mathbb{F}_q$.

The group $PSL(2,q)_\infty$ is the semidirect product of X_∞ and $PSL(2,q)_{\infty,0}$. If q is even, we have

$$|PSL(2,q)| = q(q^2-1)$$

and if q is odd, we have

$$|PSL(2,q)| = \tfrac{1}{2}q(q^2-1).$$

Since $|\mathcal{P}| = q+1$, $PSL(2,q)$ is a subgroup of the symmetric group S_{q+1}. So just by comparing orders we find that

$$PSL(2,2) \simeq S_3, \quad PSL(2,3) \simeq A_4, \quad \text{and} \quad PSL(2,4) \simeq A_5.$$

In particular, neither $PSL(2,2)$ nor $PSL(2,3)$ are simple groups.

It follows from Theorem 4.5 that A_5 is a simple group and it is easy to show (by induction) that A_n is simple for $n \geq 5$ (Exercise 4.8). It can be shown that, up to isomorphism, there is only one simple group of order 60 and from this it follows that $PSL(2,5) \simeq A_5$. (See Exercise 6.4 for a description of this isomorphism.)

The simple groups $PSL(m,q)$ and A_n whose orders are less than 1000 are given in the following table.

group	order
$PSL(2,4)$	60
$PSL(2,5)$	60
A_5	60
$PSL(2,8)$	504
$PSL(2,11)$	660

group	order
$PSL(2,7)$	168
$PSL(3,2)$	168
$PSL(2,9)$	360
A_6	360

In addition to the coincidences of orders in the above table we have

$$|PSL(3,4)| = |PSL(4,2)| = |A_8|.$$

(See Exercise 6.6.)

To see that $PSL(3,4)$ is not isomorphic to $PSL(4,2)$ notice that the centre of a Sylow 2-subgroup of $PSL(3,4)$ has order 4, whereas the centre of a Sylow 2-subgroup of $PSL(4,2)$ has order 2. In both cases the upper triangular matrices with all diagonal elements equal to 1 form a Sylow 2-subgroup. See Lemma 5.7 for the relevant commutator relations.

4.6 Theorem. *The only isomorphisms between the groups $PSL(m,q)$ and A_n are the following:*

(i) $PSL(2,3) \simeq A_4$,

(ii) $PSL(2,4) \simeq PSL(2,5) \simeq A_5$,

(iii) $PSL(2,7) \simeq PSL(3,2)$,

(iv) $PSL(2,9) \simeq A_6$,

(v) $PSL(4,2) \simeq A_8$.

These isomorphisms will be studied in detail in later chapters. A proof of this theorem can be found in Huppert (1968/69), §5 and §6 and in Artin (1955a). These proofs have been used as the basis of exercises in the following chapters. (Exercises 4.11, 5.18, 5.19 and 6.6.)

EXERCISES

4.1 Show that $\Gamma L(V)/GL(V) \simeq \mathrm{Aut}(\mathbb{F})$.

4.2 Suppose that $e_1, e_2, \ldots, e_m$ is a basis for V and that $\omega_1, \omega_2, \ldots, \omega_m$ is the dual basis for V. For $f \in GL(V)$, write

$$f(e_j) = \sum_{i=1}^{m} a_{ij} e_i.$$

(i) Show that the assignment $f \mapsto (a_{ij})$ is an isomorphism from $GL(V)$ onto the group of non-singular matrices over $\mathbb{F}$.

(ii) Show that the matrix representing the action of f on V^* with respect to $\omega_1, \omega_2, \ldots, \omega_m$ is the transposed inverse of (a_{ij}).

4.3 If $\varphi \in V^*$ and $f \in GL(V)$, show that $\ker f(\varphi) = f(\ker \varphi)$.

4.4 Show that if $\dim V \geq 3$, $PGL(V)$ does not act 3-transitively on the points of $\mathcal{P}(V)$.

4.5 Show that the number of ordered bases of a vector space of dimension m over $\mathbb{F}_q$ is

$$\prod_{i=0}^{m-1} (q^m - q^i).$$

4.6 Show that if q is even, $PSL(2,q)$ is 3-transitive on the points of the projective line over $\mathbb{F}_q$.

4.7 Prove Theorem 4.2.

4.8 Assuming that A_5 is simple, show by induction that A_m is simple for $m > 5$.

4.9 Show that if $f \in \Gamma L(V)$ commutes with every element of $SL(V)$ and $\dim V \geq 2$, then $f \in Z(V)$.

4.10 Show that the transvections $t_{\varphi,u}$ and $t_{\psi,v}$ commute if and only if $\varphi(v) = \psi(u) = 0$.

4.11 Suppose that q is a power of a prime and that p is a prime which does not divide q. If $p = 2$, suppose that $q \equiv 1 \bmod 4$. Let f be the least positive integer such that p divides $q^f - 1$.

(*i*) Show that if p divides $q^i - 1$, then f divides i.

(*ii*) Show that p^a divides i if and only if p^a divides $(q^{fi} - 1)/(q^f - 1)$.

(*iii*) Show that f divides $p - 1$.

(*iv*) Given m, let $l = \lfloor m/f \rfloor$ be the largest integer $\leq m/f$ and let D be the direct product of l copies of $GL(f, q)$. Regarding D as the set of l-tuples $(x_1, x_2, \ldots, x_l)$ with $x_i \in GL(f, q)$, show that the symmetric group S_l acts as a group of automorphisms of D, permuting the entries of the l-tuples. Let E be the semidirect product of D by S_l and show that a Sylow p-subgroup of E is isomorphic to a Sylow p-subgroup of $GL(m, q)$.

(*v*) Interpret $\mathbb{F}_{q^f}$ as a vector space of dimension f over $\mathbb{F}_q$ and show that multiplication by an element of $\mathbb{F}_{q^f}$ induces a linear transformation of this space. Deduce that the Sylow p-subgroups of $GL(f, q)$ are cyclic.

(*vi*) Show that the order of a Sylow p-subgroup of S_l is p^k, where $k = \lfloor \frac{l}{p} \rfloor + \lfloor \frac{l}{p^2} \rfloor + \cdots$, and then show that $k \leq l/(p-1)$.

(*vii*) Show that the order of a Sylow p-subgroup P of $PSL(m, q)$ is strictly less than $(q\sqrt{3})^m$ when p is odd and strictly less than $(2q)^m$ when $p = 2$. Show that a Sylow subgroup corresponding to the characteristic of $\mathbb{F}_q$ has order larger than $|P|$ except in $PSL(2, 8)$ when $p = 3$ and in those groups $PSL(2, 2^r)$ for which $p = 2^r + 1$ is prime. (When $p = 2$ and $q \equiv 3 \bmod 4$, the only additional exceptions are the groups $PSL(2, q)$ for which $q = 2^s - 1$ is prime.)

5

BN-pairs and Buildings

It was mentioned in the Preface that most of the groups studied in this book can be obtained from the classical Lie algebras by constructions due to Chevalley, Steinberg, Tits and Hertzig. Whilst this provides a uniform approach to these groups it does not immediately yield descriptions related to the geometry of the underlying vector space (see Ree (1957) or Carter (1972) for the actual connections).

Various ways of obtaining geometries associated with the groups have been studied extensively by Jacques Tits. Here we shall be particularly concerned with the notion of the *Tits building* of a group (Tits (1974)) and the related algebraic concept of *BN-pair*, also due to Tits (1962). This does not mean that we shall investigate Lie algebras or abstract buildings; rather we show directly how the BN-pairs and buildings arise from the underlying projective geometry.

The BN-pair Axioms

A *BN-pair* in a group G is a pair of subgroups B and N such that

(i) $G = \langle B, N \rangle$

(ii) *$H := B \cap N$ is a normal subgroup of N.*

(iii) *$W := N/H$ is generated by elements $\{ w_i \mid i \in I \}$ such that $w_i^2 = 1$ for all $i \in I$.*

(iv) *If $w_i = n_i H$ and $n \in N$, then*

  (a) *$n_i B n_i \neq B$ and*

  (b) *$n_i B n \subseteq (B n_i n B) \cup (B n B)$.*

Any subgroup of G conjugate to B is called a *Borel* subgroup. The group W is called the *Weyl group* and $|I|$ is the *rank* of the BN-pair. The generators for W are uniquely determined by the other conditions and therefore the rank is well-defined. (See Chapter 9 for the details.)

Some simple consequences of these axioms can be found in the exercises at the end of this chapter and in Chapter 9. More detailed treatments of groups with BN-pairs can be found in the works of Bourbaki (1968), Chapitre IV, Carter (1972, 1977 and 1985), Suzuki (1982) and Tits (1974). For the remainder of this chapter we shall concentrate on constructing the standard BN-pair for $SL(V)$.

The Tits Building

Recall from Chapter 3 that the projective geometry $\mathcal{P}(V)$ is a partially ordered set in which the partial order is inclusion of subspaces, and that a *flag* of $\mathcal{P}(V)$ is just a chain of distinct subspaces $V_1 \subset V_2 \subset \cdots \subset V_k$. A *proper* flag is one in which neither $\{0\}$ nor V occurs. Regarding flags simply as sets of subspaces we order them by inclusion and in this way turn the set $\Delta(V)$ of all proper flags of $\mathcal{P}(V)$ into a partially ordered set whose least element is the empty flag $\emptyset$.

The *type* of a proper flag $V_1 \subset V_2 \subset \cdots \subset V_k$ is the set $\{d_1, d_2, \ldots, d_k\}$ where $d_i := \dim V_i$ for $i = 1, 2, \ldots, k$. The group $SL(V)$ acts on $\Delta(V)$ and it is easy to see that two flags are in the same orbit of $SL(V)$ if and only if they have the same type. If $m := \dim V$, the *maximal* flags are those of type $\{1, 2, \ldots, m-1\}$. The group B in the BN-pair for $SL(V)$ will turn out to be the stabilizer of a maximal flag.

A *frame* in $\mathcal{P}(V)$ is a set of points $\mathcal{F} := \{\langle e_1\rangle, \langle e_2\rangle, \ldots, \langle e_m\rangle\}$, where $e_1, e_2, \ldots, e_m$ is a basis for V. The *apartment* $\Sigma(\mathcal{F})$ of the frame $\mathcal{F}$ consists of those proper flags $V_1 \subset V_2 \subset \cdots \subset V_k$ for which each V_i is spanned by some subset of $\{e_1, e_2, \ldots, e_m\}$. It is not hard to see that the maximal flags in $\Sigma(\mathcal{F})$ correspond to the $m!$ orderings of $e_1, e_2, \ldots, e_m$ (Exercise 5.3). The group $SL(V)$ is transitive on the set of all apartments and it will turn out that the stabilizer of an apartment is the group N of the BN-pair.

The set $\Delta(V)$ of all proper flags of $\mathcal{P}(V)$ together with the set $\mathcal{A}$ of all apartments is known as the *building* of $\mathcal{P}(V)$ and the maximal flags are called the *chambers* of $\Delta(V)$. (See Chapter 9 for the abstract definition of a building.) The group $P\Gamma L(V)$ acts faithfully on $\Delta(V)$ and $\mathcal{A}$ and preserves the partial order. In Chapter 7 we shall determine all the maps of $\Delta(V)$ with this property.

The BN-pair of $SL(V)$

Now suppose that $V_1 \subset V_2 \subset \cdots \subset V_{m-1}$ is a chamber (maximal flag) and choose a basis $e_1, e_2, \ldots, e_m$ for V so that $V_i = \langle e_1, e_2, \ldots, e_i\rangle$ for $1 \le i < m$. Let B be the stabilizer of this chamber in $SL(V)$ and let N be the setwise stabilizer of the frame $\{\langle e_1\rangle, \langle e_2\rangle, \ldots, \langle e_m\rangle\}$. That is, $f \in SL(V)$ is in B if and only if $f(V_i) = V_i$ for all i and $f \in N$ if and only if $f(e_i)$ is a scalar multiple of e_j for some j. Thus, with respect to the basis $e_1, e_2, \ldots, e_m$, the elements of B have upper triangular matrices and the elements of N have *monomial* matrices (i.e., exactly one non-zero entry in each row and column). The group

$$H := B \cap N$$

is the pointwise stabilizer of the frame and consequently its elements are precisely those elements of $SL(V)$ that have diagonal matrices with respect to the given basis.

Each monomial matrix factorizes (uniquely) in $GL(m, \mathbb{F})$ into a product of a diagonal matrix and a permutation matrix. A *permutation matrix* is a monomial matrix all of whose non-zero entries are 1. The permutation π of $\{1, 2, \ldots, m\}$ corresponds to the permutation matrix with 1 in the $(\pi(i), i)$-th place (for $i = 1, 2, \ldots, m$) and 0 elsewhere.

It follows that $W := N/H$ is isomorphic to the group of all $m \times m$ permutation matrices. In other words, W is the symmetric group S_m. We shall show that B, N, W and I satisfy the BN-pair axioms, where I is the set $\{1, 2, \ldots, m-1\}$ and for $i \in I$, w_i is the transposition $(i, i+1)$ of S_m.

We observed in the proof of Theorem 4.3 that multiplying $g \in SL(V)$ on the left by the transvection $\mathbf{1} + ae_{ij}$ effects an elementary row operation on the matrix of g, adding a times row j to row i. Similarly, multiplication on the right by $\mathbf{1} + ae_{ij}$ effects an elementary column operation. The element $\mathbf{1} + ae_{ij}$ belongs to B whenever $i < j$. Using the elementary row operations corresponding to these transvections the matrix of g can be reduced to one in which there is exactly one non-zero entry in the first column, say in row i_1. Then using elementary column operations corresponding to the above transvections the other entries in row i_1 can be reduced to zero. Continuing in this fashion, g can be reduced to a matrix in which there is exactly one non-zero entry in each row and each column; i.e., g can be reduced to an element of N. Thus

$$SL(V) = BNB$$

and, in particular,

$$SL(V) = \langle B, N \rangle.$$

This establishes the first of the BN-pair axioms and also shows that double coset representatives for B may be chosen from N. If $n \in N$ and $w = nH$, it is customary (and well-defined) to write BwB in place of BnB. This allows us to write

$$SL(V) = \bigcup_{w \in W} BwB. \tag{5.1}$$

Later we shall show that each double coset determines a *unique* element of W. But before doing that we shall look at a 'geometric' interpretation of (5.1).

Chambers

Let M denote the chamber

$$\{\langle e_1, e_2, \dots, e_i\rangle \mid 1 \le i < m\}$$

with stabilizer B and let $\mathcal{F}$ denote the frame

$$\{\langle e_1\rangle, \langle e_2\rangle, \dots, \langle e_m\rangle\}$$

with stabilizer N. Since $SL(V)$ is transitive on chambers we may identify them with the cosets gB of B. Then the double cosets BgB correspond to the orbits of $SL(V)$ on *pairs* of chambers: each orbit has a representative of the form $(M, g(M))$ (*cf.* Exercises 1.8 and 1.9).

5.2 Theorem. *If M_1 and M_2 are chambers of $\Delta(V)$, there is an apartment Σ such that M_1, $M_2 \in \Sigma$.*

Proof. We may suppose that $M_1 = M$, where M is defined above. If $w \in W$, then $w(M)$ is well-defined and it follows from (5.1) that the orbits of $SL(V)$ on pairs of chambers are represented by the pairs $(M, w(M))$, $w \in W$. That is, each orbit is represented by a pair of chambers of the apartment $\Sigma(\mathcal{F})$. □

This is the 'geometric' version of (5.1). Indeed Theorem 5.2 can be proved directly (Exercise 5.10) thus giving another proof of (5.1).

Flags and Apartments

We now return to the question of the uniqueness of the double coset representatives. It is possible to show directly that for w_1, $w_2 \in W$ we have $Bw_1B = Bw_2B$ if and only if $w_1 = w_2$. But instead of using a direct argument we shall obtain the result via a more detailed study of flags and apartments.

5.3 Lemma. *If Γ_1 and Γ_2 are subsets of a basis $\{e_1, e_2, \dots, e_m\}$ for V and if U_1 and U_2 are the subspaces spanned by Γ_1 and Γ_2 respectively, then $\Gamma_1 \cup \Gamma_2$ is a basis for $U_1 + U_2$ and $\Gamma_1 \cap \Gamma_2$ is a basis for $U_1 \cap U_2$.*

Proof. The set $\Gamma_1 \cup \Gamma_2$ certainly spans $U_1 + U_2$ and it is linearly independent. Therefore it is a basis for $U_1 + U_2$. A simple calculation, using Grassmann's relation, shows that

$$\dim(U_1 \cap U_2) = |\Gamma_1 \cap \Gamma_2|$$

and therefore $\Gamma_1 \cap \Gamma_2$ is a basis for $U_1 \cap U_2$. □

5.4 Theorem. *Suppose that $\mathcal{F}_1$ and $\mathcal{F}_2$ are frames and that F and F' are flags which belong to both apartments $\Sigma(\mathcal{F}_1)$ and $\Sigma(\mathcal{F}_2)$. Then there is a map $f \in SL(V)$ such that $f(F) = F$, $f(F') = F'$ and $f(\mathcal{F}_1) = \mathcal{F}_2$.*

Proof. Suppose that F is the flag $V_1 \subset V_2 \subset \cdots \subset V_k$ and that F' is the flag $V_1' \subset V_2' \subset \cdots \subset V_l'$. Since $F \in \Sigma(\mathcal{F}_1)$, $\mathcal{F}_1$ is the union of disjoint subsets $\Phi_1, \Phi_2, \ldots, \Phi_{k+1}$ such that for $1 \le i \le k$,

$$\Phi_1 \cup \Phi_2 \cup \cdots \cup \Phi_i$$

is a frame for V_i. Similarly, $\mathcal{F}_1$ is the disjoint union of subsets Φ_1', Φ_2', $\ldots$, Φ_{l+1}' such that for $1 \le i \le l$,

$$\Phi_1' \cup \Phi_2' \cup \cdots \cup \Phi_i'$$

is a frame for V_i'.

A straightforward calculation, making use of Lemma 5.3, shows that

$$|\Phi_i \cap \Phi_j'| = \dim(V_i \cap V_j' / (V_{i-1} \cap V_j' + V_i \cap V_{j-1}')).$$

By symmetry, a similar result holds for $\mathcal{F}_2$. That is, $\mathcal{F}_2$ is the disjoint union of subsets $\Gamma_1, \Gamma_2, \ldots, \Gamma_{k+1}$ and of subsets $\Gamma_1', \Gamma_2', \ldots, \Gamma_{l+1}'$ such that

$$\Gamma_1 \cup \Gamma_2 \cup \cdots \cup \Gamma_i$$

is a frame for V_i and

$$\Gamma_1' \cup \Gamma_2' \cup \cdots \cup \Gamma_i'$$

is a frame for V_i'. It follows that

$$|\Gamma_i \cap \Gamma_j'| = |\Phi_i \cap \Phi_j'|$$

and therefore we may choose $f \in SL(V)$ so that $f(\Phi_i \cap \Phi_j') = \Gamma_i \cap \Gamma_j'$ for all i and j. We have $f(V_i) = V_i$, $f(V_j') = V_j'$ and $f(\mathcal{F}_1) = \mathcal{F}_2$. This completes the proof. □

5.5 Theorem. *For $n_1, n_2 \in N$, $Bn_1B = Bn_2B$ if and only if $n_1H = n_2H$.*

Proof. We have already observed that any pair of maximal flags can be transformed into a pair of the form $(M, n(M))$, for some $n \in N$, where M is the flag

$$\{\, \langle e_1, \ldots, e_i \rangle \mid 1 \le i < m \,\}.$$

Moreover, for n_1, $n_2 \in N$ we have $Bn_1B = Bn_2B$ if and only if $bn_1(M) = n_2(M)$ for some $b \in B$. If this is the case, then M and $n_2(M)$ belong to both Σ and to $b(\Sigma)$, where Σ is the apartment of

$$\{\langle e_1\rangle, \langle e_2\rangle, \dots, \langle e_m\rangle\}.$$

By Theorem 5.4 there is a map $f \in SL(V)$ which fixes M and $n_2(M)$ and takes $b(\Sigma)$ to Σ. Thus $f \in B$ and $fb \in B \cap N$. But $B \cap N$ is a normal subgroup of N and so $n_1^{-1} fbn_1 \in B$. It follows that

$$n_1(M) = n_1(n_1^{-1} fbn_1)(M) = fbn_1(M) = n_2(M).$$

Therefore $n_2^{-1}n_1 \in B \cap N$ and hence n_1 and n_2 represent the same element of W. □

Panels

In order to complete the proof that B and N form a BN-pair we need to show that axiom (*iv*) holds. One way to do this is to introduce for each hyperplane H and projective point P of H, the *root group*

$$X_{P,H} := \{\, t_{\varphi,u} \mid u \in P,\ \ker\varphi = H \,\}$$

and to make explicit calculations involving certain root groups. This approach is outlined in the exercises. The method of proof that we shall use here involves another application of Theorem 5.4.

In preparation for this proof we introduce another term. A proper flag consisting of exactly $m-2$ subspaces is called a *panel* of $\Delta(V)$. (Suzuki (1982) calls this a wall.) Recall that a chamber of $\Delta(V)$ is a maximal flag and consists of $m-1$ distinct subspaces.

5.6 Lemma. *If A is a panel in the apartment Σ, then A is contained in exactly two chambers of Σ.*

Proof. We may choose a basis $f_1, f_2, \dots, f_m$ for V so that Σ is the apartment of the frame $\{\langle f_1\rangle, \langle f_2\rangle, \dots, \langle f_m\rangle\}$ and A consists of all the subspaces

$$\langle f_1\rangle, \quad \langle f_1, f_2\rangle, \quad \dots, \quad \langle f_1, f_2, \dots, f_{m-1}\rangle$$

except $\langle f_1, f_2, \dots, f_i\rangle$ for some i. The only way to extend A to a chamber of Σ is to adjoin either

$$\langle f_1, f_2, \dots, f_{i-1}, f_i\rangle$$

or

$$\langle f_1, f_2, \dots, f_{i-1}, f_{i+1}\rangle. \qquad □$$

Now let n_i be the element of $SL(V)$ which sends e_i to $-e_{i+1}$, e_{i+1} to e_i and fixes e_j for all $j \neq i, i+1$. Then $w_i = n_i H$ corresponds to the transposition $(i, i+1)$ of S_m and it is easy to see that $n_i B n_i \neq B$.

If M is the chamber $\{ \langle e_1, \ldots, e_j \rangle \mid 1 \leq j < m \}$, then $A = M \cap n_i(M)$ is the panel obtained from M by omitting $\langle e_1, e_2, \ldots, e_i \rangle$. If $b \in B$ and $n \in N$, then b fixes A and therefore

$$n(A) \subseteq nbn_i(M).$$

By Theorem 5.2 there is an apartment Σ' which contains both M and $nbn_i(M)$. Then Σ and Σ' both contain M and $n(A)$ and so by Theorem 5.4 there is a map $f \in SL(V)$ which fixes M and $n(A)$ and takes Σ' to Σ. Thus $f \in B$ and $fnbn_i(M)$ is a chamber of Σ which contains $n(A)$. It follows from Lemma 5.6 that $fnbn_i(M)$ is $n(M)$ or $nn_i(M)$. Thus $nbn_i \in BnB$ or Bnn_iB. On taking inverses and replacing n by n^{-1} we find that

$$n_i Bn \subseteq (BnB) \cup (Bn_i nB).$$

This completes the proof that B and N form a BN-pair for $SL(V)$.

Split BN-pairs

Recall from Chapter 1 that if x and y are elements of a group G, the *commutator* of x and y is the element

$$[x, y] := xyx^{-1}y^{-1}.$$

If X and Y are subgroups of G, we set

$$[X, Y] := \langle\, [x, y] \mid x \in X,\ y \in Y \,\rangle.$$

The *lower central series* of G consists of the groups $L_i(G)$, where $L_1(G) := G$ and $L_{i+1} := [L_i(G), G]$, for $i = 1, 2, \ldots$. The group G is said to be *nilpotent of class* c if c is the least positive integer for which $L_{c+1}(G) = 1$.

The BN-pair constructed for $SL(V)$ actually satisfies the following two conditions in addition to the BN-pair axioms.

(1) $B = UH$, where U is a normal nilpotent subgroup of B such that $U \cap H = 1$.

(2) $H = \bigcap_{n \in N} nBn^{-1}$.

The subgroup U of $SL(V)$ can be described as follows. Let $e_1, e_2, \ldots, e_m$ be a basis for the vector space V and let $\omega_1, \omega_2, \ldots, \omega_m$ be its dual basis. Recall that e_{ij} is the linear transformation defined by $e_{ij}(v) := \omega_j(v)e_i$ and that the matrix of e_{ij} has 1 in the (i, j)th position and 0 elsewhere. For $i \neq j$, let X_{ij} denote the root group $X_{\langle e_i \rangle, \ker \omega_j}$. Then X_{ij} consists of the transvections $\mathbf{1} + ae_{ij}$, $a \in \mathbb{F}$. We define

$$U := \langle X_{ij} \mid i < j \rangle.$$

The elements of U correspond to the upper triangular matrices with 1's along the diagonal. If we take B to be the stabilizer of the chamber $\{ \langle e_1, \ldots, e_j \rangle \mid 1 \leq j < m \}$ then B may be identified with the group of all upper triangular matrices (of determinant 1). Taking N to correspond to the monomial matrices, $H = B \cap N$ is the group of all diagonal matrices and (1) is clear. Note that U is the (normal) subgroup of B which acts as the identity on each factor space

$$\langle e_1, \ldots, e_{j+1} \rangle / \langle e_1, \ldots, e_j \rangle.$$

If $n_0 \in N$ corresponds to the permutation $(1, m)(2, m-1) \cdots$ which interchanges i and $m - i + 1$, then $n_0 B n_0^{-1}$ consists of lower triangular matrices and thus

$$H = B \cap n_0 B n_0^{-1};$$

a stronger form of (2). We say that $SL(V)$ has a *split* BN-pair.

Note that when $\mathbb{F}$ is a finite field of characteristic p, U is a Sylow p-subgroup of $SL(V)$. In general, U is called the *unipotent radical* of B. For a more complete discussion and a precise description of split BN-pairs in the context of algebraic groups, see Carter (1985) and the references given there.

Commutator Relations

We shall show that U is nilpotent of class $m - 1$ and to this end we prove a lemma describing the commutator relations between the X_{ij} (see Carter (1972), p. 76).

5.7 Lemma. *If $i < j$ and $k < l$, then*

$$[X_{ij}, X_{kl}] = \begin{cases} X_{il} & \text{if } j = k, \\ X_{kj} & \text{if } i = l, \\ 1 & \text{otherwise} \end{cases}$$

Proof. This is an immediate consequence of the fact that $e_{ij}e_{kl} = \delta_{jk}e_{il}$. $\square$

For $i = 1, 2, \ldots, n-1$ let $X_{(i)}$ denote the following subset of U.

$$X_{1,i+1}X_{2,i+2}\ldots X_{n-i,n}$$

5.8 Theorem. (i) *The group U is nilpotent of class $m-1$ and the i-th term of its lower central series is $L_i(U) = \langle\, X_{jk} \mid k-j \geq i \,\rangle$.*

(ii) $L_i(U) = X_{(i)}L_{i+1}(U)$.

(iii) *$L_i(U)/L_{i+1}(U)$ is isomorphic to the additive group of a vector space of dimension $m-i$ over $\mathbb{F}$.*

Proof. Part (ii) follows directly from Lemma 5.7 by induction. From this it is easy to deduce the parts (i) and (iii) of the theorem. (Consider the form of the matrices $L_i(U)$.) □

The Weyl Group

Now let us investigate the action of the Weyl group of the BN-pair on U. As before we take N to be the subgroup of $SL(V)$ whose elements have monomial matrices with respect to $e_1, e_2, \ldots, e_m$. Thus the matrix of $g \in N$ has non-zero entries only in positions $(w(j), j)$ for some permutation w in the symmetric group S_m. The mapping $g \mapsto w$ is a homomorphism from N onto S_m with kernel $B \cap N$ and so we may identify the Weyl group W of the BN-pair with the symmetric group S_m.

For $w \in W$, let n_w be an element of N which maps onto w, set

$$U_w^+ := \langle\, X_{jk} \mid j < k \quad \text{and} \quad w(j) < w(k) \,\rangle$$

and set

$$U_w^- := \langle\, X_{jk} \mid j < k \quad \text{and} \quad w(j) > w(k) \,\rangle.$$

5.9 Lemma. *We have $U = U_w^+U_w^-$, $U_w^+ \cap U_w^- = 1$ and $n_wU_w^+n_w^{-1} = U_{w^{-1}}^+$.*

Proof. Let $X_{(i)}^+$ be the product of those groups $X_{j,i+j}$ for which $w(j) < w(i+j)$ and let $X_{(i)}^-$ be the product of the $X_{j,i+j}$ for which $w(j) > w(i+j)$. These sets are not necessarily groups, but by Lemma 5.7 the sets

$$L_i^+ := X_{(i)}^+X_{(i+1)}^+\cdots X_{(n-1)}^+$$

and

$$L_i^- := X_{(i)}^-X_{(i+1)}^-\cdots X_{(n-1)}^-$$

are groups. Suppose that $L_{i+1}(U) = L_{i+1}^+L_{i+1}^-$. It follows from Theorem 5.8 that $L_i(U) = X_{(i)}^+X_{(i)}^-L_{i+1}(U)$ and hence $L_i(U) = L_i^+L_i^-$. By induction this holds for all i. Since $U_w^+ = L_1^+$ and $U_w^- = L_1^-$, we have $U = U_w^+U_w^-$.

It follows from these calculations that $L_i(U)/L_{i+1}(U)$ is isomorphic to the direct product $L_i^+/L_{i+1}^+ \times L_i^-/L_{i+1}^-$ and hence, by induction, we have $L_i^+ \cap L_i^- = 1$. In particular, $U_w^+ \cap U_w^- = 1$.

Finally, observe that $n_w X_{ij} n_w^{-1} = X_{w(i),w(j)}$ and hence $n_w U_w^+ n_w^{-1} = U_{w^{-1}}^+$. □

Whilst reading the above proof it may be helpful to choose a fairly simple example for w and to contemplate the matrix forms of the various groups encountered.

The final result is a 'normal form' theorem for elements of $SL(V)$.

5.10 Theorem. *Each element $g \in SL(V)$ can be expressed uniquely in the form $g = bn_wu$, where $b \in B$, $w \in W$ and $u \in U_w^-$.*

Proof. By (5.1) we have $g \in Bn_wB$ for some $w \in W$. Now $B = (B \cap N)U$ and N normalizes $B \cap N$, so $Bn_wB = Bn_wU$. From Lemma 5.9 we have $U = U_w^+U_w^-$ and $n_wU_w^+n_w^{-1} \subseteq B$. Thus $Bn_wB = Bn_wU_w^-$ and this establishes the required form for g.

To prove uniqueness, suppose $bn_wu = b'n_{w'}u'$, where b, $b' \in B$, w, $w' \in W$ and u, $u' \in U$. By Theorem 5.5 we have $w = w'$ and consequently

$$b^{-1}b' = n_wuu'^{-1}n_w^{-1} \in n_wU_w^-n_w^{-1} \cap B = 1.$$

That is, $b = b'$ and $u = u'$. □

EXERCISES

5.1 If $\dim V = m$, a *simplex* in $\mathcal{P}(V)$ is a set of $m+1$ points with the property that each m element subset is a frame. Show that $PGL(V)$ acts regularly on the set of ordered simplexes.

5.2 Show that two flags of $\mathcal{P}(V)$ are in the same orbit of $PSL(V)$ if and only if they have the same type.

5.3 Show that the flags of the apartment of $\{\,\langle e_1\rangle, \dots, \langle e_m\rangle\,\}$ may be identified with the ordered partitions of the set $\{1, 2, \dots, m\}$ and that under this identification the chambers correspond to the $m!$ linear orderings of $1, 2, \dots, m$. In particular, show that the Weyl group S_m acts regularly on the chambers of the apartment.

5.4 Define two chambers to be *adjacent* if they have a common panel. Show that two chambers of an apartment of $SL(V)$ are adjacent if and only if they are interchanged by an element of the Weyl group corresponding to a transposition.

5.5 A *gallery* from the chamber M to the chamber M' is a sequence of chambers $M = M_0, M_1, \dots, M_k = M'$ such that M_{i-1} is adjacent to M_i for $i = 1, 2, \dots, k$. Show that in the building of $SL(V)$ any two chambers can be connected by a gallery.

5.6 If n_1 and n_2 are monomial matrices and b is an upper triangular matrix such that $n_1^{-1} b n_2$ is upper triangular, show that $n_1^{-1} n_2$ is a diagonal matrix.

5.7 Let X be a root group of $SL(V)$. Show that for $f, g \in SL(V)$, the group

$$L := \langle\, fXf^{-1}, gXg^{-1} \,\rangle$$

is isomorphic to $\mathbb{F}^+$, $\mathbb{F}^+ \times \mathbb{F}^+$, or $SL(2,\mathbb{F})$ or else $L' = Z(L) \simeq \mathbb{F}^+$ and $L/L' \simeq \mathbb{F}^+ \times \mathbb{F}^+$.

5.8 Let $e_1, e_2, \dots, e_m$ be a basis for V and let $\omega_1, \omega_2, \dots, \omega_m$ be its dual basis. Let B be the stabilizer of the flag

$$\{\, \langle e_1, e_2, \dots, e_i \rangle \mid 1 \le i < m \,\},$$

N the stabilizer of the frame $\{\langle e_1\rangle, \dots, \langle e_m\rangle\}$ and set $H := B \cap N$.

(*i*) If U is the subgroup consisting of the elements of B whose matrices have 1's on the diagonal, show that U is a normal subgroup of B, $U \cap H = 1$ and $B = HU$.

(*ii*) Set $U_i := \{\, f \in U \mid \omega_i f e_{i+1} = 0 \,\}$ and for $i \neq j$ let X_{ij} be the root group consisting of the transvections $t_{\omega_j, a e_i}$, where $a \in \mathbb{F}$. If $X_i := X_{i,i+1}$ and $X_{-i} := X_{i+1,i}$, show that $U = U_i X_i$, $U_i = n_i U n_i^{-1} \cap U$ and $n_i X_i n_i^{-1} = X_{-i}$, where n_i is the element of N which takes e_i to $-e_{i+1}$, e_{i+1} to e_i and fixes every other basis element.

(*iii*) Suppose that the matrix of $n \in N$ corresponds to the permutation $\pi \in S_m$; that is, it has its non-zero entries in positions $(\pi(j), j)$. Show that $n X_{i,j} n^{-1} = X_{\pi(i),\pi(j)}$ and $n_i B \subseteq B n_i X_i$. If $\pi^{-1}(i) < \pi^{-1}(i+1)$, show that $n^{-1} X_i n \subseteq B$ and deduce that $n_i B n \subseteq B n_i n B$.

(iv) Show that $X_{-i} \subseteq B \cup Bn_iB$ and then apply (iii) to n_in in place of n to deduce that when $\pi^{-1}(i) > \pi^{-1}(i+1)$ we have

$$n_iBn \subseteq (Bn_inB) \cup (BnB).$$

(v) If $\mathbb{F} = \mathbb{F}_{p^a}$, where p is a prime, show that U is a Sylow p-subgroup of $SL(V)$.

5.9 Let Γ be a set of subsets of a set Ω such that every pair of elements of Ω belongs to at least one subset in Γ. Suppose that G is a group which acts transitively on Ω and on Γ. Choose $\alpha \in \Omega$, $\Delta \in \Gamma$ and set $B = G_\alpha$ and $N = G_\Delta$. Show that $G = BNB$ if and only if G is transitive on the set of pairs (β, Λ), where $\beta \in \Lambda$ and $\Lambda \in \Gamma$.

5.10 Suppose that M_1 and M_2 are maximal flags of $\mathcal{P}(V)$. Show by induction on $\dim V$ that there is a frame $\mathcal{F}$ such that M_1 and M_2 belong to $\Sigma(\mathcal{F})$.

5.11 Suppose that B and N form a BN-pair for a group G and that

$$W = N/B \cap N = \langle\, w_i \mid i \in I \,\rangle$$

is its Weyl group. If $J \subseteq I$, let N_J be the inverse image of $\langle\, w_i \mid i \in J \,\rangle$ in N. Show that $P_J := BN_JB$ is a subgroup of G. The groups P_J and their conjugates are called the *parabolic* subgroups of G.

5.12 If B and N form a BN-pair for $SL(V)$, show that their images in $PSL(V)$ form a BN-pair for $PSL(V)$ with the same Weyl group.

5.13 Show that each panel of the building $\Delta(V)$ is contained in at least three chambers and show that this result may be regarded as the 'geometric' interpretation of BN-pair axiom (iv)(a).

5.14 Show that every parabolic subgroup of $SL(V)$ is the stabilizer of a flag and that any subgroup of $SL(V)$ that contains the stabilizer of a maximal flag is parabolic. (Use elementary row and column operations.)

5.15 Let $e_{i,j}$ denote the $m \times m$ matrix which has 1 in the (i,j)-th place and 0 elsewhere, and let ξ be a generator of the multiplicative group of $\mathbb{F}_q$. Let x be the transvection $\mathbf{1} + e_{1,2}$, let w be the monomial matrix $e_{1,m} - \sum_{i=1}^{m-1} e_{i+1,i}$, and let h be the diagonal matrix with first two diagonal entries ξ and ξ^{-1} and the rest equal to 1. Show that when $q > 3$,

$$SL(m,q) = \langle\, h, xw \,\rangle$$

and when $q = 2$ or 3,

$$SL(m,q) = \langle\, x, w \,\rangle.$$

5.16 In the notation of Lemma 5.9, show that $U_w^- = U \cap n_w^{-1} U n_w$.

5.17 Use Theorem 5.8 to show that if $PSL(n_1,q_1) \simeq PSL(n_2,q_2)$ and if q_1 and q_2 are powers of the same prime, then $n_1 = n_2$ and $q_1 = q_2$.

5.18 Complete Exercise 4.11 as follows. Suppose that q is a power of an odd prime and that $q \equiv 3 \pmod 4$.

(i) Show that 2^a divides

$$(q^{2i-1} - 1)(q^{2i} - 1)/(q - 1)(q^2 - 1)$$

if and only if 2^a divides i.

(ii) Show that the order of a Sylow 2-subgroup P of $PSL(m,q)$ is strictly less than $(8(q+1))^{m/2}$ and that $|P|$ is less than the order of a Sylow subgroup corresponding to the characteristic of $\mathbb{F}_q$ except when $m = 2$ and $q = 2^s - 1$, for some s.

5.19 Use Exercises 4.11, 5.17 and 5.18 to show that the only isomorphisms between the groups $PSL(m,q)$ are those listed in Theorem 4.6.

6

The 7-Point Plane and the Group $\mathbf{A}_7$

In this chapter we study the isomorphisms

$$PSL(3,2) \simeq PSL(2,7)$$

and

$$PSL(4,2) \simeq A_8$$

mentioned in Theorem 4.6. The purpose is to illustrate the interplay between group theoretic and geometric methods and to provide hints of further developments which can be pursued elsewhere (Cameron (1976), Kantor (1981) and Neumaier (1984)).

The proofs require somewhat more familiarity with finite group theory than is necessary in the rest of the book. On the other hand, the material is not a prerequisite for what follows and this chapter could be skipped on a first reading.

The 7-Point Plane

If V is a vector space of dimension 3 over the field $\mathbb{F}_2$, then the projective plane $\mathcal{P}(V)$ has seven points and seven lines. Each line has 3 points and each point is on 3 lines. This is the 7-*point plane*, also known as the *Fano* plane. From results of the previous chapters, the collineation group of $\mathcal{P}(V)$ is the simple group $P\Gamma L(3,2) = PSL(3,2)$ of order 168.

We shall show that if Ω is a set of seven elements and if $\mathbb{B}$ is a set of seven 3-element subsets of Ω such that any pair of distinct elements of $\mathbb{B}$ have exactly one element of Ω in common, then Ω and $\mathbb{B}$ can be identified with the points and the lines of the 7-point plane.

First observe that the conditions imply that any two distinct elements P, Q of Ω belong to a unique element, say L, of $\mathbb{B}$. We let $P+Q$ denote the third element of L. Next, set

$$W := \Omega \cup \{0\}$$

and make W into a vector space over $\mathbb{F}_2$ by defining $P+Q$ as above, $P+0 = 0+P = P$ and $P+P=0$, for all $P \in W$.

Associativity is easy to check and consequently W is a vector space isomorphic to V. The points of $\mathcal{P}(W)$ may be identified with Ω and the lines with $\mathbb{B}$. We use this construction in the proof of the next theorem.

The Simple Group of Order 168

Recall that if H is a *subset* of a group G, the *normalizer* of H in G is the subgroup

$$N_G(H) := \{\, g \in G \mid gHg^{-1} = H \,\}$$

and the *centralizer* of H in G is the subgroup

$$C_G(H) := \{\, g \in G \mid ghg^{-1} = h \text{ for all } h \in H \,\}.$$

If G is finite, the number of conjugates of H in G is $|G : N_G(H)|$ and if H is a Sylow p-subgroup of G, then

$$|G : N_G(H)| \equiv 1 \pmod{p}.$$

Also, if H is any subgroup of G, then G acts transitively on the set of left cosets gH of H. The elements of G act by multiplication on the left.

6.1 Theorem. *If G is a simple group of order 168, then G is isomorphic to $PSL(3,2)$.*

Proof. Let P be a Sylow 7-subgroup of G. Then by Sylow's theorem, $|G : N_G(P)| \equiv 1 \pmod 7$, and as P is not a normal subgroup of G it follows that $|N_G(P)| = 21$. Now G acts transitively on the 8 cosets of $N_G(P)$ and as G is simple we may regard it as a subgroup of A_8. The Sylow 7-subgroups of A_8 are self-centralizing and therefore $C_G(P) = P$ (see Exercise 1.17). We have $N_G(P) = PQ$, where Q is a Sylow 3-subgroup of $N_G(P)$ and of G. Since $C_G(P) = P$, 7 cannot divide $|N_G(Q)|$ and as $|G : N_G(Q)| \equiv 1 \pmod 3$ we have $|G : N_G(Q)| = 7$ or 28.

If $|G : N_G(Q)| = 7$, then G may be regarded as a subgroup of A_7. The only coset of $N_G(Q)$ fixed by Q is $N_G(Q)$ itself and therefore, from the structure of A_7, $C_G(Q) = Q$. But $|G : N_G(Q)| = 7$ implies $|N_G(Q)| = 2^3.3$, a contradiction. It follows that $|G : N_G(Q)| = 28$ and therefore $|N_G(Q)| = 6$. This means that $N_G(Q) = Q\langle\, t\,\rangle$, where $t^2 = 1$.

If $C_G(Q) = N_G(Q)$, then $N_G(Q) \subset C_G(t)$. From Sylow's theorem we have $|C_G(t) : N_G(Q)| \equiv 1 \pmod 3$ and so $|C_G(t)| = 24$ since 4 divides $|C_G(t)|$. But then $|G : C_G(t)| = 7$ and again G can be regarded as a subgroup of A_7. In A_7 an element of order 3 which normalizes an element of order 7 has only one fixed point and cannot commute with an element of order 2. This contradiction forces $C_G(Q) = Q$.

So far we have found 48 elements of order 7, 56 elements of order 3 and 1 element of order 1. No element of order 2 can commute with any element of order 3 or 7 and A_8 does not contain any elements of order 8. Therefore the remaining 63 elements have order 2 or 4.

Suppose that S is a Sylow 2-subgroup of G. If $S \neq N_G(S)$, it follows that $|G : N_G(S)| = 7$ and yet again G can be considered as a subgroup of A_7. But the Sylow 2-subgroups of A_7 are self-normalizing. Hence $S = N_G(S)$.

In order to proceed we need the following lemma of Burnside:

6.2 Lemma. *If X and Y are normal subsets of a Sylow p-subgroup P of G and if $gXg^{-1} = Y$ for some $g \in G$, then $nXn^{-1} = Y$ for some $n \in N_G(P)$.*

Proof. We have $P \subseteq N_G(X)$ and therefore both P and gPg^{-1} are Sylow p-subgroups of $N_G(Y)$. By Sylow's theorem there is an element $h \in N_G(Y)$ such that $hgPg^{-1}h^{-1} = P$. But then $n := hg$ belongs to $N_G(P)$ and $nXn^{-1} = Y$. □

We revert to the proof of the theorem. If S were abelian, it would follow from the lemma that no two elements of S could be conjugate. But for each element $x \in S$, $x \neq 1$, we would have $C_G(x) = S$ and hence x would have 21 conjugates. Then G would contain more than 63 2-elements, a contradiction. It follows that S is non-abelian and therefore it contains an element x of order 4. We must have $C_G(x) = \langle x \rangle$ and $C_G(x^2) = S$. Thus G has one class of elements of order 2 (containing 21 elements) and one class of elements of order 4 (containing 42 elements). If y is any element of order 4 in S, both $\langle x \rangle$ and $\langle y \rangle$ are normal in S and y is conjugate to x in G. It follows from the lemma that $\langle x \rangle = \langle y \rangle$. This means that S is the dihedral group D_8.

Now S has exactly two subgroups A and B isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. Since S is not abelian, $C_G(A) = A$ and $C_G(B) = B$. Since A and B are normal in S, the lemma shows that they cannot be conjugate in G. But all elements of order 2 are conjugate in G and therefore both $N_G(A)$ and $N_G(B)$ are isomorphic to the symmetric group S_4.

Let Ω be the set of conjugates of A and let $\mathbb{B}$ be the set of conjugates of B. Define $X \in \Omega$ to be incident with $Y \in \mathbb{B}$ whenever $\langle X, Y \rangle$ is a Sylow 2-subgroup of G. As S_4 has 3 Sylow 2-subgroups, each element of $\mathbb{B}$ is incident with 3 elements of Ω and each element of Ω is incident with 3 elements of $\mathbb{B}$. The group G acts by conjugation on Ω and $\mathbb{B}$ and preserves the relation of incidence. In order to show that Ω and $\mathbb{B}$ represent points and lines of the 7-point plane it is enough to show that for each $g \notin N_G(B)$, there is a unique element of Ω incident with both B and gBg^{-1}. Every element of order 3 acts on $\mathbb{B}$ as a product of two disjoint 3-cycles and therefore $N_G(B)$ is transitive on $\mathbb{B} \setminus \{B\}$. Thus for $g \notin N_G(B)$, $N_G(B) \cap N_G(gBg^{-1})$ is a group of order 4, namely the unique element of Ω incident with B and gBg^{-1}. This shows that Ω and $\mathbb{B}$ form the points and lines of a 7-point plane and hence $G \simeq PSL(3,2)$. □

6.3 Corollary. $PSL(2,7) \simeq PSL(3,2)$. □

This provides us with the isomorphism whose existence was asserted in Theorem 4.6 (*iii*). The rest of the chapter is devoted to obtaining the isomorphism $A_8 \simeq PSL(4,2)$ by studying the action of A_7 on a set of 7-point planes.

The remaining isomorphisms of Theorem 4.6 will be dealt with in the exercises at the end of the chapter.

A Geometry of 7-Point Planes

From now on let Ω be a set of seven elements and denote the set of all 3-element subsets of Ω by $\mathcal{L}$. The 35 elements of $\mathcal{L}$ will be called *Lines*. At this stage the elements of $\mathcal{L}$ are simply the 3-element subsets of Ω. But ultimately we shall identify them with the lines of a projective geometry. This is why we call them Lines (with a capital L).

A 7-*point plane* is a set $\mathbb{B}$ of seven Lines such that any two distinct Lines of $\mathbb{B}$ have exactly one element of Ω in common. A *fan* with vertex $\alpha \in \Omega$ is a set of three Lines such that any two of them have only α in common. There are 105 fans and each fan can be extended to a 7-point plane in exactly two ways, hence there are 30 7-point planes on Ω.

The group S_7 acts on Ω and permutes the 7-point planes. We have seen that the subgroup of S_7 which fixes a 7-point plane is $PSL(3,2)$ and as $|S_7 : PSL(3,2)| = 30$, it follows that S_7 is transitive on the set of all 7-point planes. Since $PSL(3,2) \subset A_7$ and $|A_7 : PSL(3,2)| = 15$, it follows that A_7 has two orbits on the 7-point planes and both orbits have length 15.

Recall from the proof of Theorem 6.2 that $PSL(3,2)$ has precisely two conjugacy classes of subgroups of index 7 and that the orbit lengths of a group of one class acting on the other class are 3 and 4. The orbits of length 3 form the lines of a 7-point plane and it follows that in S_7 there is only one conjugacy class of subgroups isomorphic to $PSL(3,2)$ whereas in A_7 there are two.

Let $\mathcal{P}$ and $\mathcal{H}$ denote the two orbits of A_7 on 7-point planes. Call the elements of $\mathcal{P}$ Points and the elements of $\mathcal{H}$ Planes. (Note the capitals.) We shall build a 'geometry' from $\mathcal{P}$, $\mathcal{L}$ and $\mathcal{H}$ and prove that these sets are the points, lines, and planes of the projective geometry of a vector space of dimension 4 over $\mathbb{F}_2$.

To build a geometry is to define an incidence relation. We declare a Line $\lambda \in \mathcal{L}$ to be *incident* with $\mathbb{B} \in \mathcal{P} \cup \mathcal{H}$ whenever $\lambda \in \mathbb{B}$. The Point $\mathbb{B}_1$ is *incident* with the Plane $\mathbb{B}_2$ whenever $\mathbb{B}_1 \cap \mathbb{B}_2$ is a fan. As a first step towards identifying the 'geometry' of $\mathcal{P}$, $\mathcal{L}$ and $\mathcal{H}$ we investigate the incidence relations in some detail. Note that A_7 acts transitively on each of the sets $\mathcal{P}$, $\mathcal{L}$ and $\mathcal{H}$ and preserves the incidence relations. The elements of S_7 that are not in A_7 interchange $\mathcal{P}$ and $\mathcal{H}$.

6.4 Lemma. (i) *Each Line is incident with 3 Points and 3 Planes.*

(ii) *A Point $\mathbb{B}_1$ is incident with a Plane $\mathbb{B}_2$ if and only if $\mathbb{B}_1$ and $\mathbb{B}_2$ are incident with a common Line.*

Proof. Let λ be a Line. If $\mathbb{B}$ is a 7-point plane that contains λ, each element $\alpha \in \lambda$ belongs to two Lines of $\mathbb{B}$ other than λ. These two Lines partition $\Omega \setminus \lambda$ into two sets of size 2. Thus $\mathbb{B}$ determines a bijection between the elements of λ and the three partitions of $\Omega \setminus \lambda$ into two sets of size 2. Conversely, each such bijection arises from a unique 7-point plane which contains λ. Hence λ is contained in six 7-point planes.

Let L be the stabilizer of λ in S_7. Then $L = R \times T$, where $R \simeq S_3$ fixes $\Omega \setminus \lambda$ pointwise and $T \simeq S_4$ fixes λ pointwise. The group L acts on the six 7-point planes containing λ and the transpositions of L interchange the 7-point planes of $\mathcal{P}$ with those of $\mathcal{H}$. It follows that λ is incident with three Points and with three Planes. This proves (i).

Each transposition of L fixes a fan containing λ and, in particular, the action of the three transpositions of R makes each Point containing λ incident with the three Planes containing λ. This proves (ii). $\square$

If $\mathbb{B}$ is a 7-point plane, the stabilizer $H \simeq PSL(3,2)$ of $\mathbb{B}$ in S_7 has two orbits on $\mathcal{L}$: the 7 Lines of $\mathbb{B}$ and the 28 Lines not in $\mathbb{B}$. Thus

$$H \cap L \simeq \begin{cases} S_3 & \lambda \in \mathbb{B} \\ S_4 & \lambda \notin \mathbb{B}. \end{cases}$$

It follows that L has two orbits on 7-point planes: the 6 planes that contain λ and the 24 that do not.

6.5 Theorem. *The Points $\mathcal{P}$, the Lines $\mathcal{L}$ and the Planes $\mathcal{H}$ can be identified with the points, lines and planes of the projective geometry of a vector space of dimension 4 over $\mathbb{F}_2$ so that the incidence relations become the usual containment of subspaces.*

Proof. If $\mathbb{B}$ is a Point, there are 7 fans in $\mathbb{B}$, hence 7 Planes incident with $\mathbb{B}$. Also, for each Line of $\mathbb{B}$, there are exactly two points other than $\mathbb{B}$ incident with it. This accounts for all 15 Points and it follows that each pair of distinct

Points is incident with a unique Line. Dually, each pair of distinct Planes is incident with a unique Line.

This means that we can make the set $V := \mathcal{P} \cup \{0\}$ into a vector space of dimension 4 over $\mathbb{F}_2$ by defining $\mathbb{B} + \mathbb{B} = 0$, $0 + \mathbb{B} = \mathbb{B} + 0 = \mathbb{B}$ and (for $\mathbb{B}_1 \neq \mathbb{B}_2$) $\mathbb{B}_1 + \mathbb{B}_2$ to be the third Point incident with the Line determined by $\mathbb{B}_1$ and $\mathbb{B}_2$. Associativity follows from the fact that two Lines with a common Point are incident with a common Plane. By construction, $\mathcal{P}$, $\mathcal{L}$ and $\mathcal{H}$ correspond to the 1-, 2- and 3-dimensional subspaces of V. □

6.6 Theorem. *We have $A_7 \subset PGL(4,2)$ and A_7 acts doubly transitively on both the points and planes of the projective geometry of dimension 3 over $\mathbb{F}_2$.*

Proof. The group A_7 acts transitively on $\mathcal{P}$ and on the set of incident Point-Line pairs. The stabilizer in A_7 of a Line acts as S_3 on the Line and as two distinct Points are on a unique Line it follows that A_7 is doubly transitive on $\mathcal{P}$. By symmetry, A_7 is doubly transitive on $\mathcal{H}$ as well. □

6.7 Corollary. $PSL(4,2) = PGL(4,2) \simeq A_8$.

Proof. From the order formula of Chapter 4, the index of A_7 in $PSL(4,2)$ is 8. But $PSL(4,2)$ is simple and therefore $PSL(4,2) \simeq A_8$. □

A Geometry for A_8

The isomorphism $PSL(4,2) \simeq A_8$ suggests that there should be a connection between the geometry of Points, Lines and Planes constructed in the previous section and a set of size *eight.* Indeed there is.

The set of size eight is the set $W := \Omega \cup \{0\}$ considered at the beginning of the chapter. We showed that for each 7-point plane $\mathbb{B}$ on Ω, there is a vector space structure on W such that $\mathbb{B} = \mathcal{P}(W)$. Instead of the 7-point planes on Ω we now look at the 8-point affine spaces on W. We define an *8-point affine space* on W to be a set $\mathcal{A}$ of 4-element subsets of W such that every 3-element subset of W is contained in a unique element of $\mathcal{A}$. It is easy to see that $\mathcal{A}$ consists of 14 4-element sets and each point of W is in 7 of them. Moreover, the 7 elements of $\mathcal{A}$ that contain 0 correspond to a 7-point plane on Ω. It follows that $\mathcal{A}$ consists of the 2-dimensional affine subspaces of the vector space structure induced by this 7-point plane on W. In addition, this establishes a one-to-one correspondence between the 8-point affine spaces and the 7-point planes on Ω. To pass from a 7-point plane $\mathbb{B}$ to the corresponding 8-point affine space $\mathcal{A}$, adjoin 0 to each of the lines of $\mathbb{B}$ and declare the complement of each line of $\mathbb{B}$ to be a line of $\mathcal{A}$.

The objects corresponding to the 3-element subsets of Ω (which we called Lines) are the 35 partitions of W into two sets of size 4. The Line $\ell \in \mathcal{L}$

corresponds to the partition $\{\ell \cup \{0\}, \Omega \setminus \ell\}$. Thus the Line ℓ belongs to a 7-point plane if and only if the two blocks of the partition are lines in the corresponding 8-point affine space.

The fans of Ω correspond to the 105 partitions of W into four sets of size 2. Thus two 8-point spaces $\mathcal{A}_1$ and $\mathcal{A}_2$ are incident if $\mathcal{A}_1 \cap \mathcal{A}_2$ consists of the six 4-element sets obtained by taking the union of pairs of blocks of such a partition.

This completes the description of the geometry in terms of W. It follows that S_8 acts transitively on the 8-point spaces and on the sets of partitions, preserving incidence. The group A_8 has two orbits on the 8-point spaces—corresponding to the sets $\mathcal{P}$ and $\mathcal{H}$ of 7-point planes on Ω. Thus A_8 is indeed the collineation group of the projective geometry constructed from $\mathcal{P}$, $\mathcal{L}$ and $\mathcal{H}$.

The stabilizer of an 8-point space (in either A_8 or in S_8) is the affine group $\mathrm{Aff}(W) = T(W)PSL(3,2)$.

The elements of S_8 not in A_8 preserve incidence but interchange Points and Planes. In the next chapter we begin a detailed study of maps with this property.

It has been shown by Cameron and Kantor (1979) that if $G \subseteq P\Gamma L(n,q)$ acts doubly transitively on the points of the projective geometry, then either $G \simeq A_7$ (and $n = 4$, $q = 2$) or G contains $PSL(n,q)$. For earlier work, including a proof that A_7 is doubly transitive on the projective geometry, see Wagner (1961).

EXERCISES

6.1 Let $\mathbb{B}$ be a set of seven 3-element subsets of the set Ω of seven elements. Suppose that any pair of distinct elements of $\mathbb{B}$ have at most one element in common. Show that each 2-element subset of Ω is contained in a unique element of $\mathbb{B}$. Deduce that Ω and $\mathbb{B}$ are the points and lines of a 7-point plane.

6.2 Describe the two 7-point planes that contain a given fan.

6.3 Using the notation of this chapter, define a graph on the set $\mathcal{P} \cup \mathcal{L}$ as follows. Join $\mathbb{B} \in \mathcal{P}$ to each of the 7 elements of $\mathcal{L}$ incident with it. Join $\lambda \in \mathcal{L}$ to the 3 elements of $\mathcal{P}$ incident with it and to the 4 elements of $\mathcal{L}$ disjoint from it. Show that the graph has no circuits of length 3 or 4 and that any two vertices which are not joined belong to 6 circuits of length 5. This is a *Moore* graph (also known as the *Hoffman-Singleton* graph).

6.4 Let G be either a simple group of order $\frac{1}{2}p(p-1)(p+1)$, where p is an odd prime and let U be a Sylow p-subgroup, or let $G = A_6$ and let U be a Sylow 3-subgroup. In all cases set $q = |U|$. The aim of this exercise is to show that $G \simeq PSL(2,q)$.

(*i*) Show that G acts doubly transitively (by conjugation) on the set $\mathcal{P}$ of conjugates of U and that $|\mathcal{P}| = q+1$.

(*ii*) Let $B = N_G(U)$ and show that $B = UH$, where H is a cyclic group of order $\frac{1}{2}(q-1)$.

(*iii*) Show that only the identity element of G fixes more than two elements of $\mathcal{P}$.

(*iv*) Let $N = N_G(H)$ and show that $N = HW$, where $W = \langle\, t \,\rangle$, $t^2 = 1$ and $tUt \cap U = 1$.
(*a*) If $q \equiv 3 \pmod 4$, show that t fixes no element of $\mathcal{P}$ and that $C_H(t) = 1$.
(*b*) If $q \equiv 1 \pmod 4$, show that t fixes two elements of $\mathcal{P}$ and that $|C_H(t)| = 2$.
In both cases show that $tht = h^{-1}$ for all $h \in H$.

(*v*) Show that $G = BNB$ and that $BtB = UtB$.

(*vi*) Let the elements of U be $\{\, u_a \mid a \in K \,\}$, where $u_0 = 1$. Make K into an (additive) group by defining $a+b$ so that $u_a u_b = u_{a+b}$. Let U_0 be the element of $\mathcal{P} \setminus \{U\}$ fixed by H and define U_a to be $u_a U_0 u_a^{-1}$. Show that $u_b U_a u_b^{-1} = U_{a+b}$ and that $tUt = U_0$.

(*vii*) Choose the notation so that U_1 denotes an element of $\mathcal{P}\setminus\{U, U_0\}$ not fixed by t. If $h \in H$ and $hU_1h^{-1} = U_a$, write h_a in place of h and define ab by the prescription $h_a u_b h_a^{-1} = u_{ab}$. Show that h_1 is the identity element, $h_a U_b h_a^{-1} = U_{ab}$ and $a(b+c) = ab+ac$ whenever h_a is defined.

(*viii*) Let $\mathbb{F}$ be the set of group homomorphisms $f : K \to K$ such that $f(ab) = af(b)$ whenever h_a is defined. Define $f+g$ and fg by

$$(f+g)(a) = f(a) + g(a) \quad \text{and} \quad (fg)(a) = f(g(a)).$$

Show that $\mathbb{F}$ is a field and that $f \mapsto f(1)$ is a group isomorphism from $\mathbb{F}^+$ onto K such that $fg \mapsto f(1)g(1)$ whenever the product is defined in K. From now on identify K with $\mathbb{F}_q$ and note that h_a is defined if and only if a is a square.

(ix) Show that there are unique functions α, β and γ from $K \setminus \{0\}$ to itself such that for $a \neq 0$, $tu_a t = u_{\alpha(a)} t u_{\beta(a)} h_{\gamma(a)}$.

(x) Show that $tU_a t = U_{\alpha(a)}$ and that $\alpha(ab) = a^{-1}\alpha(b)$, whenever h_a is defined.

(xi) Suppose that $tU_1 t = U_d$ and then show that $-d$ is a square and that by replacing t by th_{-d} we may assume that $d = -1$.

(xii) Consider $tu_a t$ and show that

$$\beta(a) = -\alpha(-a)\gamma(a) \text{ and } \gamma(a) = \gamma(-a).$$

($xiii$) Consider $tu_{\alpha(a)} t$ and deduce that

$$\alpha^2(a) = a, \ \beta(\alpha(a)) = -\beta(a)\gamma(a)^{-1} \text{ and } \gamma(\alpha(a)) = \gamma(a)^{-1}.$$

(xiv) Show that when -1 is not a square in $\mathbb{F}_q$, then

$$\alpha(a) = -a^{-1}, \ \beta(a) = -a \text{ and } \gamma(a) = a^2.$$

(xv) Suppose that -1 is a square in $\mathbb{F}_q$ and show that $\beta(a) = -a$. In addition, if a and b are squares, apply $tu_a t$ to $U_{\alpha(b)}$ and deduce that $\alpha(a+b) = (a+b)^{-1}$ provided $a + b \neq 0$. Now show that for all a, we have $\alpha(a) = -a^{-1}$ and $\gamma(a) = a^2$.

(xvi) Identify $\mathcal{P}$ with $\mathbb{F}_q \cup \{\infty\}$ by making U correspond to ∞ and U_a correspond to a. Show that with this identification, u_a, h_b and t become the transformations $z \mapsto z + a$, $z \mapsto bz$ and $z \mapsto -z^{-1}$, respectively. Deduce that G is isomorphic to $PSL(2, q)$.

(For further hints and a considerably more general result, consult Zassenhaus (1936) or Huppert and Blackburn (1982), Chapter XI.)

6.5 Using the previous exercise, deduce that a simple group of order 60 is isomorphic to A_5, a simple group of order 168 is isomorphic to $PSL(3, 2)$ and that A_6 is isomorphic to $PSL(2, 9)$.

6.6 Let p be a prime, set $q = p^r$ and suppose that $|PSL(n, q)| = |A_m|$. The aim of this exercise is to show that the only solutions for (n, q, m) are (2,3,4), (2,4,5), (2,5,5), (2,9,6), (3,4,8) and (4,2,8).

(i) As a preliminary step show that, in general, if U is a Sylow p-subgroup of $PSL(n, q)$, then the order of $PSL(n, q)$ is less than $|U|^{2(n+1)/n}$. Also show that the order of a Sylow p-subgroup of A_m is less than $p^{m/(p-1)}$.

(*ii*) Show (by induction on m) that $m^m e^{-m} < \frac{1}{2}m!$, where $e = 2.71828\cdots$ and then deduce that $m < ep^{2(n+1)/n(p-1)}$.

(*iii*) Show that for $x > 1$, the function $x^{1/(x-1)}$ is decreasing and deduce that if $p \geq 5$, then $m \leq e.5^{3/4}$ and hence $m \leq 9$ and the only solution for (n, q, m) is (2,5,5). In this case $PSL(2,5) \simeq A_5$.

(*iv*) Show that is $p = 3$, then $m \leq 14$ and hence $rn(n-1) \leq 10$. Show that the solutions for (n, q, m) in this case are (2,3,4) and (2,9,6), corresponding to the isomorphisms $PSL(2,3) \simeq A_4$ and $PSL(2,9) \simeq A_6$.

(*v*) Finally, suppose that $p = 2$. If $n = 2$, show that the only solution is the one corresponding to the isomorphism $PSL(2,4) \simeq A_5$. If $n \geq 3$, show that there are just two solutions: (4,2,8) arising from the isomorphism $PSL(4,2) \simeq A_8$ and (3,4,8), which does *not* correspond to an isomorphism. (For further information see Artin (1955b).)

6.7 Let Ω be a set of seven elements and consider the Points and Lines of Ω as defined in this chapter. For $\alpha \in \Omega$ show that the 15 Points and the 15 Lines which contain α form a *generalized quadrangle*—that is, show that they satisfy the following axioms with $s = t = 2$.

(*i*) Each Point is incident with $t + 1$ Lines and two distinct Points are incident with at most one Line.

(*ii*) Each Line is incident with $s + 1$ Points and two distinct Lines are incident with at most one Point.

(*iii*) If $\mathbb{B}$ is a Point and λ is a Line not incident with $\mathbb{B}$, there is a unique Point $\mathbb{B}'$ and a unique Line λ' incident with $\mathbb{B}'$ such that $\mathbb{B}$ is incident with λ' and $\mathbb{B}'$ is incident with λ.

6.8 Show that the Points and Lines of the generalized quadrangle described in the previous exercise correspond to the 15 partitions of $\Omega \setminus \{\alpha\}$ into three subsets of two elements and the 15 2-element subsets of $\Omega \setminus \{\alpha\}$, respectively. Relate this to the geometry for A_8.

7

Polar Geometry

In addition to the groups studied in Chapter 4 there are essentially three further types of classical groups: symplectic, unitary and orthogonal. These groups correspond to three types of bilinear form which can be defined on a vector space. In preparation for a more detailed study of the groups in the following pages, this chapter is devoted to the classification of the forms (the Birkhoff–von Neumann (1936) theorem) and to a proof of a general result known as Witt's theorem. For a while we shall return to the assumptions of Chapter 3. That is, K will denote a division ring that is not necessarily commutative and V will denote a finite-dimensional left vector space over K.

The Dual Space

The set V^* of linear functionals $\varphi : V \to K$ is naturally a *right* vector space over K. However, we would like V^* to be a *left* vector space. To achieve this we introduce the ring K^{op} (called the *opposite ring* of K) which has the same elements as K but in which multiplication is defined by

$$a \circ b = ba.$$

Then V^* becomes a *left* vector space over K^{op} provided we define

$$(\varphi_1 + \varphi_2)(v) = \varphi_1(v) + \varphi_2(v)$$

and

$$(a\varphi)(v) = \varphi(v)a.$$

We call this left vector space V^* the *dual* space of V. Of course if K is a field, then $K = K^{\mathrm{op}}$ and V^* is the usual dual space as described in Chapter 4. In any case, if $e_1, e_2, \ldots, e_n$ is a basis for V, the dual basis for V^* is $\omega_1, \omega_2, \ldots, \omega_n$, where

$$\omega_i(a_1e_1 + a_2e_2 + \cdots + a_ne_n) = a_i$$

Therefore $\dim_{K^{\mathrm{op}}} V^* = \dim_K V$.

If X is a subset of V, the *annihilator* of X is the subspace

$$X^\circ := \{\, \varphi \in V^* \mid \varphi(x) = 0 \text{ for all } x \in X \,\}.$$

Dually, the annihilator of a subset S of V^* is

$$S^\circ := \{\, v \in V \mid \varphi(v) = 0 \text{ for all } \varphi \in S \,\}.$$

If $S \subseteq T \subseteq V^*$, then $T^\circ \subseteq S^\circ$ and we have $S \subseteq S^{\circ\circ}$. Similarly, if $X \subseteq V$, then $X \subseteq X^{\circ\circ}$.

If X is a subspace of V, restriction of linear functionals to X defines a linear transformation $V^* \to X^*$ with kernel X°. This map is surjective, hence $V^*/X^\circ \simeq X^*$ and therefore $\dim V^* - \dim X^\circ = \dim X^*$. We know that $\dim V = \dim V^*$ and consequently we have the simple but important relation:

$$\dim X + \dim X^\circ = \dim V.$$

Similarly, if S is a subspace of V^* and $v \in S^\circ$, evaluation at v defines an element of $(V^*/S)^*$ and the resulting map from S° to $(V^*/S)^*$ is an isomorphism. Consequently,

$$\dim S + \dim S^\circ = \dim V$$

and we have $\dim S = \dim S^{\circ\circ}$. Since $S \subseteq S^{\circ\circ}$, it follows that $S = S^{\circ\circ}$ and similarly $X = X^{\circ\circ}$.

Correlations

From now on we shall assume that $\dim V \geq 3$. Recall that the projective geometry $\mathcal{P}(V)$ is the set of all subspaces of V, partially ordered by inclusion, and a collineation of $\mathcal{P}(V)$ is an order-preserving bijection. A *correlation* of $\mathcal{P}(V)$ is a bijection from $\mathcal{P}(V)$ to $\mathcal{P}(V)$ which *reverses* inclusion. Thus a correlation sends points to hyperplanes and *vice versa*. The composition of two correlations is obviously a collineation and therefore the group $P\mathit{\Gamma} L(V)$ of all collineations is a subgroup of index 2 in the group $P\mathit{\Gamma} L^*(V)$ of all collineations and correlations of $\mathcal{P}(V)$.

The map from $\mathcal{P}(V)$ to $\mathcal{P}(V^*)$ which sends X to X° is a bijection which reverses inclusion. Therefore, if π is a correlation of $\mathcal{P}(V)$, the map

$$\mathcal{P}(V) \to \mathcal{P}(V^*) : X \mapsto \pi(X)^\circ$$

is a collineation. It follows from Theorem 3.1 (the Fundamental Theorem of Projective Geometry) that there is a semilinear map $f : V \to V^*$ with associated isomorphism $\sigma : K \to K^{\mathrm{op}}$ such that $\pi(X)^\circ = f(X)$ for all $X \in \mathcal{P}(V)$. In other words,

$$\pi(X) = \{\, v \in V \mid f(x)v = 0 \text{ for all } x \in X \,\}.$$

The isomorphism σ from K to K^{op} may also be regarded as an *anti-automorphism* of K. These considerations lead us to the next topic.

Sesquilinear Forms

If σ is an anti-automorphism of K, a *σ-sesquilinear form* on V is a map $\beta : V \times V \to K$ such that

$$\beta(u_1 + u_2, v) = \beta(u_1, v) + \beta(u_2, v),$$
$$\beta(u, v_1 + v_2) = \beta(u, v_1) + \beta(u, v_2),$$

and

$$\beta(au, bv) = a\beta(u, v)\sigma(b),$$

for all u, v, u_1, u_2, v_1, $v_2 \in V$ and all a, $b \in K$. If $\sigma = 1$, the form is said to be *bilinear*; note that this forces K to be a field.

A semilinear isomorphism $f : V \to V^*$ induces a σ-sesquilinear form

$$\beta(u, v) := f(v)u.$$

Moreover β is *non-degenerate* in the sense that $\beta(u, v) = 0$ for all u implies $v = 0$, or equivalently, $\beta(u, v) = 0$ for all v implies $u = 0$. Conversely, a non-degenerate σ-sesquilinear form β induces a semilinear map f from V to V^* defined by $f(v) := \beta(-, v)$.

A pair of vectors $(u,\, v)$ such that $\beta(u, v) = 0$ is said to be *orthogonal.* For $X \in \mathcal{P}(V)$, the set

$$X^\perp := \{\, u \in V \mid \beta(u, v) = 0 \text{ for all } v \in X \,\}$$

is called the *orthogonal complement* of X. Note that $X^\perp = f(X)^\circ$.

The correlation of $\mathcal{P}(V)$ corresponding to f sends X to $X^\perp$. The inverse of this correlation corresponds to the semilinear map

$$u \mapsto \sigma^{-1}\beta(u, -).$$

It follows from Theorem 3.1 (*ii*) that a σ-sesquilinear form β and a σ'-sesquilinear form β' induce the *same* correlation of $\mathcal{P}(V)$ if and only if, for some $b \in K$, we have

$$\beta'(u, v) = \beta(u, v)b \qquad \text{and} \qquad \sigma'(a) = b^{-1}\sigma(a)b.$$

Note that since $X^\perp = f(X)^\circ$, we have

$$\dim X + \dim X^\perp = \dim V,$$
$$X \subseteq Y \text{ implies } Y^\perp \subseteq X^\perp,$$
$$(X + Y)^\perp = X^\perp \cap Y^\perp \qquad \text{and} \qquad (X \cap Y)^\perp = X^\perp + Y^\perp.$$

Polarities

A *polarity* of $\mathcal{P}(V)$ is a correlation π of order 2 and the pair $(\mathcal{P}(V), \pi)$ is called a *polar geometry*. The term polar comes from the example of poles and polars with respect to a conic in the usual two-dimensional projective geometry over the real field. This is a special case of the general notion of polarity.

A correlation π with associated σ-sesquilinear form β is a polarity if and only if $\pi = \pi^{-1}$. This is equivalent to π being induced both by $v \mapsto \beta(-, v)$ and by $u \mapsto \sigma^{-1}\beta(u, -)$.

A sesquilinear form β such that $\beta(u, v) = 0$ implies $\beta(v, u) = 0$ for all $u, v \in V$ is said to be *reflexive*. Thus π is a polarity if and only if β is reflexive. And β is reflexive if and only if $X = X^{\perp\perp}$ for all $X \in \mathcal{P}(V)$.

7.1 Theorem (Birkhoff–von Neumann). *If* $\dim V \geq 3$ *and if* π *is a polarity of* $\mathcal{P}(V)$, *then* π *arises from a non-degenerate reflexive* σ*-sesquilinear form* β *of one of the following types:*

(i) *Alternating. In this case* K *is a field,* $\sigma = 1$ *and* $\beta(v, v) = 0$ *for all* $v \in V$.

(ii) *Symmetric. In this case* K *is a field,* $\sigma = 1$ *and* $\beta(u, v) = \beta(v, u)$ *for all* $u, v \in V$.

(iii) *Hermitian. In this case* $\sigma^2 = 1$, $\sigma \neq 1$ *and* $\beta(u, v) = \sigma\beta(v, u)$ *for all* $u, v \in V$.

Proof. Let β be a σ-sesquilinear form which induces π. Then the semilinear maps $v \mapsto \beta(-, v)$ and $u \mapsto \sigma^{-1}\beta(u, -)$ both induce π. Therefore, by Theorem 3.1, for some $\lambda \in K$ we have $\beta(u, v) = \sigma^{-1}\beta(v, u)\lambda$ and $\sigma(a) = \lambda^{-1}\sigma^{-1}(a)\lambda$ for all $a \in K$. Thus

$$\beta(u, v) = \sigma^{-1}(\sigma^{-1}\beta(u, v)\lambda)\lambda = \sigma^{-1}(\lambda)\sigma^{-2}\beta(u, v)\lambda$$

and by choosing u and v so that $\beta(u, v) = 1$ we find that $\sigma^{-1}(\lambda)\lambda = 1$. Hence

$$\sigma(\lambda) = \lambda^{-1} \quad \text{and} \quad \sigma^2(a) = \lambda^{-1}a\lambda.$$

If for all $a \in K$, we have $\lambda^{-1}a + \sigma(a) = 0$, then $\lambda = -1$ and $\sigma = 1$. Then K is a field and $\beta(u, v) = -\beta(v, u)$. If the characteristic of K is not 2, then β is alternating and (i) holds. If the characteristic of K is 2, then β is symmetric and (ii) holds.

We may now assume that for some $a \in K$, $b = \lambda^{-1}a + \sigma(a) \neq 0$. Then $\sigma(b) = \sigma(a)\lambda + \lambda^{-1}a\lambda = b\lambda$. If we set $\hat{\beta}(u, v) := \beta(u, v)b$ and $\hat{\sigma}(x) = b^{-1}\sigma(x)b$, then $\hat{\beta}$ is a $\hat{\sigma}$-sesquilinear form which induces π and in addition satisfies $\hat{\beta}(u,v) = \hat{\sigma}\hat{\beta}(v,u)$. If $\hat{\sigma} = 1$, then K is a field and (*ii*) holds. If $\hat{\sigma} \neq 1$, then $\hat{\sigma}^2 = 1$ and (*iii*) holds. □

The polar geometry $(\mathcal{P}(V), \pi)$ is known as a *symplectic*, *orthogonal* or *unitary* geometry according to whether (*i*), (*ii*) or (*iii*) holds.

Note that if $\beta(v,v) = 0$ for all $v \in V$, then on evaluating $\beta(u+v, u+v)$ we have $\beta(u,v) = -\beta(v,u)$ and hence $\sigma(a) = a$ for all $a \in K$. It follows that K is a field and the geometry is symplectic.

Conversely, if the form β satisfies

$$\beta(u,v) = -\beta(v,u) \qquad \text{for all } u,v \in V$$

and if the characteristic of K is not 2, then β is an alternating form.

In the case of a unitary geometry it is possible to choose $d \in K$ so that $d \neq \sigma(d)$. Then $e := d - \sigma(d)$ is *skew-symmetric* (i.e., $\sigma(e) = -e$) and $\beta'(u,v) := \beta(u,v)e$ is a *skew σ'-hermitian* form, where $\sigma'(a) := e^{-1}\sigma(a)e$. That is, $\beta'(u,v) = -\sigma'\beta'(v,u)$ for all $u,v \in V$. Of course β' also induces the polarity π.

Now that we have shown that it is only in a unitary geometry that the division ring can be non-commutative we exclude this case and henceforth assume that the division ring K is actually a field. This is no restriction in the finite case since a well-known theorem of Wedderburn (1905) asserts that every finite division ring is a field. An elementary proof of this result can be found in Taylor (1974). Further information about unitary geometries over non-commutative division rings can be found in Dieudonné (1971).

Quadratic Forms

When the characteristic of K is 2, our definition of an orthogonal geometry as a polar geometry corresponding to a symmetric bilinear form is not quite general enough for our later discussion of orthogonal groups. We rectify this as follows. A *quadratic form* on V is a function $Q : V \to K$ such that

$$Q(av) = a^2 Q(v)$$

and

$$\beta(u,v) := Q(u+v) - Q(u) - Q(v)$$

is a bilinear form.

We say that β is the *polar form* of Q or that Q *polarizes* to β. We now define an *orthogonal geometry* to be a vector space V together with a quadratic form Q which is *non-degenerate* in the sense that its polar form β has the property that $\beta(u,v) = Q(u) = 0$ for all $v \in V$ implies $u = 0$.

When the characteristic of K is not 2 this coincides with our first definition since then $Q(v) := \frac{1}{2}\beta(v, v)$ defines a quadratic form which polarizes to β. However, if the characteristic of K is 2, the polar form of Q is alternating and does not determine Q. Also, if the characteristic of K is 2, there exist symmetric forms that are *not* the polar form of any quadratic form. In general we *exclude* this case from our considerations. That is, we stipulate that when β is symmetric it arises as the polar form of a quadratic form.

Isometries

Suppose that β_1 and β_2 are (possibly degenerate) reflexive σ_1- and σ_2-sesquilinear forms on vector spaces V_1 and V_2 over the fields K_1 and K_2, respectively. A σ-semilinear map $f : V_1 \to V_2$ is called an *isometry* if it is a one-to-one function such that $\sigma_2\sigma = \sigma\sigma_1$ and

$$\beta_2(f(u), f(v)) = \sigma\beta_1(u, v)$$

for all $u, v \in V_1$.

If π_1 and π_2 are the correlations of V_1 and V_2 corresponding to β_1 and β_2, then $\pi_2\varphi = \varphi\pi_1$, where $\varphi := \mathcal{P}(f)$ is the collineation induced by f. We leave it as an exercise to discover to what extent the converse of this observation remains true.

If V_1 and V_2 are provided with quadratic forms Q_1 and Q_2, an *isometry* is defined to be a σ-semilinear map $f : V_1 \to V_2$ such that f is one-to-one and

$$Q_2(f(u)) = \sigma Q_1(u)$$

for all $u \in V_1$.

In all cases a *linear* isometry is an isometry whose associated field automorphism is the identity.

7.2 Lemma. *Suppose that $V_1 = U \oplus W$ and that $f : U \to V_2$ and $g : W \to V_2$ are isometries with the same associated field automorphism $\sigma : K_1 \to K_2$. If $\operatorname{im}(f) \cap \operatorname{im}(g) = \{0\}$ and $\beta_2(f(u), g(w)) = \sigma\beta_1(u, w)$ for all $u \in U$ and $w \in W$, then the map $f + g : V_1 \to V_2$ which sends $u + w$ to $f(u) + g(w)$ is also an isometry.*

Proof. Exercise 7.4. □

Witt's Theorem

Suppose that V is a vector space over a *field* $\mathbb{F}$ and let β be a reflexive σ-sesquilinear form on V. The *radical* of β is the subspace

$$\operatorname{rad} V := V^{\perp} = \{\, u \in V \mid \beta(u, v) = 0 \text{ for all } v \in V \,\}.$$

The form β is non-degenerate if and only if $\operatorname{rad} V = \{0\}$. If β is the polar form of a quadratic form Q, then Q is non-degenerate if and only if 0 is the only element of $\operatorname{rad} V$ which Q maps to 0.

Definitions

(*i*) A non-zero vector u is *isotropic* if $\beta(u, u) = 0$.

(*ii*) A subspace W is *totally isotropic* if $W \subseteq W^{\perp}$.

(*iii*) A non-zero vector u is *singular* if $Q(u) = 0$ and a subspace W is *totally singular* if $Q(u) = 0$ for all $u \in W$.

(*iv*) A pair of vectors (u, v) such that u and v are isotropic and $\beta(u, v) = 1$ is called a *hyperbolic pair*. The line $\langle\, u, v \,\rangle$ in $\mathcal{P}(V)$ is called a *hyperbolic line*.

(*v*) A subspace W is *non-degenerate* if $W \cap W^{\perp} = \{0\}$.

(*vi*) If $V = U \oplus W$ and $\beta(u, w) = 0$ for all $u \in U$ and $w \in W$, we write $V = U \perp W$ and say that V is the *orthogonal* direct sum of U and W.

7.3 Lemma. *Suppose that L is a non-degenerate two-dimensional subspace of V which contains an isotropic vector u with respect to the sesquilinear form β. Then $L = \langle\, u, v \,\rangle$, where (u, v) is a hyperbolic pair. Moreover, if β is the polar form of a quadratic form Q and $Q(u) = 0$, then v may be chosen so that $Q(v) = 0$.*

Proof. We may suppose that $V = L$ and that β is one of the forms listed in Theorem 7.1. Then $L = \langle\, u, w \,\rangle$ for some w such that $a := \beta(u, w) \neq 0$. If β is an alternating form take $v := a^{-1}w$. If β is the polar form of Q, take $v := -Q(w)a^{-2}u + a^{-1}w$.

This leaves the unitary case to be dealt with. As pointed out in a remark following the proof of Theorem 7.1, we may suppose that β is skew σ-hermitian. Choose d so that $d + \sigma(d) \neq 0$ and let

$$c := (d + \sigma(d))^{-1}\beta(w, w)\sigma(d).$$

Then $\beta(w, w) = c - \sigma(c)$ and we may take $v := -(a\sigma(a))^{-1}cu + \sigma(a)^{-1}w$. To complete the proof, note that if the lemma holds for β, then it holds for all multiples of β. □

Suppose that β is a possibly degenerate σ-sesquilinear form on V and that W is a complement to $\operatorname{rad} V$. Then $V = W \perp \operatorname{rad} V$ and Theorem 7.1 applies to the restriction of β to W. It is possible to use β to define a non-degenerate form $\overline{\beta}$ on $V/\operatorname{rad} V$ so that the natural map from W to $V/\operatorname{rad} V$ is an isometry (Exercise 7.7). However, if β is the polar form of a quadratic form Q, it is not always the case that Q induces a quadratic form on $V/\operatorname{rad} V$. The following theorem applies to isometries of quadratic forms and this explains why it is not possible to simply factor out $\operatorname{rad} V$ as the first step in the proof.

7.4 Theorem (Witt). *Suppose that U is a subspace of V and that the map $f : U \to V$ is a linear isometry. Then there is a linear isometry $g : V \to V$ such that $g(u) = f(u)$ for all $u \in U$ if and only if $f(U \cap \operatorname{rad} V) = f(U) \cap \operatorname{rad} V$.*

Proof. If f is the restriction of a linear isometry $g : V \to V$, then it is the case that $f(U \cap \operatorname{rad} V) = f(U) \cap \operatorname{rad} V$. For the converse we suppose that $f(U \cap \operatorname{rad} V) = f(U) \cap \operatorname{rad} V$ and then construct g.

Choose a subspace W that is a common complement in $\operatorname{rad} V$ to $U \cap \operatorname{rad} V$ and $f(U) \cap \operatorname{rad} V$ (Exercise 2.18). Then

$$U + \operatorname{rad} V = U \oplus W \qquad \text{and} \qquad f(U) + \operatorname{rad} V = f(U) \oplus W$$

and it follows from Lemma 7.2 that $f + \mathbf{1}_W$ is a linear isometry which extends f. Thus from now on we may assume that $\operatorname{rad} V \subseteq U$ and $\operatorname{rad} V \subseteq f(U)$. We shall argue by induction on $\dim U - \dim \operatorname{rad} V$. If $U = \operatorname{rad} V$ and if W is a complement to U in V, then $f + \mathbf{1}_W$ is a linear isometry which extends f. Thus we may assume that $U \neq \operatorname{rad} V$.

Let H be a hyperplane of U which contains $\operatorname{rad} V$ and let f' be the restriction of f to H. By induction, f' has an extension $g' : V \to V$. Replacing f by $g'^{-1} f$ we may assume that f fixes every element of H. If f fixes every element of U, we may take $g = \mathbf{1}$. Therefore, from now on assume that this is not the case. Then $P := \operatorname{im}(f - \mathbf{1})$ is a one-dimensional subspace of V.

For u, $v \in U$ we have

$$\begin{aligned}\beta(f(u), f(v) - v) &= \beta(f(u), f(v)) - \beta(f(u), v)\\ &= \beta(u, v) - \beta(f(u), v)\\ &= \beta(u - f(u), v).\end{aligned}$$

Hence $H \subseteq P^\perp$ and it follows that $U \subseteq P^\perp$ if and only if $f(U) \subseteq P^\perp$.

If $U \not\subseteq P^\perp$, then

$$U \cap P^\perp = f(U) \cap P^\perp = H.$$

In this case, let W be a complement to H in $P^\perp$. Then $V = W \oplus U$ and for $w \in W$ and $u \in U$, we have $\beta(w, f(u) - u) = 0$ and therefore $\beta(w, f(u)) = \beta(w, u)$. It follows from Lemma 7.2 that $\mathbf{1}_W + f$ is a linear isometry of V which extends f.

So from now on we may suppose that U and $f(U)$ are contained in $P^\perp$. It follows that $P \subseteq P^\perp$. If $U \neq f(U)$, $u \in U \setminus H$ and $v \in f(U) \setminus H$, then

$X := \langle u+v \rangle$ is a common complement to U and $f(U)$ in $U + f(U)$. Let W be a complement to $U + f(U)$ in $P^\perp$ and set $S := W + X$. Then

$$P^\perp = S \oplus U = S \oplus f(U)$$

and by Lemma 7.2, $\mathbf{1}_S + f$ is a linear isometry of $P^\perp$ which extends f. On the other hand, if $U = f(U)$, let S be any complement to U in $P^\perp$. Then again $\mathbf{1}_S + f$ is a linear isometry of $P^\perp$. In both cases the extension of f to $P^\perp$ acts as the identity on a hyperplane of $P^\perp$ that contains $\operatorname{rad} V$. Thus we may as well suppose that $U = P^\perp = f(U) \neq V$.

Let $P = \langle u \rangle$ and choose $v \in U$ such that $u = f(v) - v$. If β is the polar form of Q, then

$$Q(u) = Q(f(v)) + Q(v) - \beta(f(v), v) = 2Q(v) - \beta(v, v) = 0,$$

and in any case $P \subseteq P^\perp$, whence u is isotropic. If L is a two-dimensional subspace such that $P \subseteq L$ but $L \not\subseteq P^\perp$, then L is non-degenerate. By Lemma 7.3 we can write

$$L = \langle u, w \rangle$$

where (u, w) is a hyperbolic pair (and if β is the polar form of a quadratic form, w is singular).

Since $w \notin \operatorname{rad} V$, $\langle w \rangle^\perp$ is a hyperplane of V and similarly $L^\perp$ is a hyperplane of U. Thus $\langle w \rangle^\perp \cap U = L^\perp$ and we set $Y := f(L^\perp)$. Now $\langle w \rangle + Y$ is a hyperplane of V which contains $\operatorname{rad} V$ but not $f(u)$ and therefore $\langle w \rangle + Y = \langle w' \rangle^\perp$ for some $w' \notin U$. Thus $\langle f(u), w' \rangle$ is non-degenerate and we have $Y = \langle f(u), w' \rangle^\perp$.

By Lemma 7.3 there is an isotropic (singular) vector w'' such that

$$\langle f(u), w' \rangle = \langle f(u), w'' \rangle$$

and $(f(u), w'')$ is a hyperbolic pair.

Define an isometry $g : \langle w \rangle \to V$ by $aw \mapsto aw''$. If $x \in L^\perp$, then

$$\beta(f(x), g(w)) = 0 = \beta(x, w)$$

and we also have

$$\beta(f(u), g(w)) = 1 = \beta(u, w).$$

Since $U = \langle u \rangle \oplus L^\perp$ and $V = \langle w \rangle \oplus U$, it follows from Lemma 7.3 that $g + f$ is a linear isometry of V. This completes the proof. □

This theorem was first proved by Witt (1937) for symmetric and hermitian forms over fields of characteristic other than 2. However, the proof presented here was inspired by the version that appears in the book of Chevalley (1954). Similar proofs can be found in Higman (1978) and Gross (1979).

It is a consequence of this theorem that any two maximal totally isotropic subspaces of V have the same dimension. This common dimension is called the *Witt index* of the form β. If M is a totally isotropic subspace and β is non-degenerate, then

$$M \subseteq M^\perp$$

and the Witt index of β is therefore at most $\frac{1}{2}\dim V$. In general, the index is at most $\frac{1}{2}(\dim V + \dim \operatorname{rad} V)$.

The *Witt index* of a quadratic form is the common dimension of the maximal totally singular subspaces.

Bases of Orthogonal Hyperbolic Pairs

When working with polarities it is possible to choose the basis of V in ways which are better adapted to calculations than is the case for a random basis. The next lemma will help us to make 'nice' choices in later chapters. We assume that V has a non-degenerate reflexive σ-sesquilinear form β (which in the symmetric case is the polar form of a quadratic form Q).

7.5 Lemma. *If U and W are totally isotropic (resp. totally singular) subspaces of V such that $U^\perp \cap W = \{0\}$, then there is a totally isotropic (resp. totally singular) subspace U' containing W such that $V = U^\perp \oplus U'$. Moreover, for each basis $u_1, u_2, \ldots, u_k$ of U, there is a unique basis $u_1', u_2', \ldots, u_k'$ of U' such that (u_1, u_1'), (u_2, u_2'), $\ldots$, (u_k, u_k') are mutually orthogonal hyperbolic pairs.*

Proof. First we construct the subspace U'. If $U \cap W^\perp = \{0\}$, then $V = U^\perp \oplus W$ and we may take $U' := W$. Therefore we may suppose that $U \cap W^\perp \neq \{0\}$. In particular,

$$U \cap W^\perp \nsubseteq (W^\perp)^\perp = W$$

and it follows that there are vectors $u \in U \cap W^\perp$ and $w \in W^\perp$ such that $\beta(u, w) \neq 0$. By Lemma 7.3 the hyperbolic line $\langle\, u, w \,\rangle$ is spanned by isotropic (singular) vectors u and v.

The subspace $W' := W \oplus \langle\, v \,\rangle$ is totally isotropic (totally singular) and $U^\perp \cap W' = \{0\}$. By induction there is a totally isotropic (totally singular) subspace U' containing W' such that $V = U^\perp \oplus U'$.

Now suppose that $u_1, u_2, \ldots, u_k$ is a basis for U and let U_1 be the subspace $\langle\, u_2, \ldots, u_k \,\rangle$. Then $U^\perp$ is a hyperplane of $U_1^\perp$ and therefore $\dim(U_1^\perp \cap U') = 1$. Since $U_1^\perp \cap U'$ is not contained in $\langle\, u_1 \,\rangle^\perp$, there is a unique vector $u_1' \in U_1^\perp \cap U'$ such that $\beta(u_1, u_1') = 1$. The subspaces U_1 and $\langle\, u_1 \,\rangle^\perp \cap U'$ are contained in the non-degenerate space $\langle u_1, u_1' \rangle^\perp$ and satisfy the same conditions as U and U'. It follows by induction that u_1' may be extended to a basis of U' of the desired form. □

The Group $\Gamma L^*(V)$

In this section we shall describe the correlations and collineations of V in a way which is often more suitable for computations than the abstract definitions given earlier. The detailed verification of the assertions is straightforward and will be left as an exercise.

We begin with the observation that a finite-dimensional vector space V may be identified with its second dual V^{**} as follows. For $v \in V$, define $\hat{v} \in V^{**}$ by $\hat{v}(\varphi) := \varphi(v)$ for all $\varphi \in V^*$. Then the map $v \mapsto \hat{v}$ is an isomorphism from V onto V^{**} (Exercise 7.2 (i)).

If V_1 and V_2 are vector spaces and $f : V_1 \to V_2$ is a σ-semilinear transformation, then the map $f^* : V_2^* \to V_1^*$ defined by

$$f^*(\varphi) := \sigma^{-1}\varphi f$$

is a σ^{-1}-semilinear transformation and we have $f^{**}(\hat{v}) = f(v)\hat{}$ for all $v \in V_1$. From now on we shall identify each vector space with its second dual and each map f with f^{**}. Note that if $g : V_2 \to V_3$, then $(gf)^* = f^* g^*$.

Returning to the vector space V, we see that for the map $f : V \to V^*$ given by

$$f(v) = \beta(-, v),$$

where β is a σ-sesquilinear form on V, we have

$$f^*(v) = \sigma^{-1}\beta(v, -).$$

When β is non-degenerate, the correlations of $\mathcal{P}(V)$ induced by f and f^* are mutually inverse.

Now let $\Gamma L^*(V)$ be the set of all semilinear bijections from V to V together with all semilinear bijections from V to V^*. For $f \in \Gamma L^*(V)$, let $\overline{f}$ be the inverse of f^*.

For $f, g \in \Gamma L^*(V)$ define $f \circ g$ to be fg when $g : V \to V$ and $\overline{f}g$ when $g : V \to V^*$. In particular, if $f : V \to V^*$, the inverse of f in $\Gamma L^*(V)$ is f^*. It is straightforward, but tedious, to show that this definition of product turns $\Gamma L^*(V)$ into a group which contains $\Gamma L(V)$ as a subgroup of index 2 (Exercise 7.2 (vi)).

If $\dim V \geq 3$, then the Fundamental Theorem of Projective Geometry shows that the group $P\Gamma L^*(V)$ of all collineations and correlations of V is isomorphic to the quotient group $\Gamma L^*(V)/Z(V)$.

If p is the σ-semilinear transformation $p(v) := \beta(-, v)$ and if f is a τ-semilinear transformation in $\Gamma L(V)$, it is of particular interest (for later calculations) to compute the matrices of various transformations associated with p and f.

Begin by choosing a basis $e_1, e_2, \ldots, e_n$ for V and let $\omega_1, \omega_2, \ldots, \omega_n$ be the dual basis for V^*. We find that $p(e_j) = \sum_{i=1}^n b_{ij}\omega_i$, where $b_{ij} := \beta(e_i, e_j)$. The determinant of the matrix $J := (b_{ij})$ is called the *discriminant* of β.

If $f(e_j) = \sum_{i=1}^n a_{ij}e_i$, then $f^*(\omega_j) = \sum_{i=1}^n \tau^{-1}a_{ji}\omega_i$ and so the matrix of f^* is $\tau^{-1}(A^t)$, where A^t is the *transpose* of the matrix $A := (a_{ij})$ of f. The maps p and f have associated field automorphisms σ and τ and of course these must be taken into account when composing the corresponding matrices. In particular, the matrix of p^* is $\sigma^{-1}(J^t)$ and the matrix of the σ-semilinear transformation $\overline{p}$ is J^{-t}, the inverse of J^t. Similarly, the matrix of $\overline{f}$ is A^{-t} and the matrix of $f^\perp := p \circ f \circ p^{-1}$ $(= \overline{p}\overline{f}p^*)$ is $J^{-t}\sigma(A^{-t})\sigma\tau\sigma^{-1}(J^t)$.

Thus for $f \in GL(V)$ we have

$$f^\perp = f \qquad \text{if and only if} \qquad A^tJ\sigma(A) = J. \tag{7.6}$$

Note that $f^\perp$ is the unique $\sigma\tau\sigma^{-1}$-semilinear transformation such that

$$\beta(f^\perp(u), f(v)) = \sigma\tau\sigma^{-1}\beta(u, v) \tag{7.7}$$

for all $u, v \in V$.

Flags and Frames

From now on suppose that β is a non-degenerate reflexive σ-sesquilinear form corresponding to a polarity π.

The image of the flag F, given by

$$V_1 \subset V_2 \subset \cdots \subset V_k, \tag{7.8}$$

under the action of the polarity $\pi : X \mapsto X^\perp$ is the flag

$$V_k^\perp \subset V_{k-1}^\perp \subset \cdots \subset V_1^\perp. \tag{7.9}$$

Thus π fixes F if and only if $V_i^\perp = V_{k-i+1}$ for $i = 1, 2, \ldots, k$. If this is the case, then for $i \leq \frac{1}{2}(k+1)$ we have $V_i \subseteq V_i^\perp$ and consequently V_i is totally isotropic. Conversely, if the flag F of (7.8) consists of totally isotropic subspaces, then the flag

$$V_1 \subset V_2 \subset \cdots \subset V_k \subseteq V_k^\perp \subset \cdots \subset V_1^\perp \tag{7.10}$$

is fixed by π. This shows that the flags fixed by π may be identified with the flags of totally isotropic subspaces. We let $\Delta_\pi(V)$ denote the set of proper flags of totally isotropic subspaces of V. The maximal flags of totally isotropic subspaces are called the *chambers* of $\Delta_\pi(V)$. (Not all geometries have flags fixed by a correlation. For example, ordinary Euclidean space $\mathbb{R}^n$ with the usual inner product has no isotropic vectors and therefore no non-empty fixed flags.)

Suppose that W is a subspace of V. It follows from (7.7) that

$$f^\perp(W^\perp) = f(W)^\perp \tag{7.11}$$

and therefore, if f fixes the flag F given by (7.8), then $f^\perp$ fixes $\pi(F)$.

The map $\hat{\pi} : f \mapsto f^\perp$ is an automorphism of $\mathit{\Gamma L}(V)$ and our calculations show that if the flag F is fixed by π, then the stabilizer of F in $\mathit{\Gamma L}(V)$ is fixed by $\hat{\pi}$.

Now consider a frame $\mathcal{F} = \{P_1, P_2, \dots, P_n\}$ of $\mathcal{P}(V)$, where $P_i = \langle e_i \rangle$, and let $\omega_1, \omega_2, \dots, \omega_n$ be the dual basis to $e_1, e_2, \dots, e_n$. For $i = 1, 2 \dots, n$, let $e_i^* := \overline{p}(\omega_i)$. Then for $1 \le i, j \le n$ we have

$$\beta(e_i^*, e_j) = \delta_{ij}. \tag{7.12}$$

We define $\mathcal{F}^*$ to be the frame $\{P_1^*, P_2^*, \dots, P_n^*\}$, where $P_i^* := \langle e_i^* \rangle$ and we note that $P_i^* = \langle \mathcal{F} \setminus \{P_i\} \rangle^\perp$. Thus each point P_i^* of $\mathcal{F}^*$ is *not* orthogonal to exactly one point of $\mathcal{F}$, namely P_i. It follows that $P_i^{**} = P_i$ for all i. From this description it is easy to see that if $\Phi \subseteq \mathcal{F}$ and $W = \langle \Phi \rangle$, then $W^\perp = \langle \Phi' \rangle$, where $\Phi' := \{ P_i^* \mid P_i \notin \Phi \}$. In particular, if F is a flag in the apartment $\Sigma(\mathcal{F})$, then $\pi(F)$ belongs to $\Sigma(\mathcal{F}^*)$.

Suppose that f fixes $\mathcal{F}$. That is, $f(P_i) = P_{\phi(i)}$, where ϕ is a permutation of $\{1, 2, \dots, n\}$. Then (7.7) shows that

$$f^\perp(P_i^*) = P_{\phi(i)}^* \tag{7.13}$$

and therefore $f^\perp$ fixes $\mathcal{F}^*$.

Polar Frames

Let m be the Witt index of β and let

$$U := \langle e_1, e_2, \dots, e_m \rangle$$

be a maximal totally isotropic subspace of V. By Lemma 7.5 there is a totally isotropic subspace
$$U' := \langle f_1, f_2, \dots, f_m \rangle$$
such that $(e_1, f_1), \dots, (e_m, f_m)$ are mutually orthogonal hyperbolic pairs. It follows that $V = (U \oplus U') \perp U''$, where U'' contains no isotropic vectors. Let $u_1, u_2, \dots, u_h$ be a basis for U'' such that for $i \neq j$, $\beta(u_i, u_j) = 0$. If for $1 \leq i \leq m$, $P_i := \langle e_i \rangle$, then $P_i^* = \langle f_i \rangle$. For $1 \leq i \leq h$, let $Q_i := \langle u_i \rangle$. Then $Q_i^* = Q_i$ and the set

$$\mathcal{F} := \{ P_i, P_i^* \mid 1 \leq i \leq m \} \cup \{ Q_i \mid 1 \leq i \leq h \} \tag{7.14}$$

is a frame such that $\mathcal{F} = \mathcal{F}^*$.

The points Q_i are not isotropic and therefore they never occur in flags of totally isotropic subspaces. Consequently we restrict our attention to the set $\{ P_i, P_i^* \mid 1 \leq i \leq m \}$ and call this a *polar frame* of V (with respect to π).

The Building of a Polarity

Suppose that $m \neq 0$. If $\mathcal{F}$ is a polar frame, the *apartment* of $\mathcal{F}$ in $\Delta_\pi(V)$ is the set of flags
$$\Sigma_\pi(\mathcal{F}) := \Sigma(\mathcal{F}) \cap \Delta_\pi(V).$$
The assumption that $m \neq 0$ ensures that $\Sigma_\pi(\mathcal{F})$ is not empty. The *building* of π is the set $\Delta_\pi(V)$ of all proper flags of totally isotropic subspaces together with the set $\mathcal{A}_\pi$ of all apartments obtained from polar frames.

At this point we may prove the polar analogue of Theorem 5.3.

7.15 Theorem. *If M_1 and M_2 are chambers of $\Delta_\pi(V)$, there is an apartment Σ_π which contains both M_1 and M_2.*

Proof. Suppose that M_1 and M_2 are the flags $V_1 \subset V_2 \subset \cdots \subset V_m$ and $W_1 \subset W_2 \subset \cdots \subset W_m$, respectively. By Theorem 5.2 there is a frame $\mathcal{F}$ such that $\Sigma(\mathcal{F})$ contains both
$$V_1 \subset V_2 \subset \cdots \subset V_m \subseteq V_m^\perp \subset \cdots \subset V_1^\perp$$
and
$$W_1 \subset W_2 \subset \cdots \subset W_m \subseteq W_m^\perp \subset \cdots \subset W_1^\perp.$$
These flags are fixed by π and therefore the results of the previous section show that they belong to $\Sigma(\mathcal{F}^*)$.

Let Φ be the subset of $\mathcal{F}$ which is a frame for $\mathcal{P}(V_m)$ and let Ψ be the subset of $\mathcal{F}^*$ which is a frame for $\mathcal{P}(W_m)$. Now set
$$\Gamma := \Phi \cup \{ P \in \Psi \mid P \notin \mathcal{P}(V_m) \}.$$

By Lemma 5.3, $V_m \cap W_i$ is spanned by a subset of Φ and by a subset of Ψ. It follows that W_i is spanned by a subset of Γ and that V_i is spanned by a subset of $\Phi \subseteq \Gamma$.

We shall show that Γ can be extended to a polar frame. Suppose that $P \in \mathcal{F}$ is such that $P^* \in \Psi$. If $P^* \notin \mathcal{P}(V_m)$, then P^* cannot be orthogonal to every element of Φ otherwise $\langle\, \Phi, P^* \,\rangle$ would be totally isotropic, contrary to the maximality of V_m. Therefore $P \in \mathcal{P}(V_m)$. This means that we may choose the notation so that

$$\Phi = \{P_1, P_2, \dots, P_m\}$$

and so that the remaining points of Γ are P_1^*, P_2^*, $\dots$, P_r^* for some $r \le m$. Applying Lemma 7.5 with $U = \langle\, \Phi \,\rangle$ and $W = \langle\, P_1^*, \dots, P_r^* \,\rangle$ we see that Γ extends to a polar frame whose apartment contains M_1 and M_2. □

We shall use the following lemma in the proof of the polar analogue of Theorem 5.4.

7.16 Lemma. *If $g \in SL(V)$ and if $\mathcal{F}$ is a polar frame such that $g^\perp(P) = g(P)$ for all $P \in \mathcal{F}$, then there is a linear isometry f such that $f(P) = g(P)$ for all $P \in \mathcal{F}$.*

Proof. Put $\mathcal{F}' := \{\, g(P) \mid P \in \mathcal{F} \,\}$ and recall that for $P \in \mathcal{F}$, P^* is the unique element of $\mathcal{F}$ not orthogonal to P. It follows from (7.7) that $g(P^*)$ is the unique element of $\mathcal{F}'$ not orthogonal to $g^\perp(P) = g(P)$. Thus $\mathcal{F}'$ is a polar frame. It follows from Theorem 7.4 that there is a linear isometry f such that $f(P) = g(P)$ for all $P \in \mathcal{F}$. □

7.17 Theorem. *Suppose that $\mathcal{F}$ and $\mathcal{F}'$ are polar frames and that F and G are flags which belong to both apartments $\Sigma_\pi(\mathcal{F})$ and $\Sigma_\pi(\mathcal{F}')$. Then there is a linear isometry f such that $f(F) = F$, $f(G) = G$ and $f(\mathcal{F}) = \mathcal{F}'$.*

Proof. Suppose at first that F is a chamber of $\Delta_\pi(V)$. The sets $\hat{F} := F \cup \pi(F)$ and $\hat{G} := G \cup \pi(G)$ are flags of V fixed by π and it follows from Theorem 5.4 that for some $g \in SL(V)$ we have

$$g(\hat{F}) = \hat{F}, \quad g(\hat{G}) = \hat{G} \quad \text{and} \quad g(\mathcal{F}) = \mathcal{F}'.$$

From (7.11) we have $g^\perp(\hat{F}) = \hat{F}$ and $g^\perp(\hat{G}) = \hat{G}$. It also follows from (7.11) that $g^\perp(\mathcal{F}) = \mathcal{F}'$. But now $g^{-1}g^\perp$ fixes $\mathcal{F}$ and the flag $\hat{F}$. Since F is a chamber of $\Sigma_\pi(\mathcal{F})$, it follows that $g^{-1}g^\perp$ fixes every point of $\mathcal{F}$ and hence by Lemma 7.16 there is a linear isometry f which coincides with g on $\mathcal{F}$. In particular, $f(F) = F$, $f(G) = G$ and $f(\mathcal{F}) = \mathcal{F}'$.

Now consider the general case in which F is not necessarily a chamber. Let F_1 be a chamber of $\Sigma_\pi(\mathcal{F})$ which contains F and let G_1 be a chamber of

$\Sigma_\pi(\mathcal{F}')$ which contains G. By Theorem 7.15 there is a polar frame $\mathcal{F}''$ such that F_1 and G_1 belong to $\Sigma_\pi(\mathcal{F}'')$. By the previous paragraph we can find isometries f_1 and f_2 such that

$$f_1(\mathcal{F}) = \mathcal{F}'', \quad f_2(\mathcal{F}'') = \mathcal{F}',$$
$$f_1(F_1) = F_1, \quad f_1(G) = G, \quad f_2(F) = F \quad \text{and} \quad f_2(G_1) = G_1.$$

Thus $f := f_2 f_1$ satisfies our requirements. □

In Chapter 5 we showed that the stabilizer of a frame and the stabilizer of a chamber in the apartment of that frame form a BN-pair. A similar result holds in the polar case. We postpone the proof of the general result to Chapter 9 but consider the special case of a symplectic geometry in Chapter 8.

Throughout the exercises assume that all the vector spaces which occur are finite-dimensional.

EXERCISES

7.1 Let X and Y be subspaces of the vector space V and show that

$$(X+Y)^\circ = X^\circ \cap Y^\circ \qquad \text{and} \qquad (X \cap Y)^\circ = X^\circ + Y^\circ.$$

7.2 (*i*) Show that the map $V \to V^{**} : v \mapsto \hat{v}$, where $\hat{v}(\varphi) := \varphi(v)$, is an isomorphism.

(*ii*) If $f : V_1 \to V_2$ is a σ-semilinear map and $f^* : V_2^* \to V_1^*$ is the σ^{-1}-semilinear map defined by $f^*(\varphi) := \sigma^{-1}\varphi f$, show that

$$(f^*(f(X)^\circ))^\circ = X + \ker(f)$$

for all subspaces X of V.

(*iii*) If $f : V_1 \to V_2$ is a σ-semilinear map and if $\hat{v}$ is defined as in (*i*), show that $f^{**}(\hat{v}) = f(v)\hat{}$.

Henceforth we identify each vector space with its second dual.

(*iv*) If β is a σ-sesquilinear form on V and if $f : V \to V^*$ is the map $f(v) := \beta(-, v)$ induced by β, show that $f^*(v) = \sigma^{-1}\beta(v, -)$.

(*v*) If β is non-degenerate, show that the correlations induced by f and f^* are mutually inverse.

(*vi*) Verify that $\mathit{\Gamma L}^*(V)$ is a group.

7.3 For $i = 1, 2$, let V_i be a vector space with basis $e_1^{(i)}, e_2^{(i)}, \ldots, e_{n_i}^{(i)}$ and dual basis $\omega_1^{(i)}, \omega_2^{(i)}, \ldots, \omega_{n_i}^{(i)}$. If $f : V_1 \to V_2$ is a σ-semilinear map such that

$$f(e_j^{(1)}) = \sum_{i=1}^{n} a_{ij} e_i^{(2)},$$

show that

$$f^*(\omega_j^{(2)}) = \sum_{i=1}^{n} \sigma^{-1} a_{ji} \omega_i^{(1)}.$$

7.4 Prove Lemma 7.2.

7.5 Let $e_1, e_2, \ldots, e_n$ be a basis for V and let $\omega_1, \omega_2, \ldots, \omega_n$ be the dual basis for V^*.

(i) If the σ-semilinear map $p : V \to V^*$ in $\mathit{\Gamma L}^*(V)$ has matrix J and the τ-semilinear map $f : V \to V$ has matrix A with respect to these bases, show that

$$f^{-1} \circ p \circ f \quad \text{has matrix} \quad \tau^{-1}(A^t J \sigma(A)).$$

(ii) If p corresponds to the σ-sesquilinear form β, show that

$$f^{-1} \circ p \circ f \quad \text{corresponds to} \quad \tau^{-1}\beta(f-, f-),$$

which is a $\tau^{-1}\sigma\tau$-sesquilinear form.

7.6 Suppose that V is a vector space of dimension at least 3 over a field $\mathbb{F}$ and let $\Delta(V)$ be the building of V.

(i) Show that the group of order preserving bijections of $\Delta(V)$ is isomorphic to $P\mathit{\Gamma L}^*(V)$.

(ii) Using Sylow's theorem and Exercise 5.8, show that, when $\mathbb{F}$ is finite, every automorphism of $PSL(V)$ induces an automorphism of the building $\Delta(V)$. Deduce that $\mathrm{Aut}(PSL(V)) = P\mathit{\Gamma L}^*(V)$.

7.7 Let β be a σ-sesquilinear form on V. Show that

$$\overline{\beta}(u + \operatorname{rad} V, v + \operatorname{rad} V) := \beta(u, v)$$

is a well-defined non-degenerate σ-sesquilinear form on $V/\operatorname{rad} V$.

7.8 Let V and W be vector spaces over a field $\mathbb{F}$. A *σ-sesquilinear pairing* is a function

$$\beta : V \times W \to \mathbb{F}$$

such that $\beta(-, w) \in V^*$ and $\sigma^{-1}\beta(v, -) \in W^*$ for all $v \in V$ and $w \in W$, where σ is an automorphism of $\mathbb{F}$. For each subspace H of V, let

$$H^\perp := \{\, w \in W \mid \beta(h, w) = 0 \text{ for all } h \in H \,\}$$

and for each subspace K of W, let

$${}^\perp K := \{\, v \in V \mid \beta(v, k) = 0 \text{ for all } k \in K \,\}.$$

Show that

(i) the maps $W/H^\perp \to (H/{}^\perp W \cap H)^* : w + H^\perp \mapsto \beta(-, w)$ and ${}^\perp K/{}^\perp W \to (W/K + V^\perp)^* : u + {}^\perp W \mapsto \sigma^{-1}\beta(u, -)$ are isomorphisms,

(ii) ${}^\perp(H^\perp) = H + {}^\perp W$,

(iii) $\dim W + \dim(H \cap {}^\perp W) = \dim H + \dim H^\perp$,

(iv) $\dim {}^\perp K + \dim(K + V^\perp) = \dim W + \dim {}^\perp W$, and

(v) $\dim H^\perp + \dim(H + {}^\perp W) = \dim V + \dim V^\perp$.

7.9 Suppose that β is a symmetric bilinear form on a vector space over a field of characteristic 2. Show that either β is alternating or else the set of isotropic vectors together with 0 form a proper subspace. If the field is finite (or, more generally, *perfect*) show that the subspace of isotropic vectors is a hyperplane.

7.10 Define a gallery in $\Delta_\pi(V)$ as in Exercise 5.5 and then show that any two chambers of $\Delta_\pi(V)$ are connected by a gallery.

7.11 Show that each panel in an apartment Σ of $\Delta_\pi(V)$ is contained in exactly two chambers of Σ.

7.12 Let Σ be an apartment of $\Delta_\pi(V)$ and let M be a chamber of Σ. By Theorem 7.15, for each flag $F \in \Delta_\pi(V)$ there is an apartment Σ' containing F and M. By Theorem 7.17 there is a linear isometry f such that $f(\Sigma') = \Sigma$ and $f(M) = M$. Define $\rho : \Delta_\pi(V) \to \Sigma$ by $\rho(F) := f(F)$. Show that

(i) ρ is well-defined.

(ii) If F and G are adjacent chambers, then either $\rho(F) = \rho(G)$ or $\rho(F)$ and $\rho(G)$ are adjacent chambers of Σ.

(iii) If $\rho(F) \subseteq M$, then $\rho(F) = F \subseteq M$.

8

Symplectic Groups

With this chapter we begin the investigation of the groups associated with polar geometries. The symplectic groups have been chosen first because in many ways they resemble the groups $SL(V)$ and consequently it is possible to organize the material along the same lines as Chapter 4.

Thus, after defining the groups, we compute the orders of the finite examples and then prove the simplicity of the projective symplectic groups. Our approach uses Iwasawa's criterion combined with a study of symplectic transvections. Next we study embeddings of the symmetric groups in the symplectic groups and finish the chapter with the construction of symplectic BN-pairs.

Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and suppose that π is a symplectic polarity induced by the (non-degenerate) alternating form β. Thus if X is a subspace of V, then $\pi(X) = p(X)^\circ$, where $p\colon V \to V^*$ is the linear transformation $p(v) := \beta(-, v)$.

The elements of $GL(V)$ which commute with p form the *symplectic group* $Sp(V)$. It follows from the calculations of the last section of Chapter 7 (equation (7.7)) that $f \in GL(V)$ commutes with p if and only if

$$\beta(f(u), f(v)) = \beta(u, v) \quad \text{for all } u, v \in V.$$

More generally, a σ-semilinear transformation $f \in \mathit{\Gamma L}(V)$ induces a collineation which commutes with π if and only if, for some $a \in \mathbb{F}$, we have

$$\beta(f(u), f(v)) = a\sigma\beta(u, v) \quad \text{for all } u, v \in V.$$

(cf. Exercise 7.5 (*ii*).)

The transformations which satisfy this condition form the group $\mathit{\Gamma Sp}(V)$ which contains $Sp(V)$ as a normal subgroup.

Matrices

Let $e_1, e_2, \ldots, e_n$ be a basis for V and let $\omega_1, \omega_2, \ldots, \omega_n$ be the dual basis for V^*. Then

$$J := (\beta(e_i, e_j))$$

is the matrix of p with respect to these bases and if A is the matrix of $f \in GL(V)$, it follows from (7.6) that f is in $Sp(V)$ if and only if

$$A^t JA = J.$$

Since $\beta(v, v) = 0$ for all $v \in V$ we find, by expanding $\beta(u + v, u + v) = 0$, that $\beta(u, v) = -\beta(v, u)$. Thus $J^t = -J$ and $p \circ p = -\mathbf{1}$ (in the notation of Chapter 7).

We have just shown that $Sp(V)$ is isomorphic to the group $Sp(n, \mathbb{F})$ of $n \times n$ matrices A such that $A^t JA = J$. If $\mathbb{F}$ is the finite field of q elements, we write $Sp(n, q)$ instead of $Sp(n, \mathbb{F})$.

It is easy to see that the only scalar matrices in $Sp(n, \mathbb{F})$ are I and $-I$ and therefore the group $PSp(V)$ of collineations of $\mathcal{P}(V)$ induced by $Sp(V)$ is isomorphic to $Sp(V)/\{\pm\mathbf{1}\}$.

Symplectic Bases

It is possible to choose a basis of V better adapted to calculations. Choose $e_1 \neq 0$ and f_1 so that $\beta(e_1, f_1) \neq 0$. Replacing f_1 by $\beta(e_1, f_1)^{-1} f_1$ we may suppose that (e_1, f_1) is a *hyperbolic pair* (as defined in Chapter 7). Then

$$V = \langle e_1, f_1 \rangle \perp \langle e_1, f_1 \rangle^\perp.$$

Continuing in this fashion, choosing a hyperbolic pair (e_2, f_2) in $\langle e_1, f_1 \rangle^\perp$ and so on, we find that V can be written as an orthogonal direct sum of the hyperbolic lines $\langle e_i, f_i \rangle$. That is,

$$V = \langle e_1, f_1 \rangle \perp \langle e_2, f_2 \rangle \perp \ldots \perp \langle e_m, f_m \rangle$$

where $\beta(e_i, e_j) = \beta(f_i, f_j) = 0$ and $\beta(e_i, f_j) = \delta_{ij}$. It follows that $\dim V = 2m$ is even and for each m there is only one symplectic geometry of dimension $2m$ over $\mathbb{F}$ (up to isomorphism).

The basis e_1, f_1, e_2, f_2, $\ldots$, e_m, f_m is called a *symplectic basis* of V and the corresponding frame

$$\{\langle e_1 \rangle, \langle f_1 \rangle, \ldots, \langle e_m \rangle, \langle f_m \rangle\}$$

is called a *symplectic frame*. When the basis is ordered as e_1, e_2, $\ldots$, e_m, f_1, f_2, $\ldots$, f_m the matrix J of β has the particularly simple form

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

The subspace $M := \langle e_1, e_2, \ldots, e_m \rangle$ is totally isotropic and, as $\dim M = \frac{1}{2} \dim V$, it follows that M is a *maximal* totally isotropic subspace. Thus the Witt index of β is m.

Order Formulae

The group $Sp(V)$ acts regularly on the set of ordered symplectic bases of V and hence in order to compute $|Sp(V)|$ when $\mathbb{F}$ is finite we simply count the number of symplectic bases.

If $\mathbb{F} = \mathbb{F}_q$, the number of hyperbolic pairs in V is

$$(q^{2m}-1)(q^{2m}-q^{2m-1})/(q-1) = (q^{2m}-1)q^{2m-1}.$$

It follows that the number of ordered symplectic bases is

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

and therefore

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

Putting $m=1$ we see that $|Sp(2,q)| = |SL(2,q)|$. This is also a consequence of the following more general result.

8.1 Theorem. *The groups $Sp(2,\mathbb{F})$ and $SL(2,\mathbb{F})$ are isomorphic.*

Proof. Check that a 2×2 matrix A satisfies

$$A^t\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

if and only if $\det A = 1$. □

The Action of $PSp(V)$ on $\mathcal{P}(V)$

Recall from Chapter 1 that the rank of a transitive permutation group is the number of orbits of the stabilizer of a point.

8.2 Theorem. *If $\dim V \geq 4$, then $PSp(V)$ acts as a permutation group of rank 3 on the points of $\mathcal{P}(V)$.*

Proof. By Witt's theorem (Theorem 7.4), $PSp(V)$ is transitive on the points of $\mathcal{P}(V)$ and if P is a point, the orbits of $PSp(V)_P$ are

$$\{P\},\ \{\,Q \mid Q\in P^{\perp},\ Q\neq P\,\} \text{ and } \{\,Q \mid Q\notin P^{\perp}\,\}. \qquad □$$

8.3 Theorem. *The action of $PSp(V)$ on the points of $\mathcal{P}(V)$ is primitive.*

Proof. If $\dim V = 2$, then by Theorem 8.1, $PSp(V) = PSL(V)$ and it is a consequence of Theorem 4.1 that this group is primitive. Therefore we may assume that $\dim V \geq 4$.

Suppose that B is a block of imprimitivity such that $|B| > 1$ and choose $P \in B$. By the proof of Theorem 8.2, if $B \cap P^\perp$ contains a point other than P, then $P^\perp \subseteq B$. In this case, if $R \notin P^\perp$, choose $Q \in (P+R)^\perp$. Then $Q \in B$ and $R \in Q^\perp$, hence $R \in B$. Thus B consists of all the points of $\mathcal{P}(V)$ in this case.

Now suppose that B contains a point not in $P^\perp$. Then B contains all the points not in $P^\perp$. Suppose that $R \in P^\perp$, $R \neq P$ and choose $Q \notin P^\perp \cup R^\perp$. Then $Q \in B$ and since $R \notin Q^\perp$, it follows that $R \in B$. Again B consists of all the points of $\mathcal{P}(V)$ and this proves that $PSp(V)$ is primitive. □

Symplectic Transvections

Recall from Chapter 4 that if $\varphi \in V^*$, $u \in V$ and $\varphi(u) = 0$, the transvection $t_{\varphi,u}$ is defined by

$$t_{\varphi,u}(v) := v + \varphi(v)u.$$

We have $t_{\varphi,u} \in Sp(V)$ if and only if $\beta(t_{\varphi,u}(v), t_{\varphi,u}(w)) = \beta(v,w)$ for all $v, w \in V$. This is the case if and only if $\varphi(w)\beta(v,u) + \varphi(v)\beta(u,w) = 0$. Choose v so that $\beta(u,v) = 1$. Then $\varphi(w) = \varphi(v)\beta(u,w)$ and it follows that $\ker\varphi = \langle u \rangle^\perp$. Thus the transvections that belong to $Sp(V)$ have the form

$$t(v) = v + a\beta(v,u)u \tag{8.4}$$

for some $a \in \mathbb{F}$ and $u \in V$. We call these maps *symplectic transvections*.

The *root group* $X_{P,P^\perp}$ consists of those symplectic transvections of the form (8.4) for which $P = \langle u \rangle$. If $f \in Sp(V)$ and W is a subspace of V, then $f(W)^\perp = f(W^\perp)$ and consequently

$$fX_{P,P^\perp}f^{-1} = X_{f(P),f(P)^\perp}.$$

Thus $X_{P,P^\perp}$ is a normal subgroup of the stabilizer $Sp(V)_P$. Notice that $X_{P,P^\perp} = X_P \cap Sp(V)$, where X_P is the group which occurs in Theorem 4.3.

In order to use Iwasawa's criterion to show that $PSp(V)$ is simple we need to prove that $Sp(V)$ is generated by the conjugates of $X_{P,P^\perp}$ and that it is equal to its derived group $Sp(V)'$

8.5 Theorem. *The symplectic transvections generate $Sp(V)$.*

Proof. Let T be the subgroup of $Sp(V)$ generated by the transvections. We first show that T is transitive on $V\setminus\{0\}$. If $u_1,u_2 \in V$ and $\beta(u_1,u_2)\neq 0$, define

$$t(v) := v - \beta(u_1,u_2)^{-1}\beta(u_1-u_2,v)(u_1-u_2).$$

Then $t\in T$ and $t(u_1)=u_2$. If $u_1\neq u_2$ and $\beta(u_1,u_2)=0$, choose w so that $\beta(u_1,w)\neq 0$ and $\beta(u_2,w)\neq 0$. By the previous case we can find $t_1,t_2\in T$ such that $t_1(u_1)=w$ and $t_2(w)=u_2$. Thus $t_2t_1(u_1)=u_2$ and this completes the proof that T is transitive on $V\setminus\{0\}$.

Next we show that T is transitive on hyperbolic pairs. Suppose that (u_1,v_1) and (u_2,v_2) are hyperbolic pairs. Since T is transitive on vectors we may suppose that $u_1=u_2=u$. If $\beta(v_1,v_2)\neq 0$, the transvection

$$t(v) := v - \beta(v_1,v_2)^{-1}\beta(v_1-v_2,v)(v_1-v_2)$$

fixes u and takes v_1 to v_2. If $\beta(v_1,v_2)=0$, then $(u,u+v_1)$ is a hyperbolic pair such that $\beta(v_1,u+v_1)\neq 0$ and $\beta(u+v_1,v_2)\neq 0$. We have just shown that this implies that there are elements of T which take (u,v_1) to $(u,u+v_1)$ and $(u,u+v_1)$ to (u,v_2). Hence T is transitive on hyperbolic pairs.

Now suppose that $f\in Sp(V)$ and that (u,v) is a hyperbolic pair. Then for some $t\in T$ we have $tf(u)=u$ and $tf(v)=v$. Hence tf acts as the identity on $L := \langle\, u,v\,\rangle$ and fixes $L^\perp$. By induction, the restriction of tf to $L^\perp$ is the product of transvections t_1, t_2, …, t_k of $L^\perp$. Since $V=L+L^\perp$, each t_i extends to a transvection $t_i' := \mathbf{1}_L + t_i$ of V. It follows that $f=t^{-1}t_1'\dots t_k'\in T$. ◻

8.6 Corollary. $Sp(V)\subseteq SL(V)$.

Proof. The determinant of every transvection is 1 and $Sp(V)$ is generated by its transvections. ◻

The Simplicity of $PSp(V)$

8.7 Theorem. *$Sp(2m,\mathbb{F})' = Sp(2m,\mathbb{F})$ except for $Sp(2,\mathbb{F}_2)$, $Sp(2,\mathbb{F}_3)$ and $Sp(4,\mathbb{F}_2)$.*

Proof. We argue by induction on the dimension. So suppose that

$$Sp(2m,\mathbb{F})' = Sp(2m,\mathbb{F}),\ \ t\in X_{P,P^\perp}\subseteq Sp(2m+2,\mathbb{F})$$

and let L be a hyperbolic line in $P^\perp$. Then the restriction s of t to $L^\perp$ belongs to $Sp(L^\perp)$ and by assumption $Sp(L^\perp)'=Sp(L^\perp)$. Since t acts as the identity on L, it follows that $t = \mathbf{1}_L + s$ and therefore it belongs to $Sp(2m+2, \mathbb{F})'$. Thus $Sp(2m+2, \mathbb{F})'$ contains all the symplectic transvections and by Theorem 8.5 we have $Sp(2m+2, \mathbb{F})' = Sp(2m+2, \mathbb{F})$.

From Theorem 8.1 we have $Sp(2, \mathbb{F}) = SL(2, \mathbb{F})$ and from Theorem 4.4 it follows that $Sp(2, \mathbb{F})' = Sp(2, \mathbb{F})$ unless $\mathbb{F}$ is $\mathbb{F}_2$ or $\mathbb{F}_3$. Thus to complete the proof we need only check that $Sp(4,3)' = Sp(4,3)$ and $Sp(6,2)' = Sp(6,2)$.

Let $e_1, f_1, \ldots, e_m, f_m$ be a symplectic basis ($m = 2$ or 3) and use the ordering $e_1, e_2, \ldots, e_m, f_1, f_2, \ldots, f_m$ so that a matrix M is in $Sp(2m, \mathbb{F})$ if and only if $M^t J M = J$, where

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

In particular, the matrices

$$\begin{pmatrix} A^{-1} & 0 \\ 0 & A^t \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} I & B \\ 0 & I \end{pmatrix}$$

belong to $Sp(2m, \mathbb{F})$ for all $A \in GL(m, \mathbb{F})$ and for all $m \times m$ matrices B such that $B = B^t$.

By definition, $Sp(2m, \mathbb{F})'$ contains

$$\begin{pmatrix} A^{-1} & 0 \\ 0 & A^t \end{pmatrix}^{-1} \begin{pmatrix} I & B \\ 0 & I \end{pmatrix}^{-1} \begin{pmatrix} A^{-1} & 0 \\ 0 & A^t \end{pmatrix} \begin{pmatrix} I & B \\ 0 & I \end{pmatrix}$$

and a straightforward calculation shows that this commutator is equal to

$$\begin{pmatrix} I & B - ABA^t \\ 0 & I \end{pmatrix}.$$

In the case of $Sp(4,3)$, set $A := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ and $B := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ to get $B - ABA^t = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Thus $Sp(4,3)'$ contains a transvection. It follows that $Sp(4,3)'$ contains all symplectic transvections and hence it equals $Sp(4,3)$.

For $Sp(6,2)$, set

$$A := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

to get

$$B - ABA^t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus $Sp(6,2)'$ contains one, and hence all, of the symplectic transvections and it follows that $Sp(6,2)' = Sp(6,2)$ as required. □

8.8 Theorem. *The groups $PSp(2m,\mathbb{F})$ are simple except for $PSp(2,\mathbb{F}_2)$, $PSp(2,\mathbb{F}_3)$ and $PSp(4,\mathbb{F}_2)$.*

Proof. This follows from Theorem 1.2 (Iwasawa's criterion) together with Theorems 8.3, 8.5 and 8.7 □

The Symmetric Groups

We know from Theorem 8.1 and the description of $PSL(2,q)$ in Chapter 4 that $PSp(2,2) \simeq S_3$ and that $PSp(2,3) \simeq A_4$. It remains for us to show that $PSp(4,2)$ is truly an exception to Theorem 8.8. This is a corollary to the next theorem.

8.9 Theorem. *If $m \geq 2$, then the symmetric group S_{2m+2} is a subgroup of $Sp(2m,2)$.*

Proof. Let Ω be a set of size $2m+2$ and let V be the set of partitions $\{\Gamma,\Delta\}$, such that $\Omega = \Gamma \cup \Delta$, $\Gamma \cap \Delta = \emptyset$ and $|\Gamma|$ is even. Make V into a vector space over $\mathbb{F}_2$ by defining

$$\{\Gamma_1,\Delta_1\} + \{\Gamma_2,\Delta_2\} := \{\Gamma_1+\Gamma_2, \Gamma_1+\Delta_2\},$$

where $\Gamma_1 + \Gamma_2 := (\Gamma_1 \cup \Gamma_2) \setminus (\Gamma_1 \cap \Gamma_2)$ denotes the *symmetric difference* of Γ_1 and Γ_2. This is a well-defined vector space (Exercise 8.4) and S_{2m+2} acts on it faithfully as a group of linear transformations. Next define

$$\beta(\{\Gamma_1,\Delta_1\},\{\Gamma_2,\Delta_2\}) := |\Gamma_1 \cap \Gamma_2| \pmod 2$$

and note that β is a well-defined non-degenerate alternating form on V preserved by S_{2m+2}. Since $\dim V = 2m$, we have

$$S_{2m+2} \subseteq Sp(2m,\mathbb{F}_2).$$ □

8.10 Corollary. $Sp(4,2) \simeq S_6$.

Proof. We have $S_6 \subseteq Sp(4,2)$ by Theorem 8.9 and both these groups have order 720. □

An alternative approach to this isomorphism is to continue the analysis of A_7 and A_8 acting on the projective geometry of Points, Lines and Planes begun in Chapter 6. A transposition of S_8 interchanges Points and Planes and therefore acts as a polarity of the geometry. Moreover, each Point is incident with its polar Plane and therefore we have a *symplectic* polarity. The elements of A_8 that commute with a transposition of S_8 form a subgroup isomorphic to S_6. It follows that $S_6 \simeq Sp(4,2)$.

Symplectic BN-pairs

There are several ways to construct a BN-pair for $Sp(V)$. One way is to use the building $\Delta_\pi(V)$ introduced in Chapter 7, where π is a symplectic polarity. Then N is defined to be the stabilizer of a polar frame and B is defined to be the stabilizer of a maximal flag (chamber) which is in the apartment of that frame. Theorems 7.15 and 7.17 can be used to prove the BN-pair axioms in the same way that Theorems 5.2 and 5.4 were used in the construction of the BN-pair for $SL(V)$. The details of this approach can be found in the next chapter.

The method we employ here involves a careful study of the action of a symplectic polarity on a BN-pair for $SL(V)$. The key result that we use is Theorem 5.10 which gives a normal form for elements of $SL(V)$.

Let π be a symplectic polarity of V and let J be the matrix of the associated alternating form β with respect to the basis $e_1, e_2, \ldots, e_m, f_m, f_{m-1}, \ldots, f_1$, where $e_1, f_1, \ldots, e_m, f_m$ is a symplectic basis for β. Then

$$J = \begin{pmatrix} 0 & Q \\ -Q & 0 \end{pmatrix}, \quad \text{where} \quad Q := \begin{pmatrix} 0 & & 1 \\ & \iddots & \\ & 1 & \\ 1 & & 0 \end{pmatrix}.$$

The calculations on p. 61 at the end of Chapter 7 show that the symplectic group $Sp(V)$ consists of the elements of $SL(V)$ fixed by the automorphism $\hat{\pi} : f \mapsto f^\perp$ induced by β. Note that if the matrix of f is A, then the matrix of $f^\perp$ is $-JA^{-t}J$ (since $J^t = J^{-1} = -J$ in this case).

Let B be the subgroup of $SL(V)$ corresponding to the upper triangular matrices and let N be the subgroup corresponding to the monomial matrices. That is, N is the stabilizer of the symplectic frame obtained from the basis chosen above and B is the stabilizer of the flag

$$V_1 \subset V_2 \subset \ldots \subset V_m \subset V_{m-1}^\perp \subset \ldots \subset V_1^\perp \tag{8.11}$$

where, for $1 \le i \le m$, $V_i = \langle e_1, \ldots, e_i \rangle$. Note that the subspaces V_i are totally isotropic and that B and N form a BN-pair for $SL(V)$. The calculations in the section 'Flags and Frames' at the end of the last chapter show that B and N are fixed by $\hat{\pi}$. It follows that $\hat{\pi}$ induces an automorphism of the Weyl group $W := N/B \cap N$. The Weyl group is the symmetric group S_{2m}, acting on the frame $\{ \langle e_i \rangle, \langle f_i \rangle \mid 1 \le i \le m \}$ and it follows from (7.13) that the automorphism induced by $\hat{\pi}$ coincides with conjugation by the permutation that interchanges $\langle e_i \rangle$ and $\langle f_i \rangle$ for all i.

For $1 \le i < m$, let n_i be the element of N that sends e_i to $-e_{i+1}$, e_{i+1} to e_i and fixes all other basis elements. Then n_i corresponds to the transposition $(\langle e_i \rangle, \langle e_{i+1} \rangle)$ of S_{2m}. In addition, let $n_m \in N$ correspond to the transposition $(\langle e_m \rangle, \langle f_m \rangle)$ and suppose that $n_m(e_m) = -f_m$ and $n_m(f_m) = e_m$.

For $i \neq m$, it follows from (7.7) that $\hat{\pi}(n_i)$ sends f_i to $-f_{i+1}$ and f_{i+1} to f_i. Thus $\hat{\pi}(n_i)$ commutes with n_i and therefore $\hat{n}_i := n_i\hat{\pi}(n_i)$ is fixed by $\hat{\pi}$. It is also the case that n_m is fixed by $\hat{\pi}$. Put $\hat{n}_m := n_m$ and for $1 \leq i \leq m$, let $\hat{w}_i$ be the image of $\hat{n}_i$ in W.

The subgroup W_π of elements fixed by $\hat{\pi}$ consists of the permutations that fix the partition

$$\{\, \{\langle e_i \rangle, \langle f_i \rangle\} \mid 1 \leq i \leq m \,\}$$

of the symplectic frame. The 2^m permutations that fix all the blocks of this partition form a normal subgroup of W_π whose quotient group is the symmetric group S_m (acting faithfully on the m blocks). In fact W_π is the semidirect product of the normal subgroup of order 2^m and S_m. This is often expressed by saying that W_π is the *wreath product* $\mathbb{Z}_2 \wr S_m$. From this description it is clear that W_π is generated by $\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_m$.

Now let B_π (resp. N_π) be the elements of B (resp. N) fixed by $\hat{\pi}$.

8.12 Theorem. *The groups B_π and N_π form a BN-pair for $Sp(V)$ with Weyl group W_π.*

Proof. We need to check that the BN-pair axioms given in Chapter 5 hold for B_π and N_π. If $g \in Sp(V)$, then $g \in SL(V)$ and therefore, by Theorem 5.10, $g = bn_wu$, where $b \in B$, $w \in W$ and $u \in U_w^-$. Since $\hat{\pi}$ fixes B it follows from Theorem 5.5 that $\hat{\pi}$ fixes w and therefore we may choose $n_w \in N_\pi$. But now $\hat{\pi}$ fixes U_w^- (see Exercise 5.16) and as

$$g = \hat{\pi}(b)\hat{\pi}(n_w)\hat{\pi}(u)$$

it follows from the uniqueness part of Theorem 5.10 that b and u are in B_π. In particular, $Sp(V) = B_\pi N_\pi B_\pi$ and axiom (i) holds.

We have shown that W_π is generated by $\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_m$ and consequently $W_\pi = N_\pi / B_\pi \cap N_\pi$. This leaves axiom (iv) to be proved. Certainly $\hat{n}_i B_\pi \hat{n}_i \neq B_\pi$ for all i and so we must show that axiom $(iv)(b)$ holds.

From the corresponding result for $SL(V)$ we have

$$\hat{n}_i B_\pi n \subseteq (B\hat{n}_i nB) \cup (B\hat{\pi}(n_i)nB) \cup (Bn_i nB) \cup (BnB)$$

where $1 \leq i < m$ and $n \in N_\pi$. In the first part of this proof we showed that every element of $Sp(V)$ has a unique expression of the form bn_wu, where $b \in B_\pi$, $w \in W_\pi$ and where $u \in U_w^-$ is fixed by $\hat{\pi}$. It follows that

$$\hat{n}_i B_\pi n \subseteq (B_\pi \hat{n}_i n B_\pi) \cup (B_\pi n B_\pi).$$

Similarly, $n_m B_\pi n \subseteq (B_\pi n_m n B_\pi) \cup (B_\pi n B_\pi)$, and this completes the proof. □

Symplectic Buildings

The *symplectic building* of V is the building $\Delta_\pi(V)$ (see pp. 61 and 63), where π is a symplectic polarity. The elements of $\Delta_\pi(V)$ are the proper flags of totally isotropic subspaces. It follows from (7.11) that B_π is the stabilizer in $Sp(V)$ of the flag

$$V_1 \subset V_2 \subset \ldots \subset V_m$$

defined above in (8.11) and N_π is the stabilizer of the symplectic frame

$$\{\, \langle\, e_i \,\rangle, \langle\, f_i \,\rangle \mid 1 \leq i \leq m \,\}.$$

It follows from Witt's Theorem (Theorem 7.4) that two elements of $\Delta_\pi(V)$ are in the same orbit of $Sp(V)$ if and only if they have the same *type* (as defined in Chapter 5 for $\Delta(V)$).

Generalized Quadrangles

Consider the special case of a symplectic geometry V of dimension 4. Then the building $\Delta_\pi(V)$ has just three types of elements. In projective terminology these are *points* P, *lines* ℓ and *flags* (P, ℓ), where P is a point of the line ℓ. We have already noted that $Sp(V)$ is transitive on the incident point-line pairs (P, ℓ). In addition we have:

Q1) *Any two distinct points are incident with at most one line.*

Q2) *Any two distinct lines are incident with at most one point.*

Q3) *Given a point P and a line ℓ not incident with P, there is a unique flag (P', ℓ') such that P' is on ℓ and P is on ℓ'. (Take $P' := \ell \cap P^\perp$ and $\ell' := P + P'$.)*

A collection of points and lines satisfying these conditions is called a *generalized quadrangle* (cf. Exercise 6.7). See Payne and Thas (1984) for a survey of the finite case.

EXERCISES

8.1 Let V be a symplectic geometry of dimension $2m$ over the finite field $\mathbb{F}_q$.

(i) Show that the number of hyperbolic pairs in V is $(q^{2m}-1)q^{2m-1}$.

(ii) Show that the number of totally isotropic subspaces of dimension k is

$$\prod_{i=0}^{k-1}(q^{2m-2i}-1)/(q^{i+1}-1).$$

(iii) If P is any point of $\mathcal{P}(V)$, show that the stabilizer of P in $PSp(V)$ has orbits of lengths

$$1, \quad (q^{2m-2}-1)/(q-1) \quad \text{and} \quad q^{2m-1}$$

on the points of $\mathcal{P}(V)$.

(iv) Let E be a maximal totally isotropic subspace of V and show that the number of maximal totally isotropic subspaces F such that $E \cap F = \{0\}$ is $q^{\frac{1}{2}m(m+1)}$.

(v) Let E be a maximal totally isotropic subspace of V and show that the number of maximal totally isotropic subspaces F of V such that $\dim(E \cap F) = k$ is

$$\prod_{i=1}^{m-k}(q^{m+1}-q^i)/(q^i-1).$$

8.2 Let E be a vector space of dimension m over $\mathbb{F}$.

(i) Set $V := E^* \oplus E$, where E^* is the dual space of E, and define a bilinear form β on V by

$$\beta((\varphi, v), (\psi, w)) := \varphi(w) - \psi(v).$$

Show that β is a non-degenerate alternating form and that if e_1, e_2, $\ldots$, e_m is a basis of E with dual basis ω_1, ω_2, $\ldots$, ω_m, then

$$(\omega_1, 0), (0, e_1), (\omega_2, 0), (0, e_2), \ \ldots, \ (\omega_m, 0), (0, e_m)$$

is a symplectic basis for V.

(*ii*) Let S^2E be the set of all (possibly degenerate) symmetric bilinear forms on E. For $\gamma_1, \gamma_2 \in S^2E$ and $a_1, a_2 \in \mathbb{F}$, define

$$(a_1\gamma_1 + a_2\gamma_2)(u, v) := a_1\gamma_1(u, v) + a_2\gamma_2(u, v)$$

and show that S^2E is a vector space of dimension $\frac{1}{2}m(m+1)$ over $\mathbb{F}$.

(*iii*) Identify E^* with the subspace $\{\, (\varphi, 0) \mid \varphi \in E^* \,\}$ of V. For $\gamma \in S^2E$ show that the linear transformation $\tilde{\gamma}$ defined by

$$\tilde{\gamma}(\varphi, v) := (\varphi + \gamma(v, -), v)$$

is an element of the subgroup $Sp(V)(E^*)$ of $Sp(V)$ which fixes every vector of E^*. Conversely, show that every element of $Sp(V)(E^*)$ has this form and deduce that $Sp(V)(E^*)$ is isomorphic to S^2E. (cf. Exercise 2.16.)

(*iv*) Show that $Sp(V)(E^*)$ acts regularly on the set of maximal totally isotropic subspaces F of V such that $E^* \cap F = \{0\}$. (Hint. Use Lemma 7.5.)

(*v*) For $f \in GL(E)$ and $\gamma \in S^2E$, define $f\gamma$ by

$$(f\gamma)(u, v) := \gamma(f^{-1}(u), f^{-1}(v)).$$

Show that in this way $GL(E)$ may be regarded as a group of linear transformations of S^2E.

(*vi*) For $f \in GL(E)$, let $\bar{f}$ be the element of $GL(E^*)$ that takes $\varphi \in E^*$ to φf^{-1}. Show that every element of the subgroup $Sp(V)_{E^*,E}$ fixing both E^* and E can be written in the form

$$(\varphi, v) \mapsto (\bar{f}(\varphi), f(v)) \qquad \text{for some } f \in GL(E).$$

Deduce that $Sp(V)_{E^*}$ is isomorphic to the semidirect product $(S^2E).GL(E)$.

(*vii*) If V' is any symplectic geometry of dimension $2m$ over $\mathbb{F}$ and if E' and F' are maximal totally isotropic subspaces of V' such that $E' \cap F' = \{0\}$, show that there is a linear isometry $g\colon V' \to V$ such that $g(E') = E$ and $g(F') = F$.

(*viii*) If F_1 and F_2 are maximal totally isotropic subspaces of V such that $\dim(E^* \cap F_1) = \dim(E^* \cap F_2)$, show that there is an element of $Sp(V)_{E^*}$ that takes F_1 to F_2.

8.3 Let V be a symplectic geometry defined by a non-degenerate alternating form β and let π be the corresponding polarity.

(i) Suppose that $\beta(u,v)=0$ and set $\varphi := \beta(-,v)$ and $\psi := \beta(-,u)$. Show that $t_{\varphi,u}^{\perp} = t_{\psi,v}$ and that $t_{\varphi,u}$ and $t_{\psi,v}$ commute. (See Exercise 4.10.)

(ii) Let X be a root group of $SL(V)$ and let X_π be the fixed elements of $\hat{\pi}$ in $\langle X, \hat{\pi}(X) \rangle$. Show that either

(a) $X_\pi = X = \hat{\pi}(X)$ consists of symplectic transvections, or

(b) there are linearly independent vectors u and v such that $\beta(u,v)=0$ and X_π consists of the transformations

$$x \mapsto x + a\beta(x,u)v + a\beta(x,v)u, \qquad a \in \mathbb{F}.$$

In case (a) X_π is called a *long root group* and in case (b) it is called a *short root group* of $Sp(V)$.

(iii) If $P = \langle u \rangle$, let X_P be the long root group of symplectic transvections $x \mapsto x + a\beta(x,u)u$, where $a \in \mathbb{F}$. show that $\langle X_P, X_Q \rangle$ is isomorphic to $\mathbb{F}^+$, $\mathbb{F}^+ \oplus \mathbb{F}^+$ or $SL(2,\mathbb{F})$ according to whether $P = Q$, $P+Q$ is a totally isotropic line or $P+Q$ is a hyperbolic line, respectively. If $P+Q$ is a hyperbolic line, show that $\langle X_P, X_Q \rangle$ acts transitively on the points of $P+Q$.

8.4 Show that the vector space introduced in the proof of Theorem 8.9 is well-defined and that β is a non-degenerate alternating form.

8.5 Let V be symplectic geometry over the field $\mathbb{F}$ defined by an alternating form β. Suppose that $\beta(u,v)=1$ and set $W := \langle u,v \rangle^\perp$. Make the set $E := \mathbb{F} \times W$ into a group by defining the product of (a_1,w_1) and (a_2,w_2) to be $(a_1+a_2+\beta(w_1,w_2), w_1+w_2)$. Define an action of E on V by setting

$$\begin{aligned}(a,w)u &:= u,\\ (a,w)v &:= au + v + w,\\ (a,w)w' &:= \beta(w,w')u + w' \qquad \text{for } w' \in W\end{aligned}$$

and extending (a,w) to V by linearity. Show that

(i) E may be regarded as a normal subgroup of $Sp(V)_u$ which acts regularly on the set $\{ v' \in V \mid \beta(u,v') = 1 \}$.

(ii) $Sp(V)_u$ is the semidirect product $E.Sp(W)$.

(iii) $(a_1,w_1)^{-1}(a_2,w_2)^{-1}(a_1,w_1)(a_2,w_2) = (2\beta(w_1,w_2),0)$.

(*iv*) $\{(a,0) \mid a \in \mathbb{F}\}$ acts on V as the long root group $X_{\langle u \rangle}$ of symplectic transvections $x \mapsto x - a\beta(x,u)u$.

(*v*) For $b \in \mathbb{F} \setminus \{0\}$ let h_b be the element of $Sp(V)$ that takes u to bu, v to $b^{-1}v$ and fixes every element of W. Show that

$$h_b(a,w)h_b^{-1} = (b^2a, bw)$$

and that h_b commutes with every element of $Sp(W)$. Moreover, if $H := \{h_b \mid b \in \mathbb{F} \setminus \{0\}\}$, then $Sp(V)_{\langle u \rangle} = E.Sp(W).H$ and this group is the normalizer of $X_{\langle u \rangle}$ in $Sp(V)$.

8.6 By following steps analogous to those of Exercise 6.6, show that for $n \geq 2$, $PSp(2n,q)$ is never isomorphic to an alternating group.

8.7 Suppose that q is a power of a prime and that p is a prime that does not divide q. Let f be the least positive integer such that p divides $q^{2f} - 1$.

(*i*) Deduce from Exercise 4.11 that if p divides $q^{2i}-1$, then f divides i and that p^a divides $(q^{2fi}-1)/(q^{2f}-1)$ if and only if p^a divides i.

(*ii*) Given m, let $\ell = \lfloor m/f \rfloor$ and let D be the direct product of ℓ copies of $Sp(2f,q)$. As in Exercise 4.11 show that S_ℓ acts as a group of automorphisms of D and that a Sylow p-subgroup of the semidirect product of D by S_ℓ is isomorphic to a Sylow p-subgroup of $Sp(2m,q)$. Show that if p is odd, a Sylow p-subgroup of $Sp(2f,q)$ is cyclic.

(*iii*) Show that the order of a Sylow p-subgroup P of $PSp(2m,q)$ is strictly less than $(\sqrt{3}\,(q+1))^m$, when p is odd, or $(4(q+1))^m$ when $p=2$. Moreover, for $m \geq 2$, $|P|$ is always less than q^{m^2}, the order of a Sylow subgroup corresponding to the characteristic of $\mathbb{F}_q$.

8.8 Let $\mathcal{F}$ be a symplectic frame of $\mathcal{P}(V)$, where V is a symplectic geometry of dimension $2m$. Show that the apartment of $\mathcal{F}$ contains $2^m m!$ chambers and that the Weyl group acts regularly on this set of chambers.

8.9 Let A be a panel of the apartment Σ of the symplectic building $\Delta_\pi(V)$. Show that A is in exactly two chambers of Σ.

8.10 Let V be the symplectic geometry of dimension $2m$ over $\mathbb{F}$, where the characteristic of $\mathbb{F}$ is not 2 and write

$$V = L_1 \perp L_2 \perp \ldots \perp L_m,$$

where $L_1, L_2, \ldots, L_m$ are hyperbolic lines. For $k = 1, 2, \ldots, m$ let t_k be the element of $Sp(V)$ that acts as multiplication by -1 on $W := L_1 \perp \ldots \perp L_k$ and as the identity on $W^\perp$. Show that every element of order 2 in $Sp(V)$ is conjugate to one of $t_1, t_2, \ldots, t_m$ and that the centralizer of t_k in $Sp(2m, \mathbb{F})$ is $Sp(2k, \mathbb{F}) \times Sp(2m - 2k, \mathbb{F})$.

8.11 Let V be a symplectic geometry of dimension $2m$ over the field $\mathbb{F}$ and let U be a totally isotropic subspace of dimension k.

(i) If $f \in Sp(V)$, show that $\ker(\mathbf{1} - f)^\perp = \operatorname{im}(\mathbf{1} - f)$.

(ii) Let $Sp(V)(U^\perp) = \{ f \in Sp(V) \mid f(v) = v \text{ for all } v \in U^\perp \}$. If $f \in Sp(V)(U^\perp)$, show that

$$\gamma_f(v_1 + U^\perp, v_2 + U^\perp) := \beta(v_1, v_2) - \beta(v_1, f(v_2))$$

is a well-defined symmetric bilinear form on $V/U^\perp$ and the map $f \mapsto \gamma_f$ is an isomorphism from $Sp(V)(U^\perp)$ onto the additive group $S^2(V/U^\perp)$ of symmetric bilinear forms on $V/U^\perp$. (See Exercise 8.2.)

(iii) Let H be the subgroup of $Sp(V)$ consisting of those elements that fix U and $U^\perp/U$ pointwise. For $h \in H$, define $\bar{h} : U^\perp/U \to U$ by $\bar{h}(v + U) := v - h(v)$. Show that $h \mapsto \bar{h}$ is a well-defined homomorphism of H onto $\operatorname{Hom}(U^\perp/U, U)$ whose kernel is $Sp(V)(U^\perp)$.

(iv) Show that the group H of (iii) acts regularly on the set of totally isotropic complements to $U^\perp$.

(v) Show that $Sp(V)_U$ is isomorphic to the semidirect product

$$H.(GL(U) \times Sp(U^\perp/U))$$

(cf. Exercise 8.5).

(vi) If $\mathbb{F} = \mathbb{F}_q$, show that the number of totally isotropic complements to $U^\perp$ is

$$q^{k(2m-2k)+\binom{k+1}{2}}.$$

8.12 Show that, up to isomorphism, there is only one generalized quadrangle with 15 points and 15 lines in which each line has 3 points. Hence, or otherwise, identify the generalized quadrangle described in Exercise 6.7 with the generalized quadrangle of $Sp(4, 2)$. Deduce that $S_6 \simeq Sp(4, 2)$. (cf. J. J. Sylvester (1844))

9

BN-pairs, Diagrams and Geometries

In Chapters 5 and 8 we described the buildings associated with the groups $SL(V)$ and $Sp(V)$ and in each case constructed a BN-pair. Here we construct BN-pairs for a large class of groups associated with polar geometries. Later in the chapter we indicate connections between our examples and various abstract notions of geometry.

The first part of the chapter is mainly concerned with those properties that have a uniform proof based solely on the BN-pair axioms. In particular, we give a purely algebraic account of the elementary theory of the Weyl group of a BN-pair. This leads us to consider the somewhat more general notion of a Coxeter group and its associated simplicial complex—its Coxeter complex.

After this we give the axioms for a building in a form that explicitly states that the apartments are Coxeter complexes. This is sufficient to cover all our examples. For the general theory of buildings, consult Tits (1976), Brown (1989) or Ronan (1989). For the material of this chapter see Bourbaki (1968). A more detailed study of unitary and orthogonal groups is carried out in the following chapters.

The BN-pair of a Polar Building

In Chapter 8 the BN-pair for $Sp(V)$ was obtained as the fixed points of the automorphism of $SL(V)$ induced by a symplectic polarity acting on the BN-pair for $SL(V)$. In this section the verification of the BN-pair axioms is based on properties of the flags and frames of the polar building. In particular, we use Theorems 7.15 and 7.17 of Chapter 7.

Suppose that V is a finite-dimensional vector space over a field $\mathbb{F}$ and that β is a (non-degenerate) alternating, symmetric or σ-hermitian form on V. Let π be the polarity of $\mathcal{P}(V)$ corresponding to β. The *building* of the polar geometry $(\mathcal{P}(V), \pi)$ consists of the set $\Delta := \Delta_\pi(V)$ of proper flags of totally isotropic subspaces of V together with the apartments obtained from the polar frames of V. In order that the building be non-empty we shall assume that $\dim V \geq 2$ and that V contains isotropic vectors. We shall also suppose that if β is symmetric it arises from a quadratic form. These assumptions exclude certain orthogonal and unitary geometries from our considerations.

In addition to these buildings we also consider certain buildings constructed as follows. Let β be the polar form of a non-degenerate quadratic form $Q : V \to \mathbb{F}$ and suppose that the characteristic of $\mathbb{F}$ is 2. In this case we assume that V contains singular vectors. The building Δ consists of the proper flags of *totally singular* subspaces. The apartments are the sets of flags obtained from the polar frames of *singular* points. We leave it as an exercise to show that appropriate analogues of Theorems 7.15 and 7.17 hold for Δ.

In all cases, let m be the Witt index of the geometry. Let

$$\mathcal{F} := \{ P_i, P_i^* \mid 1 \le i \le m \}$$

be a polar frame and recall from Chapter 7 that $\Phi \subseteq \mathcal{F}$ spans a totally isotropic (resp. totally singular) subspace if and only if $P \in \Phi$ implies $P^* \notin \Phi$ (for all P). In the case of a quadratic form we *define* P^* to be the unique element of $\mathcal{F}$ not orthogonal to P. Let Σ be the apartment of $\mathcal{F}$.

A group of isometries of the polar geometry is said to be *strongly transitive* if it is transitive on the pairs $(M', \mathcal{F}')$, where $\mathcal{F}'$ is a polar frame and M' is a chamber of the apartment $\Sigma(\mathcal{F}')$.

9.1 Theorem. *The group of all linear isometries of a polar geometry is strongly transitive.*

Proof. Suppose that M' is a chamber of the apartment Σ' of the polar frame $\{ Q_i, Q_i^* \mid 1 \le i \le m \}$. We may choose the notation so that

$$M' := \{ \langle Q_1, \dots, Q_i \rangle \mid 1 \le i \le m \}.$$

Similarly, we may suppose that

$$M := \{ \langle P_1, \dots, P_i \rangle \mid 1 \le i \le m \}.$$

It is easily seen that there is a linear isometry f from $\langle \mathcal{F}' \rangle$ to $\langle \mathcal{F} \rangle$ such that $f(Q_i) = P_i$ and $f(Q_i^*) = P_i^*$. By Witt's theorem, f extends to a linear isometry of V such that $f(\mathcal{F}') = \mathcal{F}$ and $f(M') = M$. □

Let G be a strongly transitive group of isometries, let N be the stabilizer of $\mathcal{F}$ and let B be the stabilizer of the chamber

$$M := \{ \langle P_1, \dots, P_i \rangle \mid 1 \le i \le m \}.$$

We shall show that, except for certain orthogonal groups, the subgroups B and N form a BN-pair for G. And even in the case of the exceptions it is only Axiom $(iv)(a)$ that fails. Recall that the BN-pair axioms are:

(*i*) $G = \langle B, N\rangle$

(*ii*) $H := B \cap N$ *is a normal subgroup of* N.

(*iii*) $W := N/H$ *is generated by elements* $\{\, w_i \mid i \in I \,\}$ *such that* $w_i^2 = 1$ *for all* $i \in I$.

(*iv*) *If* $w_i = n_i H$ *and* $n \in N$, *then*

  (*a*) $n_i B n_i \neq B$ *and*

  (*b*) $n_i B n \subseteq (B n_i n B) \cup (BnB)$.

Axiom (*ii*) is consequence of the following lemma.

9.2 Lemma. *The kernel of the action of* N *on* $\mathcal{F}$ *is* $B \cap N$ *and the group* $N/B \cap N$ *acts regularly on the chambers of* Σ.

Proof. If $n \in N$ fixes every point of $\mathcal{F}$, then n fixes M and hence $n \in B \cap N$. Conversely, suppose that $n \in B \cap N$. Then n fixes $P_1, P_2, \ldots, P_m$ and by (7.13), it fixes $P_1^*, P_2^*, \ldots, P_m^*$. In particular, $B \cap N$ is the kernel of the action of N on $\mathcal{F}$. By assumption G is strongly transitive, and therefore N is transitive on the chambers of Σ. The stabilizer in N of the chamber M is $B \cap N$ and therefore $N/B \cap N$ is regular. □

9.3 Theorem. $G = BNB$.

Proof. Given $g \in G$, consider the chamber $g(M)$. By Theorem 7.15 there is an apartment Σ' that contains both M and $g(M)$. By the strong transitivity of G we can find b such that $b(M) = M$ and $b(\Sigma') = \Sigma$. Then $bg(M) \in \Sigma$ and by Lemma 9.2 there exists $n \in N$ such that $nbg(M) = M$. We have b, $nbg \in B$ and hence $g \in BNB$, as required. □

This establishes the first of the BN-pair axioms. In order to proceed we need to find a suitable set of generators of order 2 for the group $W := N/H$, where $H := B \cap N$.

The pairs $\{P_i, P_i^*\}$, $1 \le i \le m$ are blocks of imprimitivity for the action of W on $\mathcal{F}$. Thus W is a subgroup of the wreath product $\mathbb{Z}_2 \wr S_m$ of order $2^m\, m!$ described in Chapter 8 in the context of symplectic BN-pairs. On the other hand, the number of chambers of $\Sigma(\mathcal{F})$ is $2m(2m-2)\cdots 4 \cdot 2 = 2^m\, m!$ and so by Lemma 9.2, $|W| = 2^m\, m!$. It follows that $W = \mathbb{Z}_2 \wr S_m$.

For $1 \le i < m$ we may choose $n_i \in N$ so that $w_i := n_i H$ is the permutation $(P_i, P_{i+1})(P_i^*, P_{i+1}^*)$. These permutations generate the symmetric group S_m. Choose $n_m \in N$ so that $w_m := n_m H$ is the transposition (P_m, P_m^*). Then $R := \{w_1, \ldots, w_m\}$ generates W. By construction, the elements of R have order 2 and so the BN-pair Axiom (*iii*) holds. It remains to prove Axiom (*iv*).

As in Chapter 5 we consider panels of chambers. By definition a *panel* is a proper flag of $m-1$ totally isotropic subspaces.

9.4 Lemma. *If A is a panel in the apartment Σ, then A is contained in exactly two chambers of Σ.*

Proof. The proof is similar to that of Lemma 5.6. We may suppose that A consists of all the subspaces

$$P_1,\ \langle\, P_1, P_2\,\rangle,\ \ldots,\ \langle\, P_1, \ldots, P_m\,\rangle \tag{9.5}$$

except $\langle\, P_1, \ldots, P_i\,\rangle$ for some i. If $i < m$, the only way to extend A to a chamber is to adjoin either

$$\langle\, P_1, \ldots, P_{i-1}, P_i\,\rangle \quad \text{or} \quad \langle\, P_1, \ldots, P_{i-1}, P_{i+1}\,\rangle.$$

If $i = m$, the only possibility is to adjoin

$$\langle\, P_1, \ldots, P_{m-1}, P_m\,\rangle \quad \text{or} \quad \langle\, P_1, \ldots, P_{m-1}, P_m^*\,\rangle. \qquad \square$$

9.6 Theorem. *Axiom $(iv)(b)$ holds for B and N.*

Proof. Suppose that $n \in N$ and that $n_1, \ldots, n_m$ are the generators of N defined above. We shall show that $nBn_i \subseteq Bnn_iB \cup BnB$. As in the corresponding proof for $SL(V)$ in Chapter 5, $A = M \cap n_i(M)$ is the panel obtained from M by omitting $\langle\, P_1, \ldots, P_i\,\rangle$. If $b \in B$, then b fixes A and therefore $n(A) \subseteq nbn_i(M)$. By Theorem 7.15 there is an apartment $\Sigma' := \Sigma(\mathcal{F}')$ that contains both M and $nbn_i(M)$. Then Σ and Σ' both contain M and $n(A)$. By Theorem 7.17 there is an isometry f that fixes M and $n(A)$ and takes $\mathcal{F}'$ to $\mathcal{F}$. Also, by the strong transitivity of G, there is an element $g \in G$ that fixes M and takes $\mathcal{F}'$ to $\mathcal{F}$. Then $g^{-1}f$ fixes M and $\mathcal{F}'$. By Lemma 9.2, applied to the group of all isometries, $g^{-1}f$ fixes $n(A)$. Thus $g \in B$ and $gnbn_i(M)$ is a chamber of Σ which contains $n(A)$. It follows from Lemma 9.4 that $gnbn_i(M)$ is $n(M)$ or $nn_i(M)$. Thus $nbn_i \in BnB \cup Bnn_iB$. On taking inverses we have

$$n_iBn \subseteq (BnB) \cup (Bn_inB). \qquad \square$$

9.7 Lemma. *Except for orthogonal geometries of Witt index m and dimension $2m$, every panel is in at least three chambers of Δ.*

Proof. We may suppose the panel A to be given by (9.5) of Lemma 9.4. If $i < m$, it is clear that A is in at least three chambers. So suppose that

$i = m$. It is now a question of determining the number of totally isotropic (resp. totally singular) subspaces containing $E := \langle P_1, \dots, P_{m-1} \rangle$. Passing to $E^\perp/E$, we may suppose that $m = 1$ and then count the isotropic (resp. singular) points.

Suppose that $P_1 = \langle e_1 \rangle$ and $P_1^* = \langle f_1 \rangle$, where $\beta(e_1, f_1) = 1$. If β is alternating, then $\langle e_1 \rangle$, $\langle f_1 \rangle$ and $\langle e_1 + f_1 \rangle$ are three distinct isotropic points. If β is σ-hermitian, choose $a \in \mathbb{F}$ such that $a \neq \sigma(a)$ and set $b = a - \sigma(a)$. Then $\langle e_1 \rangle$, $\langle f_1 \rangle$ and $\langle e_1 + bf_1 \rangle$ are three distinct isotropic points.

Finally, suppose that the geometry is defined by a quadratic form Q. By hypothesis there exists $u \in \langle e_1, f_1 \rangle^\perp$ such that $a = Q(u) \neq 0$. Then $\langle e_1 \rangle$, $\langle f_1 \rangle$ and $\langle -ae_1 + f_1 + u \rangle$ are distinct singular points. □

9.8 Theorem. *Except for orthogonal geometries of Witt index m and dimension $2m$, the groups B and N form a BN-pair for G.*

Proof. All that remains to be done is to show that Axiom $(iv)(a)$ holds. Retaining the notation of the previous lemmas we consider the panel $A = M \cap n_i(M)$. By Lemma 9.7 there is a chamber M' distinct from M and $n_i(M)$ which contains A. Let Σ' be an apartment containing M and M'. By the strong transitivity of G there exists b such that $b(M) = M$ and $b(\Sigma') = \Sigma$. Then b fixes A and so A is contained in $b(M')$. By Lemma 9.4, we must have $b(M') = n_i(M)$ and therefore $b^{-1}n_i(M) = M' \neq M$. It follows that $n_i b^{-1} n_i(M) \neq M$ and consequently $n_i B n_i \neq B$. □

The case of an orthogonal geometry of Witt index m and dimension $2m$ is dealt with in Chapter 11.

The Weyl Group

If the Weyl group of a BN-pair is finite, it turns out that N is uniquely determined by B. In the language of buildings, this means that there is only one way to define apartments (see Tits (1974), p. 54). We shall not prove that here but we shall show that the generators $\{ w_i \mid i \in I \}$ of the Weyl group W are uniquely determined by the other axioms. So throughout this section we suppose that G is a group with a BN-pair and that W is its Weyl group.

For $w \in W$, let $\ell(w)$ be the length of the shortest expression of w as a product of w_i's and call $\ell(w)$ the *length* of w. If $\ell(w) = k$ and $w = w_{i_1} w_{i_2} \cdots w_{i_k}$, then the word $w_{i_1} w_{i_2} \cdots w_{i_k}$ is called a *reduced expression* for w.

For each subset $J \subseteq I$, let $W_J := \langle w_i \mid i \in J \rangle$ and let N_J be the subgroup of N such that $N_J/H = W_J$, where $H := B \cap N$.

9.9 Lemma. *For all $J \subseteq I$, BN_JB is a subgroup of G.*

Proof. If $n \in N_J$, then $n = n_{i_1} \cdots n_{i_k}$, where $w_{i_j} := n_{i_j}H$ and $i_j \in J$. We shall show by induction on k that $nBN_JB \subseteq BN_JB$. This is certainly true when $n = 1$ and by induction we may suppose that

$$n_{i_2} \cdots n_{i_k} BN_JB \subseteq BN_JB. \tag{9.10}$$

By Axiom $(iv)(b)$ we have $n_{i_1}BN_JB \subseteq BN_JB$ and from (9.10) $nBN_JB \subseteq n_{i_1}BN_JB \subseteq BN_JB$. It is clear that $(BN_JB)^{-1} = BN_JB$ and therefore BN_JB is a group. □

A subgroup of G is said to be *parabolic* if it contains a conjugate of B. The subgroups $P_J := BN_JB$ are the *standard parabolic* subgroups of G. The lemma just proved shows that B and N_J form a BN-pair for P_J with Weyl group W_J. We shall show that the standard parabolic subgroups are the only subgroups of G that contain B and that the minimal elements of the set of subgroups that properly contain B are the $P_{\{i\}}$, $i \in I$. But first we need some results connecting the multiplication of double cosets with the multiplication in W. Note that throughout the following proofs we make constant use of the fact that for (B,B) double cosets X, Y and Z it is the case that $X \subseteq YZ$ if and only if $Y \subseteq XZ$.

We know from Theorem 9.3 that every (B,B) double coset has the form BwB for some $w \in W$. The next result shows that the element of W is unique.

9.11 Lemma. *If $w, w' \in W$ and $BwB = Bw'B$, then $w = w'$.*

Proof. We may suppose that $\ell(w) \leq \ell(w')$ and prove the result by induction on $\ell(w)$. If $\ell(w) = 0$, then $w = 1$ and it is clear that $w' = 1$ as well.

Suppose that $\ell(w) > 0$ and write $w = w_iw''$, where $\ell(w'') = \ell(w) - 1$. Then $Bw'B \subseteq (Bw_iB)(Bw''B)$ and consequently

$$Bw''B \subseteq (Bw_iB)(Bw'B) \subseteq (Bw_iw'B) \cup (Bw'B)$$

by Axiom $(iv)(b)$. If $Bw''B = Bw'B$, then by induction $w'' = w'$, contrary to $\ell(w'') < \ell(w')$. Hence $Bw''B = Bw_iw'B$. We have $\ell(w'') < \ell(w)$ and so $w'' = w_iw'$ by induction. Thus $w = w_iw'' = w'$. □

The next theorem is a strengthened form of Axiom $(iv)(b)$.

9.12 Theorem. *For all $i \in I$ and all $w \in W$ it is never the case that $\ell(w_iw) = \ell(w)$ or $\ell(ww_i) = \ell(w)$. In addition we have*

$$(i) \quad (Bw_iB)(BwB) = \begin{cases} Bw_iwB & \text{if } \ell(w_iw) = \ell(w) + 1 \\ (Bw_iwB) \cup (BwB) & \text{if } \ell(w_iw) = \ell(w) - 1 \end{cases}$$

$$(ii) \quad (BwB)(Bw_iB) = \begin{cases} Bww_iB & \text{if } \ell(ww_i) = \ell(w) + 1 \\ (Bww_iB) \cup (BwB) & \text{if } \ell(ww_i) = \ell(w) - 1 \end{cases}$$

Proof. The theorem is true if $w = 1$ and so we may suppose that $\ell(w) \neq 0$ and write $w = w'w_j$, where $\ell(w') = \ell(w) - 1$. Suppose that $\ell(w_iw) \geq \ell(w)$ but that $(Bw_iB)(BwB) \neq Bw_iwB$. Then $\ell(w') \leq \ell(w_iw) - 1 \leq \ell(w_iw')$ and by induction $(Bw_iB)(Bw'B) = Bw_iw'B$. We always have $Bw_iwB \subseteq (Bw_iB)(BwB)$ and in this case we are assuming that equality does not hold; therefore $(Bw_iB)(BwB) = (Bw_iwB) \cup (BwB)$. It follows that

$$\begin{aligned} BwB &\subseteq (Bw_iB)(BwB) \\ &\subseteq (Bw_iB)(Bw'B)(Bw_jB) \\ &= (Bw_iw'B)(Bw_jB) \end{aligned}$$

and hence $Bw_iw'B \subseteq (BwB)(Bw_jB)$. It follows from Axiom $(iv)(b)$ that $(BwB)(Bw_jB) \subseteq (Bw'B) \cup (BwB)$ and therefore $Bw_iw'B$ is $Bw'B$ or BwB. By Lemma 9.11 we cannot have $Bw_iw'B = Bw'B$ and therefore $Bw_iw'B = BwB$. But then $w_iw' = w$ and hence $w' = w_iw$, contrary to the assumption that $\ell(w_iw) \geq \ell(w)$.

We have shown that $\ell(w_iw) \geq \ell(w)$ implies $(Bw_iB)(BwB) = Bw_iwB$. Now suppose that $\ell(w_iw) \leq \ell(w)$. Then $(Bw_iB)(Bw_iwB) = BwB$. By Axiom $(iv)(b)$, $(Bw_iB)(BwB) \subseteq (Bw_iwB) \cup (BwB)$. If equality does not hold, then $(Bw_iB)(BwB) = Bw_iwB$ because in any case the product contains w_iw. From Axiom (iv) we have $(Bw_iB)^2 = B \cup (Bw_iB)$ and so $BwB = (Bw_iB)^2(BwB) = (BwB) \cup (Bw_iB)(BwB)$. But then $BwB = Bw_iwB$, a contradiction to Lemma 9.11. This shows that $\ell(w_iw) \leq \ell(w)$ implies $(Bw_iB)(BwB) = (Bw_iwB) \cup (BwB)$. In particular, it is never the case that $\ell(w_iw) = \ell(w)$. This completes the proof of (i).

Part (ii) follows from (i) by taking inverses. □

We are now able to obtain a useful description of the standard parabolic subgroups.

9.13 Theorem. *For $n \in N$, let $w := nH$ and suppose that $w_{i_1} \cdots w_{i_k}$ is a reduced expression for w. If $J := \{i_1, \ldots, i_k\}$, then*

$$P_J = \langle B, n \rangle = \langle B, nBn^{-1} \rangle.$$

Proof. It is clear that $\langle B, nBn^{-1} \rangle \subseteq \langle B, n \rangle \subseteq P_J$. For all j, choose $n_{i_j} \in N$ such that $w_{i_j} = n_{i_j}H$. We have $\ell(w_{i_1}w) < \ell(w)$ and so by Theorem 9.12 we have $(Bw_{i_1}B)(BwB) = (Bw_{i_1}wB) \cup (BwB)$. It follows that for some $b \in B$, $n_{i_1}bn \in BnB$ and hence $n_{i_1} \in \langle B, nBn^{-1} \rangle$. It follows by induction on $\ell(w)$ that $P_{J\setminus\{i_1\}} = \langle B, n_{i_1}^{-1}Bn^{-1}n_{i_1} \rangle \subseteq \langle B, nBn^{-1} \rangle$ and hence $P_J = \langle B, nBn^{-1} \rangle$. □

At last we have enough information to prove that the generators of W are uniquely determined by the subgroup B.

9.14 Theorem. *For all $w \in W \setminus \{1\}$, we have $w \in \{ w_i \mid i \in I \}$ if and only if $B \cup (BwB)$ is a group.*

Proof. Certainly $B \cup (Bw_iB)$ is a group for all i. Conversely, suppose that $B \cup (BwB)$ is a group and let $w = w_{i_1} \cdots w_{i_k}$ be a reduced expression for w. By Theorem 9.13 $B \cup (BwB) = P_J$, where $J = \{i_1, \ldots, i_k\}$. But then $Bw_{i_1}B \subseteq B \cup (BwB)$ whence $w = w_{i_1}$ by Lemma 9.11. □

9.15 Theorem. *If Q is a subgroup of G and $B \subseteq Q$, then $Q = P_J$ for some $J \subseteq I$.*

Proof. Let $J := \{ i \in I \mid Bw_iB \subseteq Q \}$. Then $P_J \subseteq Q$. If $h \in Q$, then $BhB = BwB$ for some $w \in W$ and $BwB \subseteq Q$. If $w_{i_1} \cdots w_{i_k}$ is a reduced expression for w, it follows from Theorem 9.13 that $Bw_{i_j}B \subseteq Q$ for all j. Hence $i_j \in J$ and therefore $h \in P_J$. This proves that $Q = P_J$. □

Further properties of the parabolic subgroups can be found in the exercises at the end of the chapter.

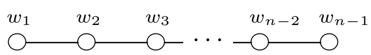
Coxeter Groups

It is customary to describe the Weyl group W of a BN-pair by a graph Γ called the *Coxeter-Dynkin diagram* of W. These diagrams arise in many contexts and are variously associated with the names Coxeter, Dynkin and Witt (see Hazewinkel et al. (1977)).

The vertices of Γ are the generators $\{ w_i \mid i \in I \}$ of W characterized in Theorem 9.14. For $i \neq j$ the vertex w_i is joined to w_j by $m_{ij} - 2$ edges, where m_{ij} is the order of w_iw_j. The matrix $M := (m_{ij})$ is symmetric with 1's on the diagonal. It is called the *Coxeter matrix* of Γ.

In Chapter 5 we found that the Weyl group of $SL(n, \mathbb{F})$ is the symmetric group S_n with generators the $n-1$ transpositions $w_i := (i, i+1)$, $(1 \leq i < n)$. Therefore its Coxeter-Dynkin diagram (said to be of type A_{n-1}) is simply

A_{n-1}

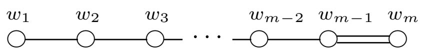


In the first section of the present chapter we showed that the Weyl group of the BN-pair associated with a polar geometry of Witt index m is the wreath product $\mathbb{Z}_2 \wr S_m$. This group acts on a set $\{1, 1', 2, 2', \ldots, m, m'\}$ of $2m$ elements and is generated by the permutations

$$\begin{aligned} w_i &:= (i, i+1)(i', (i+1)') \quad \text{for } 1 \le i < m, \quad \text{and} \\ w_m &:= (m, m'). \end{aligned}$$

Its Coxeter-Dynkin diagram (said to be of type C_m) is

C_m



Given a graph Γ with vertex set $R := \{x_1, x_2, \ldots, x_n\}$, let $m_{ij} - 2$ be the number of edges joining x_i to x_j when $i \neq j$ and let $m_{ii} = 1$ for all i. The *Coxeter group* $W := W(\Gamma)$ corresponding to this graph is the abstract group with generators R and relations $(x_i x_j)^{m_{ij}} = 1$. The pair (W, R) is a *Coxeter system* and the matrix $M := (m_{ij})$ is its *Coxeter matrix*.

It turns out that the Weyl group of a BN-pair is the Coxeter group of its Coxeter-Dynkin diagram. To prove this we shall first show that the Weyl groups satisfy a certain Exchange Condition and then show that every group which satisfies the Exchange Condition is a Coxeter group.

The Exchange Condition

In order to state this condition we need to define a length function for any group W generated by a subset R of elements of order 2. As in the case of a Weyl group, the *length* $\ell(w)$ of $w \in W$ is the number of elements in a shortest expression for w as a product of elements of R.

An expression $w := r_1 r_2 \cdots r_k$ with $r_1, r_2, \ldots, r_k \in R$ is said to be *reduced* if $k = \ell(w)$. (Strictly speaking an expression for w is a *sequence* $(r_1, r_2, \ldots, r_k)$ of elements of R such that $w = r_1 r_2 \cdots r_k$.)

We say that (W, R) satisfies the *Exchange Condition* if:

> *for each reduced expression $r_1 r_2 \cdots r_\ell$ for $w \in W$ and for each $r \in R$ such that $\ell(rw) \le \ell(w)$, there exists i such that*
>
> $$r r_1 \cdots r_{i-1} = r_1 \cdots r_i,$$
>
> *(equivalently, $rw = r_1 \cdots r_{i-1} r_{i+1} \cdots r_\ell$).*

It follows immediately that if (W, R) satisfies the Exchange Condition, then i is unique and $\ell(rw) = \ell(w) - 1$.

9.16 Lemma. *If W is the Weyl group of a BN-pair and if R is its distinguished set of generators, then (W, R) satisfies the Exchange Condition.*

Proof. Suppose that $r_1 r_2 \cdots r_\ell$ is a reduced expression for $w \in W$ and that for some $r \in R$, $\ell(rw) \leq \ell(w)$. Choose i minimal so that $rr_1 \cdots r_i$ is not reduced. Then $u := rr_1 \cdots r_{i-1}$ and $v := r_1 \cdots r_i$ are reduced expressions and $ur_i = rv$. It follows from Theorem 9.12 that

$$\begin{aligned}(Bur_iB) \cup (BuB) &= (BrB)(Br_1B) \cdots (Br_iB)\\ &= (BrvB) \cup (BvB).\end{aligned}$$

Hence $BuB = BvB$ and so $u = v$, as required. □

In order to prove that a pair (W, R) which satisfies the Exchange Condition is a Coxeter system we first show that any two reduced expressions for an element of W are related by a sequence of simple transformations.

Suppose that W is a group generated by a subset R of elements of order 2. If $r, s \in R$ and if m is the order of rs, then $rsr \cdots = srs \cdots$, where both expressions have length m. In other words, the sequences $(r, s, r, \dots)$ and $(s, r, s \dots)$ of length m represent the same element of W. If $\mathbf{r} = (r_1, r_2, \dots, r_\ell)$ and $\mathbf{s} = (s_1, s_2, \dots, s_\ell)$ are reduced expressions for w we say that the sequences are (r, s)-*related* if $\mathbf{s}$ can be obtained from $\mathbf{r}$ by replacing a subsequence $(r, s, r, \dots)$ of length m by the sequence $(s, r, s \dots)$, also of length m. We say that $\mathbf{r}$ is *homotopic* to $\mathbf{s}$ if $\mathbf{r} = \mathbf{s}$ or if there are elements $\mathbf{r} = \mathbf{r}_1$, $\mathbf{r}_2$, …, $\mathbf{r}_k = \mathbf{s}$ such that for $1 \leq i < k$, $\mathbf{r}_i$ is (r, s)-related to $\mathbf{r}_{i+1}$ for some $r, s \in R$, (cf. Tits (1969)).

9.17 Theorem. *If (W, R) satisfies the Exchange Condition, then for all $w \in W$, every two reduced expressions for w are homotopic.*

Proof. We shall prove the theorem by induction on $\ell = \ell(w)$. It is certainly true when $\ell = 1$ and we may suppose that it is true for all elements w' such that $\ell(w') < \ell$.

Suppose, by way of contradiction, that

$$\mathbf{r} := (r_1, r_2, \dots, r_\ell) \quad \text{and} \quad \mathbf{s} := (s_1, s_2, \dots, s_\ell)$$

are two reduced expressions for w which are not homotopic. It follows from the induction hypothesis that $r_1 \neq s_1$ and $r_\ell \neq s_\ell$.

Now $\ell(s_1 w) = \ell(w) - 1$ and from the Exchange Condition it follows that $s_1 r_1 \cdots r_{i-1} = r_1 \cdots r_i$, for some i. Let

$$\mathbf{s}' := (s_1, r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_\ell).$$

Then $\mathbf{s}'$ is also a reduced expression for w and by induction it is homotopic to $\mathbf{s}$. If $i < \ell$, then $\mathbf{r}$ and $\mathbf{s}'$ both end with r_ℓ and so $\mathbf{r}$ is homotopic to $\mathbf{s}'$ by induction. But then $\mathbf{r}$ is homotopic to $\mathbf{s}$, contrary to assumption. Thus $i = \ell$ and $\mathbf{s}' = (s_1, r_1, \ldots, r_{\ell-1})$.

Applying the same argument to $\mathbf{s}'$ and $\mathbf{r}$ instead of $\mathbf{r}$ and $\mathbf{s}$ we find that there is a reduced expression $\mathbf{r}' := (r_1, s_1, r_1, \ldots, r_{\ell-2})$ for w not homotopic to $\mathbf{s}'$. Continuing in this way we eventually obtain

$$\mathbf{r}^\circ := (r_1, s_1, r_1, s_1, \ldots, \ ) \quad \text{and} \quad \mathbf{s}^\circ := (s_1, r_1, s_1, r_1, \ldots, \ )$$

such that $\mathbf{r}^\circ$ and $\mathbf{s}^\circ$ are reduced expressions for w but $\mathbf{r}^\circ$ and $\mathbf{s}^\circ$ are not homotopic. But then $(r_1 s_1)^\ell = 1$ and consequently $\mathbf{r}^\circ$ and $\mathbf{s}^\circ$ are (r_1, s_1)-related, a contradiction. □

9.18 Corollary. *If (W, R) satisfies the Exchange Condition, then (W, R) is a Coxeter system. In fact, W is isomorphic to the group with generators $\{ x_i \mid i \in I \}$ and relations $(x_i x_j)^{m_{ij}} = 1$, where m_{ij} is the order of $x_i x_j$ in W.*

Proof. Suppose that $R := \{ x_i \mid i \in I \}$ and that m_{ij} is the order of $x_i x_j$. We must show that all relations satisfied by the elements of W are consequences of the relations $(x_i x_j)^{m_{ij}} = 1$. We do this by showing that if G is a group and if $f : R \to G$ satisfies

$$(f(x_i) f(x_j))^{m_{ij}} = 1$$

for all $i, j \in I$, then f extends to a homomorphism from W to G.

If $r_1 r_2 \cdots r_\ell$ is a reduced expression for w, then the element $f(w)$ is defined to be $f(r_1) f(r_2) \cdots f(r_\ell)$. Any two reduced expressions for w are homotopic and therefore $f(w)$ is well-defined.

If $r \in R$ and $\ell(rw) = \ell(w) + 1$, then $f(rw) = f(r) f(w)$. On the other hand, if $\ell(rw) = \ell(w) - 1$, write $w' := rw$ and observe that $f(rw') = f(r) f(w')$. That is, $f(rw) = f(r) f(w)$ in both cases. It follows by induction on $\ell(v)$ that $f(vw) = f(v) f(w)$ and so f is a homomorphism. □

Presently we shall see that the converse of this result is true: namely, every Coxeter system satisfies the Exchange Condition.

9.19 Corollary. *The Weyl group of a BN-pair is a Coxeter group* □

It is easy to give a direct proof that the particular Weyl groups S_n and $\mathbb{Z}_2 \wr S_n$ are Coxeter groups (see Exercises 9.5 and 9.7.)

The Coxeter systems that occur in Corollary 9.18 have the property that m_{ij} is the order of $x_i x_j$. In fact, if (W, R) is any Coxeter system with generators $R := \{x_1, \ldots, x_n\}$ and relations $(x_i x_j)^{m_{ij}} = 1$, it is always the case that m_{ij} is the order of $x_i x_j$ (Exercise 9.8).

Reflections and the Strong Exchange Condition

Suppose that (W, R) is a Coxeter system and consider the set

$$T := \{\, wrw^{-1} \mid w \in W,\ r \in R\,\}. \tag{9.20}$$

The elements of T are called the *reflections* of W. (See Exercise 9.8, Tits (1966) or Grove and Benson (1985) for a justification of this terminology.) The group W acts on T by conjugation and hence it also acts on the set $\mathcal{P}(T)$ of subsets of T.

The set $\mathcal{P}(T)$ becomes an abelian group when addition of subsets $E, F \subseteq T$ is defined to be their symmetric difference, which in this case we write as $E + F$. The action of W preserves this addition and therefore we can form the semidirect product $\mathcal{P}(T)W$. Multiplication in $\mathcal{P}(T)W$ is given by

$$(E_1, w_1)(E_2, w_2) = (E_1 + w_1 E_2 w_1^{-1}, w_1 w_2).$$

This approach to the study of W is due to Dyer (1987). His fundamental observation is

9.21 Theorem. *If (W, R) is a Coxeter system, there is a homomorphism $\delta : W \to \mathcal{P}(T)W$ such that $\delta(r) = (\{r\}, r)$ for all $r \in R$.*

Proof. First define $\delta : R \to \mathcal{P}(T)W$ by $\delta(r) = (\{r\}, r)$. Then for all $i, j \in I$ and $k > 0$ we have

$$(\delta(x_i)\delta(x_j))^k = (\{x_i\} + \{x_i x_j x_i\} + \cdots + \{(x_i x_j)^{2k-1} x_i\}, (x_i x_j)^k).$$

But $(x_i x_j)^h x_i = (x_i x_j)^{m_{ij}+h} x_i$ and therefore $(\delta(x_i)\delta(x_j))^{m_{ij}} = (\emptyset, 1)$, the identity element of $\mathcal{P}(T)W$. It follows that δ extends to a homomorphism $W \to \mathcal{P}(T)W$. □

We can write $\delta(w) = \big(D(w), w\big)$, where $D : W \to \mathcal{P}(T)$ satisfies

$$\begin{aligned} D(r) &= \{r\}, \quad \text{for all } r \in R, \text{ and} \\ D(w_1 w_2) &= D(w_1) + w_1 D(w_2) w_1^{-1}. \end{aligned} \tag{9.22}$$

Until further notice we assume that W is a group generated by a set R of elements of order 2 and that there is a function $D : W \to \mathcal{P}(T)$ satisfying (9.22), where T is given by (9.20).

Given a sequence $(r_1, r_2, \ldots, r_k)$ of elements of R, define $t_1, t_2, \ldots\ t_k \in T$ by

$$t_i := r_1 \cdots r_{i-1} r_i r_{i-1} \cdots r_1 \tag{9.23}$$

From (9.22) we have

$$\begin{aligned} D(r_1r_2\cdots r_k) &= \{t_1\}+\{t_2\}+\cdots+\{t_k\} \\ &\subseteq \{t_1,t_2,\ldots,t_k\}. \end{aligned} \tag{9.24}$$

9.25 Lemma. *The expression $w := r_1r_2\cdots r_k$ is reduced if and only if the t_i are distinct, in which case $D(w)=\{t_1,t_2,\ldots,t_k\}$.*

Proof. It follows from (9.24) that $|D(w)|\le \ell(w)$ and hence if $r_1r_2\cdots r_k$ is not reduced, the t_i are not distinct. On the other hand, if $t_i=t_j$ for some $i<j$, then

$$w = r_1\cdots r_{i-1}r_{i+1}\cdots r_{j-1}r_{j+1}\cdots r_k$$

and so $r_1r_2\cdots r_k$ is not reduced. □

9.26 Corollary. $\ell(w)=|D(w)|$. □

Note that from (9.22) we have $|D(w_1w_2)|\equiv |D(w_1)|+|D(w_2)| \pmod 2$ and so the map

$$W\to\{\pm1\} : w\mapsto (-1)^{\ell(w)}$$

is a homomorphism onto $\{\pm1\}$.

9.27 Lemma. *For all $t\in T$, $t\in D(t)$.*

Proof. For all $r\in R$ we have $r\in D(r)$ by assumption. If $t\in T$, then $t=wsw^{-1}$ for some $s\in R$ and by induction on $\ell(w)$ we may suppose that $t=rt'r$, where $r\in R$, $t'\in T$ and $t'\in D(t')$. Then $D(t)=\{r\}+rD(t')r+\{rt'rt'r\}$. As $rt'r=r$ if and only if $rt'r=rt'rt'r$, it follows that $t\in D(t)$. □

9.28 Theorem. $D(w)=\{\, t\in T \mid \ell(tw)<\ell(w)\,\}$.

Proof. Suppose that $w := r_1r_2\cdots r_k$ is a reduced expression for w. Then by Lemma 9.25, $D(w)=\{t_1,t_2,\ldots,t_k\}$, where the t_i are given by (9.23). Thus $t\in D(w)$ implies $tw=r_1\cdots r_{i-1}r_{i+1}\cdots r_k$ for some i and so $\ell(tw)<\ell(w)$.

Conversely, suppose that $t\notin D(w)$, for some $t\in T$. Then $D(tw)=D(t)+tD(w)t$ and therefore $t\in D(tw)$ by Lemma 9.27. From the previous paragraph we have $\ell(w)<\ell(tw)$, and this completes the proof. □

This result is known as the *Strong Exchange Condition*. The Exchange Condition itself can be reworded as:

for all $r\in R$, $\ell(rw)\le\ell(w)$ implies $r\in D(w)$.

However, we know from the remark following Corollary 9.26 that $\ell(rw)\equiv \ell(w)+1 \pmod 2$ and so it is never the case that $\ell(rw)=\ell(w)$. Thus the Strong Exchange Condition implies the Exchange Condition considered previously.

9.29 Theorem. *If W is a group generated by a set R of elements of order 2, if T is the set of conjugates of the elements of R, and if $D : W \to \mathcal{P}(T)$ satisfies (9.22), then (W, R) is a Coxeter system.*

Proof. This follows from Corollary 9.18 $\square$

This brings us full circle and shows that Coxeter groups are precisely those groups satisfying the Exchange Condition.

Parabolic Subgroups of Coxeter Groups

Let (W, R) be a Coxeter system. For $J \subseteq R$, let W_J be the subgroup of W generated by J. The groups W_J are the *parabolic subgroups* of W.

9.30 Theorem. *For all $J \subseteq R$, (W_J, J) is a Coxeter system. The length of $w \in W_J$ in W_J coincides with its length in W.*

Proof. The restriction of the map D, considered in the previous section, to W_J satisfies (9.22) and takes its values in $\mathcal{P}(T_J)$, where $T_J := \{\, wrw^{-1} \mid r \in J,\ w \in W_J \,\}$. Thus by Theorem 9.29, (W_J, J) is a Coxeter system. The length of w in W_J is $|D(w)| = \ell(w)$. $\square$

9.31 Lemma. *For all reduced expressions $r_1 r_2 \cdots r_k$ for $w \in W$, the set $\{r_1, r_2, \ldots, r_k\}$ depends only on w and not on the sequence $(r_1, r_2, \ldots, r_k)$.*

Proof. Let $J := \{r_1, r_2, \ldots, r_k\}$. Then $w \in W_J$. If $r'_1 r'_2 \cdots r'_k$ is another reduced expression for w, then $\ell(r'w) < \ell(w)$ and so by Theorem 9.28, $r'_1 \in D(w) \subseteq T_J$. But $\ell(r'_1) = 1$ and so by Theorem 9.30, $r'_1 \in J$. It follows by induction that $\{r'_1, r'_2, \ldots, r'_k\} = J$. $\square$

9.32 Corollary. *If $r_1 r_2 \cdots r_k$ is a reduced expression for $w \in W_J$, then $r_i \in J$, for all i.* $\square$

9.33 Corollary. *For $J, K \subseteq R$,*

(i) $W_{J \cap K} = W_J \cap W_K$, *and*

(ii) $W_{J \cup K} = \langle\, W_J, W_K \,\rangle$ $\square$

9.34 Lemma. *If $w := r_1 r_2 \cdots r_k$ with $r_i \in R$ for all i, then there is a sequence $1 \le i_1 < i_2 < \cdots < i_\ell \le k$ such that $w = r_{i_1} r_{i_2} \cdots r_{i_\ell}$ is a reduced expression for w.*

Proof. Choose i as small as possible so that $r_i \cdots r_k$ is reduced. If $i = 1$, then w is reduced as written. If $i > 1$, then $\ell(r_{i-1} r_i \cdots r_k) < \ell(r_i \cdots r_k)$ and by the Exchange Condition there exists j such that

$$r_{i-1} r_i \cdots r_{j-1} = r_i \cdots r_j.$$

But then $w = r_1 \cdots r_{i-2} r_i \cdots r_{j-1} r_{j+1} \cdots r_k$ and the result follows by induction. □

9.35 Corollary. *For $J, K, L \subseteq R$,*

(i) $W_J \cap W_K W_L = (W_J \cap W_K)(W_J \cap W_L)$,

(ii) $W_J(W_K \cap W_L) = W_J W_K \cap W_J W_L$.

Proof. Exercise 9.11 □

Complexes

In Chapters 5 and 7 we described the buildings derived from projective and polar geometries. In both cases the building was defined to be a set of flags of subspaces together with a collection of apartments; the apartments being certain distinguished sets of flags. We have seen that this provides a setting for the BN-pair structure of the group of isometries of the geometry. In all cases the Weyl group of the BN-pair acts regularly on the set of maximal flags of an apartment. It is possible to axiomatize this situation in various ways: see, for example, Tits (1981), Tits (1986) or Shult (1983).

Let I be a finite set. Following Tits (1981) we define a *complex over I* to consist of a set V (of *vertices*), a map $\tau : V \to I$, and a set Δ of subsets of V (called *simplexes*), such that

(i) *Every singleton $\{v\}$, for $v \in V$, is a simplex.*

(ii) *Every subset of a simplex is a simplex.*

(iii) *For every simplex A, the restriction of τ to A is one-to-one.*

(A set of subsets of a set V satisfying (i) and (ii) is a *simplicial complex*.)

If A is a simplex, the set $\tau(A)$ is called the *type* of A. The *rank* of A is $|A|$ and the *corank* is $|I \setminus \tau(A)|$. The *rank* of Δ is $|I|$. The simplexes of type I are the *chambers* of Δ and the simplexes of corank 1 are the *panels*.

The complex is said to be *thin* if each panel is in exactly two chambers. It is *thick* if each panel is in at least three chambers.

For example, the set Δ of flags of totally isotropic subspaces of a polar geometry of Witt index m is the set of simplexes of a complex over $\{1, 2, \ldots, m\}$. The vertices are the totally isotropic subspaces and the type of a subspace is its dimension. The complex is thick (Lemma 9.7) unless it comes from an orthogonal geometry of dimension $2m$.

Each apartment of Δ is a thin complex over I (Lemma 9.4) and the Weyl group W acts transitively on the flags of a given type. Thus the flags of type J can be identified with the cosets of $W_{I \setminus J}$. (Here we identify I with the set of generators for W given after Theorem 9.3.)

Coxeter Complexes

We shall construct a thin complex $\Sigma(W,R)$ over I for every Coxeter group W with generators $R := \{\, r_i \mid i \in I \,\}$. For $J \subseteq I$ we define W_J to be the subgroup $\langle\, r_i \mid i \in J \,\rangle$. This is consistent with our earlier notation if we identify I with R. For $i \in I$ we let $W^i := W_{I\setminus\{i\}}$ and we define the vertices of $\Sigma(W,R)$ to be the cosets wW^i, where $w \in W$ and $i \in I$. The type map is defined by $\tau(wW^i) := i$. A subset S of V is a simplex if its intersection $\bigcap_{\alpha\in S}\alpha$ is non-empty. It is easy to verify that this is a complex over I: the *Coxeter complex* of (W,R). Another, possibly more useful, description of $\Sigma(W,R)$ can be obtained from the following lemma.

9.36 Lemma. *The map $S \mapsto \bigcap_{\alpha\in S}\alpha$ is an inclusion reversing bijection between the simplexes of $\Sigma(W,R)$ and the set of cosets $\{\, wW_J \mid J \subseteq I, w \in W \,\}$.*

Proof. Each simplex S has the form $S := \{wW^{i_1}, wW^{i_2}, \dots, wW^{i_k}\}$ for some $w \in W$. If $K := \{i_1, i_2, \dots, i_k\}$, then it follows from Corollary 9.33 $\bigcap_{\alpha\in S}\alpha = wW_{I\setminus K}$. If $wW_{I\setminus K} \subseteq w'W^i$, then by Corollary 9.32, $i \in K$. Thus $S \mapsto \bigcap_{\alpha\in S}\alpha$ is a bijection and it is clear that it reverses inclusion. □

In particular, the chambers of $\Sigma(W,R)$ correspond to the elements of W and the panels correspond to pairs $\{w, ws\}$, $s \in R$. Thus $\Sigma(W,R)$ is thin. Note that the reflection $t := wsw^{-1}$ fixes the panel $\{w, ws\}$ and interchanges w and ws.

A *morphism* from a complex Δ' over I with vertex set V' and type map τ' to the complex Δ, also over I, is a map $\varphi\colon V' \to V$ which takes each simplex of Δ' to a simplex of Δ and which preserves the type, i.e., $\tau' = \tau\varphi$.

For example, each element of W acts on $\Sigma(W,R)$ by multiplication on the left and this action is an automorphism of $\Sigma(W,R)$.

In general, if A is a simplex of the complex Δ, the *link* of A is the set

$$\operatorname{link}(A) := \{\, B \in \Delta \mid A \cap B = \emptyset \text{ and } A \cup B \in \Delta \,\}.$$

The link of A is a complex over $I \setminus \tau(A)$.

In Theorem 9.30 we showed that for each $J \subseteq R$, (W_J, J) is a Coxeter system. Thus $\Sigma(W_J, J)$ is a well-defined Coxeter complex and moreover,

9.37 Theorem. *If $S \in \Sigma(W,R)$ and $\bigcap_{\alpha\in S}\alpha = wW_J$, then $\operatorname{link}(S)$ is isomorphic to $\Sigma(W_J, J)$.*

Proof. Since we are working up to isomorphism we may suppose that $\bigcap_{\alpha\in S}\alpha = W_J$. For $T \in \operatorname{link}(S)$, $T \mapsto \bigcap_{\alpha\in S\cup T}\alpha$ gives a bijection between link(S) and $\{ wW_K \mid w \in W_J,\ K \subseteq J \}$ which reverses inclusion. Combining this with the inverse of the bijection of Lemma 9.34 (applied to $\Sigma(W_J, J)$) gives the desired isomorphism. □

If $J := \{r_1, r_2\}$ and $r_i r_j$ has order m_{ij}, then W_J is the dihedral group $\langle r_i, r_j \mid r_i^2 = r_j^2 = (r_i r_j)^{m_{ij}} = 1 \rangle$ of order $2m_{ij}$. The complex $\Sigma(W_J, J)$ has m_{ij} vertices of type i, m_{ij} vertices of type j and $2m_{ij}$ edges (i.e., simplexes of rank 1). Thus $\Sigma(W_J, J)$ is a $2m_{ij}$-gon. This means that for any simplex $S \in \Sigma(W, R)$ of type $I \setminus \{i, j\}$, link(S) is a $2m_{ij}$-gon. It follows that the Coxeter matrix of (W, R) is determined by $\Sigma(W, R)$ and therefore $\Sigma(W, R)$ determines (W, R) up to isomorphism.

Buildings

Suppose that M is the Coxeter matrix of the Coxeter system (W, R), where $R := \{ r_i \mid i \in I \}$. An *apartment of type* M of a complex Δ over I is the image of an injective morphism $\Sigma(W, R) \to \Delta$.

A *building of type* M is a complex Δ over I together with a collection of apartments of type M such that for all simplexes $A, B \in \Delta$,

(i) *there is an apartment containing* A *and* B.

(ii) *if* Σ *and* Σ' *are apartments containing* A *and* B, *there is an isomorphism* $\Sigma \to \Sigma'$ *fixing* A *and* B.

Tits (1976) has determined all buildings of rank at least 3 with finite apartments.

These axioms correspond to Theorems 5.2 and 5.4 in the case of the building of a projective geometry and to Theorems 7.15 and 7.17 for polar geometries. Of course it is first necessary to identify the apartments as defined in Chapters 5 and 7 with the apartments as defined here.

For the buildings associated with projective and polar geometries the isomorphisms required in (ii) above can be taken to be automorphisms of the building (see Theorems 5.4 and 7.17). We shall see below that if Σ and Σ' have a common chamber C, there is a unique isomorphism $\varphi_C : \Sigma \to \Sigma'$ fixing C.

Two chambers C and C' are said to be *i-adjacent* if $C = C'$ or if $C \cap C'$ is a panel of type $I \setminus \{i\}$.

9.38 Theorem. $\mathrm{Aut}(\Sigma(W, R)) \simeq W$.

Proof. As in Lemma 9.36, the chambers of $\Sigma(W, R)$ correspond to the elements of W. Moreover, w and w' are i-adjacent if and only if $w' = wr_i$. If φ is and automorphism of $\Sigma(W, R)$, then φ preserves the type. Hence if w

and w' are i-adjacent, then $\varphi(w)$ and $\varphi(w')$ are i-adjacent. That is, $\varphi(wr_i) = \varphi(w)r_i$ for all i. It follows by induction on $\ell(w)$ that $\varphi(w) = \varphi(1)w$. The action of φ on $\Sigma(W,R)$ is completely determined by its action on chambers and therefore φ is left multiplication by $\varphi(1) \in W$. Thus $\varphi \mapsto \varphi(1)$ is an isomorphism between $\mathrm{Aut}(\Sigma(W,R))$ and W. □

9.39 Theorem. *If the apartments Σ and Σ' of the building Δ have a common chamber C, there is a unique isomorphism $\varphi^{\Sigma,\Sigma'} : \Sigma \to \Sigma'$ fixing C. Moreover, $\varphi^{\Sigma,\Sigma'}$ fixes every simplex of $\Sigma \cap \Sigma'$.*

Proof. The existence of $\varphi^{\Sigma,\Sigma'}$ comes from axiom (ii). If $\psi : \Sigma \to \Sigma'$ is another isomorphism fixing C, then $\psi^{-1}\varphi^{\Sigma,\Sigma'}$ is an automorphism of Σ fixing C. But from Theorem 9.38 the only automorphism fixing C is the identity. Hence $\psi = \varphi^{\Sigma,\Sigma'}$. If A is a simplex in $\Sigma \cap \Sigma'$, then again from axiom (ii) there is an isomorphism $\varphi : \Sigma \to \Sigma'$ fixing A and C. But then $\psi = \varphi^{\Sigma,\Sigma'}$ and therefore $\varphi^{\Sigma,\Sigma'}$ fixes A. □

If Σ is an apartment and C is a chamber of Σ, we can now produce a morphism $\rho_C^\Sigma : \Delta \to \Sigma$ which fixes every simplex of Σ. If $A \in \Delta$, then by axiom (i) there is an apartment Σ' containing A and C. Define $\rho_C^\Sigma(A)$ to be $\varphi^{\Sigma',\Sigma}(A)$. If Σ' is another apartment containing A and C, then $\varphi^{\Sigma'',\Sigma'}$ fixes A and C and we have $\varphi^{\Sigma'',\Sigma} = \varphi^{\Sigma',\Sigma}\varphi^{\Sigma'',\Sigma'}$, by Theorem 9.39. Thus $\varphi^{\Sigma'',\Sigma}(A) = \varphi^{\Sigma',\Sigma}(A)$ and hence $\rho_C^\Sigma(A)$ is well-defined. The morphism

$$\rho_C^\Sigma : \Delta \to \Sigma$$

is called the *retraction* of Δ onto Σ *centred* at C.

To explore these ideas further it is useful to introduce a notion of distance for Δ. First of all, a *gallery* is a sequence of chambers $C_0, C_1, \ldots, C_m$ such that for all i, C_{i-1} and C_i are k_i-adjacent for some k_i. The gallery is said to *stutter* if $C_{i-1} = C_i$ for some i.

The *distance* $d(C,D)$ between two chambers C and D is the least integer m such that there is a gallery $C = C_0, C_1, \ldots, C_m = D$. The galleries of length $d(C,D)$ from C to D are said to be *minimal* (or *geodesic*).

9.40 Theorem. *If C and D are chambers in the apartment Σ of Δ, every geodesic from C to D lies entirely in Σ.*

Proof. Suppose that $C = C_0, C_1, \ldots, C_m = D$ is a gallery of length $m = d(C,D)$ from C to D and that for some i, $C_i \in \Sigma$ but $C_{i-1} \notin \Sigma$. Let $A := C_{i-1} \cap C_i$. The complex Σ is thin and therefore A is in exactly one other chamber of Σ other than C_i, say E. Let $\rho := \rho_E^\Sigma$ be the retraction of Δ onto Σ centred at E. As ρ is a morphism it preserves i-adjacency and therefore $C = \rho(C_0), \rho(C_1), \ldots, \rho(C_m) = D$ is also a gallery from C to D. But

$\rho(C_{i-1})$ is a chamber of Σ containing A and therefore $\rho(C_{i-1}) = E$ because, by construction, the only chamber mapped onto E by ρ is E itself. Thus $\rho(C_0)$, $\rho(C_1)$, $\ldots$, $\rho(C_m)$ stutters, contradicting the minimality of m. □

9.41 Corollary. *If C and D are chambers in the apartment Σ of Δ, then the distance from C to D calculated in Σ coincides with the distance from C to D in Δ.* □

It turns out that we have already investigated geodesics in Σ, but in another language. Using the correspondence between the chambers of $\Sigma(W, R)$ and the elements of W given by Lemma 9.36 and the fact that for $w \neq w'$, w is i-adjacent to w' if and only if $w' = wr_i$ we see that the non-stuttering galleries from w to w' correspond to the sequences $(r_{i_1}, r_{i_2}, \ldots, r_{i_m})$ such that $w^{-1}w' = r_{i_1}r_{i_2}\cdots r_{i_m}$. The geodesics correspond to the reduced expressions for $w^{-1}w'$. In particular, $d(w, w') = \ell(w^{-1}w')$.

Using Theorem 9.40 and the following lemma we transfer these observations to the building Δ.

9.42 Lemma. *If $C = C_0$, C_1, $\ldots$, $C_m = D$ is a non-stuttering gallery in which, for $1 \le k \le m$, C_{k-1} is i_k-adjacent to C_k, then the expression $w := r_{i_1}r_{i_2}\cdots r_{i_m}$ is reduced if and only if the gallery is minimal. Moreover, the element $w(C, D) := r_{i_1}r_{i_2}\cdots r_{i_m} \in W$ is the same for all minimal galleries from C to D.*

Proof. If the gallery is minimal, then it lies in an apartment containing C and D and we have just seen that in this case $r_{i_1}r_{i_2}\cdots r_{i_m}$ is reduced. Conversely, suppose the expression is reduced. By induction, $C_1, C_2, \ldots, C_m$ is a minimal gallery and hence it lies in the Σ, where Σ is any apartment containing C_1 and C_m. Let ρ be the retraction of Δ onto Σ centred at C_1. Then $\rho(C_0)$, C_1, $\ldots$, C_m is a non-stuttering gallery of Σ and it is minimal because $r_{i_1}r_{i_2}\cdots r_{i_m}$ is reduced. But ρ cannot increase distances and therefore C_0, C_1, $\ldots$, C_m is minimal also.

The last part of the lemma follows from the fact that the minimal galleries from C to D lie in every apartment containing C and D and we have already seen that the assertion is true in an apartment. □

Notice that Coxeter complexes and buildings are *connected* in the sense that for any two chambers C and D, there is a gallery from C to D.

Chamber Systems

There is another approach to the study of buildings and related geometries via the chamber systems of Tits (1981). A *chamber system* over the set I

is a set $\mathcal{C}$ (whose elements are called *chambers*) together with a collection $\{\, \pi_i \mid i \in I \,\}$ of partitions of $\mathcal{C}$. Chambers C and D are said to be *i-adjacent* if $C, D \in \pi_i$. We shall always assume that $|\pi_i| \geq 2$ for all i.

If $f := (i_1, i_2, \ldots, i_\ell)$ is a sequence of elements of I, a *gallery* of type f is a sequence of chambers $C_0, C_1, \ldots, C_\ell$ such that for $1 \leq j \leq \ell$, the chambers C_{j-1} and C_j are i_j-adjacent. Given $J \subseteq I$, the gallery is said to be of *type* J if all $i_j \in J$. Being connected by a gallery of type J is an equivalence relation on the chambers. The equivalence classes are called the *J-residues* of $\mathcal{C}$.

For example, if (W, R) is a Coxeter system, where $R := \{\, r_i \mid i \in I \,\}$, then the elements of W are the chambers of a chamber system in which w and w' are defined to be i-adjacent if $w = w'$ or $w = w' r_i$. The J-residues are the cosets wW_J.

If Δ is a building, its set $\mathcal{C}$ of chambers is a chamber system in which two chambers are i-adjacent if they have a common panel of type $I \setminus \{i\}$. Moreover, two chambers are connected by a gallery of type $I \setminus \{i\}$ if and only if they have a common vertex of type i. (It is enough to check this in an apartment, in which case it follows from Theorem 9.37.) Consequently the complex Δ can be recovered from $\mathcal{C}$. The vertices of type i can be identified with the $I \setminus \{i\}$-residues of $\mathcal{C}$. A set S of residues corresponds to a simplex if and only if $\bigcap_{\alpha \in S} \alpha \neq \emptyset$. Notice that the notation introduced for chamber systems is consistent with that for complexes.

If $f := (i_1, i_2, \ldots, i_\ell)$ is a sequence of elements of I, let $r_f := r_{i_1} r_{i_2} \cdots r_{i_\ell}$ be the corresponding element of W. Call f *reduced* if $r_{i_1} r_{i_2} \cdots r_{i_\ell}$ is a reduced expression for r_f.

At the end of the previous section we showed that if Δ is a building with apartments isomorphic to $\Sigma(W, R)$ and if $\mathcal{C}$ is its set of chambers, there is a map $w : \mathcal{C} \times \mathcal{C} \to W$ such that $w(C, D) = r_f$ with f reduced if and only if there is a gallery of type f from C to D. It turns out that this property suffices to characterize those chamber systems which arise from buildings. See Tits (1981) and Tits (1986) for the details.

It is not hard to reconstruct Δ from $\mathcal{C}$ and the map $w : \mathcal{C} \times \mathcal{C} \to W$. Two chambers C and D are i-adjacent if and only if $C = D$ or $w(C, D) = r_i$ and the chambers of the simplexes of Δ of type J are the maximal subsets $\mathcal{D} \subseteq \mathcal{C}$ such that $w(\mathcal{D} \times \mathcal{D}) \subseteq W_J$.

For the building Δ of flags of proper subspaces of a vector space V, the map $w : \mathcal{C} \times \mathcal{C} \to W$ can be described as follows. The group $G := SL(V)$ acts transitively on the set $\mathcal{C}$ of maximal flags and if B is the stabilizer of the maximal flag M, then the orbits of G on $\mathcal{C} \times \mathcal{C}$ correspond to the double cosets of B in G. Theorems 5.2 and 5.5 show that each orbit has a unique representative of the form $(M, w(M))$ for some $w \in W$. Thus $w = w(C, D)$, where (C, D) is any member of the orbit of $(M, w(M))$. Theorem 9.3 and

Lemma 9.11 show that the same interpretation holds for the building of flags of totally isotropic subspaces of a polar geometry.

Diagram Geometries

The examples of buildings we have seen so far are sets of flags of an underlying geometry. In fact it is possible to describe all buildings in this way.

This approach, which leads to geometries considerably more general than buildings, has been investigated extensively by Buekenhout (1979, 1981, 1982, ...) and others. See Tits (1980) for some of the history.

To each geometry there is an associated diagram generalizing the Coxeter-Dynkin diagrams introduced earlier in this chapter. This diagram turns out to be a convenient and succinct way to give the axioms for the geometry. In the very brief account given here we limit ourselves to geometries associated with Coxeter-Dynkin diagrams.

Once again we follow Tits (1981) but with some notation taken from Shult (1983).

Let I be a finite set. A *geometry over* I is a multipartite graph Γ whose components are indexed by elements of I. That is, there is a map $\tau : V \to I$, where V is the set of vertices of Γ, such that, for all $i \in I$, no two vertices of $\tau^{-1}(i)$ are joined by an edge. Thus there are no loops in Γ. Two vertices are said to be *adjacent* if they are joined by an edge. All that matters is the adjacency relation and so we may suppose that Γ has no multiple edges.

We call τ the *type* map and for $x \in V$, $\tau(x)$ is called the *type* of x. Similarly, for $X \subseteq V$, $\tau(X)$ is the *type* of X. The elements of $\tau^{-1}(i)$ are sometimes called the *varieties* of type i.

A *flag* F of Γ is a complete subgraph; i.e., every pair of distinct vertices of F are adjacent. The *rank* of F is $|F|$ and the *corank* is $|I \setminus \tau(F)|$. The *rank* of the geometry Γ is $|I|$. A *chamber* is a flag of type I and a *panel* is a flag of corank 1.

The *residue* of a flag F is the geometry $\mathrm{Res}(F)$ over $I \setminus \tau(F)$ whose vertices are the elements of V which are adjacent to every element of F. Two elements of $\mathrm{Res}(F)$ are defined to be adjacent if they are adjacent in Γ. The type map of $\mathrm{Res}(F)$ is the restriction of τ.

The set of flags of Γ are the simplexes of a complex over I with the same vertex set and type map as Γ. We call this the *flag complex* $\Delta(\Gamma)$. If F is a flag, the flag complex of $\mathrm{Res}(F)$ is the link of F in $\Delta(\Gamma)$.

We want to restrict our attention to geometries which resemble buildings and to this end we introduce some additional axioms. In order to describe these axioms it will be convenient to have available the following standard definitions from graph theory.

A *path* of length m from x to y is a sequence of vertices $x = x_0$, x_1, $\ldots$, $x_m = y$ such that for $1 \leq i \leq m$, x_{i-1} is adjacent to x_i. If x and y are connected by a path, the *distance* $d(x, y)$ from x to y is the minimum length of a path from x to y. The relation of being joined by a path is an equivalence relation on the vertices and the equivalence classes are the *connected components* of the graph. The graph is *connected* if there is only one connected component. The *diameter* of a connected graph is the maximum value of $d(x, y)$. A *circuit* is a path x_0, x_1, $\ldots$, $x_m = x_0$ such that $|\{x_1, \ldots, x_m\}| = m \geq 3$. The *girth* is the minimum length of a circuit.

A geometry Γ is said to be *residually connected* if

(i) *every panel is contained in a chamber, and*

(ii) *for every flag F of corank ≥ 2, $\mathrm{Res}(F)$ is non-empty and connected.*

Diagram geometries are obtained by imposing additional conditions on the rank 2 residues. Essentially, we require the rank 2 residues to be of certain fixed types. Thus to obtain interesting diagram geometries in general we need a good supply of rank 2 geometries.

Projective planes and the generalized quadrangles defined at the end of Chapter 8 provide some examples. In both cases we have a set of points and a set of lines, which we define to be the varieties of type 0 and type 1 respectively; i.e., we take $I := \{0, 1\}$. A point is adjacent to a line if it is on the line. The geometry of a projective plane is thus a graph of diameter 3 and girth 6, whereas the geometry of a generalized quadrangle has diameter 4 and girth 8.

More generally, for $m \geq 2$, a geometry of rank 2 is called a *generalized m-gon* if it has diameter m, girth $2m$ and every vertex is adjacent to at least two other vertices. We obtain a generalized m-gon from an ordinary m-gon by taking the varieties of type 0 to be the vertices of the m-gon and the varieties of type 1 to be the edges. It is easy to see that a generalized 3-gon is either a triangle or a projective plane. A generalized 2-gon is a complete bipartite graph with at least 2 vertices in each component (see Tits (1976)).

Now let $M := (m_{ij})$ be a Coxeter matrix whose rows and columns are indexed by I. A *diagram geometry of type M* is a residually connected geometry over I such that, for all $i, j \in I$ and $i \neq j$, the residue of any flag of type $I \setminus \{i, j\}$ is a generalized m_{ij}-gon. The *diagram* of the geometry is the Coxeter-Dynkin diagram of M. That is, it is a graph with vertex set I such that for all $i \neq j$ there are $m_{ij} - 2$ edges from i to j.

It is easy to see that if Γ is a diagram geometry and F is a flag of type J, then $\mathrm{Res}(F)$ is a diagram geometry whose diagram is obtained by removing the elements of J (and the edges incident with them) from the diagram for Γ.

To illustrate the type of argument that can be used to study these geometries we prove the following lemmas of Tits (1981) and Buekenhout (1979).

9.43 Lemma. *If Γ is a residually connected geometry over I and if $J \subseteq I$ has at least two elements, any two varieties x, y of Γ can be joined by a path all of whose vertices, except possibly x and y, have their type in J.*

Proof. By assumption Γ is connected and so the lemma is true when $J = I$. If $J \neq I$ and $k \in I \setminus J$, then by induction on $|I \setminus J|$, the lemma is true in the residue of every variety of type k. Again by induction, there is a path $x = x_0$, $x_1, \ldots, x_m = y$ such that for $0 < i < m$, $\tau(x_i) \in J \cup \{k\}$. If $\tau(x_i) = k$, then x_{i-1} and x_{i+1} are in the residue of x_i and so there is a path from x_{i-1} to x_{i+1} all of whose internal vertices have their type in J. Thus there is a path from x to y with the same property. □

9.44 Lemma. *If Γ is a diagram geometry of type M and if i and j belong to different connected components of the diagram of Γ, then every variety of type i is adjacent to every variety of type j.*

Proof. Suppose that $\tau(x) = i$ and $\tau(y) = j$. We argue by induction on $|I|$. If $|I| = 2$, the geometry is a generalized 2-gon; i.e., a complete bipartite graph, and so the lemma is true in this case. So we may suppose that $|I| > 2$ and without loss of generality there is an element $k \neq j$ in I not in the same component as i. By the previous lemma there is a path $x = x_0$, x_1, $\ldots$, $x_m = y$ such that the types of $x_1, x_2, \ldots, x_{m-1}$ belong to $\{j, k\}$ and where m is chosen to be as small as possible. If $m > 1$, then $x, x_2 \in \mathrm{Res}(x_1)$ and $\tau(x)$ and $\tau(x_2)$ are in different components of the diagram of $\mathrm{Res}(x_1)$. By induction, x is adjacent to x_2, contradicting the minimality of m. Thus $m = 1$ and x is adjacent to y. □

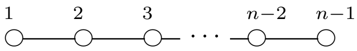
If (W, R) is a Coxeter system, where $R := \{\, r_i \mid i \in I \,\}$, we have seen that the cosets of $W^i := W_{I\setminus\{i\}}$ are the varieties of type i in the Coxeter complex $\Sigma(W, R)$. The *Coxeter geometry* $\Gamma(W, R)$ has the same vertices and type map as $\Sigma(W, R)$ and we define wW^i and $w'W^j$ to be adjacent whenever they are distinct and $wW^i \cap w'W^j \neq \emptyset$ (in which case, by Corollary 9.33, $wW^i \cap w'W^j$ is a coset of $W_{I\setminus\{i,j\}}$).

9.45 Lemma. *A set of vertices is a flag of the geometry $\Gamma(W, R)$ if and only if it is a simplex of $\Sigma(W, R)$.*

Proof. From the definition of $\Sigma(W, R)$, every simplex is a flag of $\Gamma(W, R)$. Conversely, suppose that $\{w_1W^{i_1}, \ldots, w_mW^{i_m}\}$ is a flag. By induction $\{w_1W^{i_1}, \ldots, w_{m-1}W^{i_{m-1}}\}$ is a simplex and therefore the intersection of its vertices is $w'W_{I\setminus\{i_1,\ldots,i_{m-1}}$ for some w'. Thus we may suppose that we have vertices W^{i_1}, …, $W^{i_{m-1}}$ whose intersection is $W_{I\setminus\{i_1,\ldots,i_{m-1}}$ and a coset wW^{i_m} adjacent to each of $W^{i_1}, \ldots, W^{i_{m-1}}$. For each k, such that $1 \leq k < m$, $W^{i_k} \cap wW^{i_m} \neq \emptyset$ and so $w \in W^{i_k}W^{i_m}$. It follows from Corollary 9.35 that $w \in (\cap_{k=1}^{m-1} W^{i_k})W^{i_m}$. Thus $w = uv$, where $u \in \cap_{k=1}^{m-1} W^{i_k}$ and $v \in W^{i_m}$. Therefore $u \in W^{i_1} \cap \ldots \cap W^{i_{m-1}} \cap wW^{i_m}$ and it follows that the flag is a simplex. □

Theorem 9.37 implies that the rank 2 residues of type $\{i, j\}$ of $\Gamma(W, R)$ are isomorphic to the Coxeter geometry of $\langle r_i, r_j \rangle$. The group $\langle r_i, r_j \rangle$ is a dihedral group of order $2m_{ij}$ and its geometry is the geometry of an ordinary m_{ij}-gon in which we may take the points to be the cosets of $\langle r_i \rangle$ and the edges to be the cosets of $\langle r_j \rangle$. Thus the diagram of $\Gamma(W, R)$ is just the Coxeter-Dynkin diagram of (W, R).

In the case of the symmetric group S_n acting on $\Omega := \{1, 2, \ldots, n\}$ we have seen that the diagram is

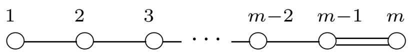


and we may identify the varieties of type i with the subsets of Ω of cardinality i. Two subsets are adjacent if one is properly contained in the other. If $W := S_n$, then W^i is the stabilizer of $\{1, 2, \ldots, i\}$.

It is easy to see that the geometry of subspaces of a projective space of dimension n has the above diagram. Conversely, it is a theorem of Tits (1981) that every geometry corresponding to this diagram is a projective geometry of dimension n.

If Δ is a building of type M, then we can construct a geometry of type M whose vertices are the vertices of Δ and whose type map is that of Δ. Two vertices v_1 and v_2 are adjacent if $\{v_1, v_2\}$ is a simplex of rank 2. We leave it as an exercise to check that the rank 2 residues of type $\{i, j\}$ are generalized m_{ij}-gons and that the flags of the geometry are the simplexes of Δ.

The diagram of the geometry of totally isotropic subspaces of a polar space of Witt index m is



The varieties of type i are the totally isotropic subspaces of dimension i. The residue corresponding to a totally isotropic subspace U is the geometry of $U \times U^\perp/U$.

Earlier in this chapter we showed that the Coxeter group of the polar building is the set of all permutations g of $\Omega := \{1, 1', 2, 2', \ldots, m, m'\}$ such that for all $i \in \Omega$, $g(i') = g(i)'$, where $i'' = i$. The varieties of type i are the subsets $A \subseteq \Omega$ of cardinality i with the property that $k \in A$ implies $k' \notin A$.

To conclude this section we describe a diagram geometry with diagram $\overset{1}{\circ}\!\!-\!\!\!-\!\!\!-\!\!\overset{2}{\circ}\!\!=\!\!=\!\!\overset{3}{\circ}$ but which is not a building. The varieties of type 1 (*points*) are the elements of a set Ω of cardinality 7; the varieties of type 2 (*Lines*) are the 35 3-element subsets of Ω; the varieties of type 3 (*Planes*) are 15 7-point planes forming an orbit for the alternating group A_7 as described in Chapter 6. Each point is adjacent to the 15 Lines which contain it, each Line is adjacent to the 3 Planes which contain it and every point is adjacent to every Plane. The residue of a point is the generalized quadrangle of totally isotropic subspaces of the symplectic geometry of dimension 4 over $\mathbb{F}_2$. The residue of a Plane is a 7-point plane. (See Exercises 6.8 and 8.12 and Neumaier (1984).)

Abstract Polar Spaces

Instead of considering the flags of totally isotropic subspaces of a polar geometry one may ask for an axiomatic characterization of the set of all totally isotropic subspaces themselves. This leads to the study of abstract polar spaces first considered by Veldkamp (1959) and later refined by Tits (1974). Tits defines an *abstract polar space* of rank m to be a set S together with a set of subsets (called *subspaces*) satisfying.

- P1. Any subspace, together with the subspaces it contains, is a projective space of projective dimension at most $m - 1$.
- P2. The intersection of any two subspaces is a subspace.
- P3. Given a subspace L of projective dimension $m - 1$ and a point $P \in S \setminus L$, there is a unique subspace M containing P such that $M \cap L$ has projective dimension $m - 2$ and contains every point of L collinear with P.
- P4. There exist two disjoint subspaces of projective dimension $m - 1$.

(The axioms for projective geometry can be found at the end of Chapter 3.)

It is not difficult to see that if V is a polar geometry of Witt index m, the set S of isotropic (or singular, in the case of an orthogonal geometry) points of $\mathcal{P}(V)$ together with the set of totally isotropic (resp. singular) subspaces of V, is an abstract polar space of rank m. In these examples two points are collinear if and only if they are orthogonal. Axioms P1 and P2 are immediate and the subspace M required in axiom P3 is $L \cap P^{\perp} + P$. Axiom P4 follows from Lemma 7.5.

The abstract polar spaces of rank at least 3 have been completely determined by Veldkamp (1959) and Tits (1974). As well as polar spaces arising from 'classical' polar geometries defined on vector spaces over fields, as indicated in the previous paragraph, there is a class of examples associated with the pseudo-quadratic forms of Tits (1968) and, when the rank is three, additional examples whose maximal subspaces are projective planes over non-commutative division rings or Cayley-Dickson division algebras.

An abstract polar space of rank two is a generalized quadrangle, as defined in Chapter 8. As well as the examples corresponding to the classical polar geometries of Witt index 2 there are many 'non-classical' examples. For a survey see Payne and Thas (1984).

A very beautiful characterization of polar spaces solely in terms of points and lines has been obtained by Buekenhout and Shult (1974). They considered a set S of *points* together with a non-empty set of subsets of S called *lines* such that the following axioms hold.

BS1. Every line has at least three points.

BS2. No point is collinear with all points of S.

BS3. If x is a point that does not belong to a line L, then either

a) exactly one point of L is collinear with x, or

b) every point of L is collinear with x.

It is a consequence of these axioms that two distinct points belong to at most one line. A subset X of S is called a (totally isotropic) *subspace* if every pair of distinct points is collinear and if whenever a line L contains two distinct points of X, every point of L belongs to X. If S does not contain infinite chains of distinct subspaces, it has been shown by Buekenhout and Shult (1974) that S together with its sets of totally isotropic subspaces satisfies the abstract polar space axioms. For additional information see Shult (1975) and Buekenhout (1979). An interesting application to finite permutation groups can be found in Kantor (1975). A far reaching generalization of this type of characterization has been obtained in the combined work of Cooperstein, Cohen, Shult and others. See Shult (1983) for a survey.

EXERCISES

9.1 Suppose that β is the polar form of a non-degenerate quadratic form $Q : V \to \mathbb{F}$ and suppose that the characteristic of $\mathbb{F}$ is 2. Assume that V contains singular vectors. The building Δ consists of the proper flags of *totally singular* subspaces and the apartments are the sets of flags obtained from the polar frames of *singular* points. Formulate and prove analogues of Theorems 7.15 and 7.17 for Δ.

9.2 Let V be a polar geometry of Witt index $m > 0$ and let G be the group of isometries of V. Excluding the case of an orthogonal geometry of dimension $2m$ and Witt index m, show that a subgroup of G is the stabilizer of a flag of totally isotropic subspaces if and only if it is a parabolic subgroup of G.

9.3 Let G be a group with a BN-pair and suppose that the Weyl group W has distinguished generators $\{ w_i \mid i \in I \}$.

(*i*) For $w \in W$ let $\ell'(w)$ be the least non-negative integer k such that $BwB \subseteq (Bw_{i_1}B)\cdots(Bw_{i_k}B)$, where $i_1, \dots, i_k \in I$. Show that $\ell'(w) = \ell(w)$.

(*ii*) If $k = \ell(w)$ and $BwB \subseteq (Bw_{i_1}B)\cdots(Bw_{i_k}B)$, show that

$$BwB = (Bw_{i_1}B)\cdots(Bw_{i_k}B) = Bw_{i_1}\cdots w_{i_k}B.$$

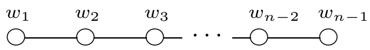
(*iii*) If $J, K \subseteq I$ and $gP_Jg^{-1} \subseteq P_K$, show that $J \subseteq K$ and $g \in P_K$.

(*iv*) Show that $P_{J\cap K} = P_J \cap P_K$ and $P_{J\cup K} = \langle P_J, P_K \rangle$.

(*v*) Show that the map $P_JwP_K \mapsto W_JwW_K$, $(w \in W)$ is a bijection.

9.4 Let G be the group of isometries of a polar geometry of Witt index $m \geq 1$, excluding the case of an orthogonal geometry of dimension $2m$. let M be a maximal flag of totally isotropic subspaces and let B be its stabilizer in G. If F is a flag and $F \subseteq M$, show that the stabilizer of F in G is a standard parabolic subgroup P_J for some J. Show that every standard parabolic subgroup is the stabilizer of a flag contained in M and deduce that there is a one-to-one correspondence between the flags of the polar building and the cosets of the standard parabolic subgroups.

9.5 Let W be the Coxeter group corresponding to the diagram



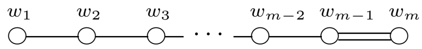
Consider the effect of each generator on the n cosets

$$H,\ Hw_{n-1},\ Hw_{n-1}w_{n-2},\ \dots,\ Hw_{n-1}w_{n-2}\cdots w_1$$

of the subgroup $H := \langle w_1, w_2, \dots, w_{n-2} \rangle$ and show by induction on n that W is isomorphic to the symmetric group S_n.

9.6 Let A be the group with generators x_1, x_2, $\dots$, x_{n-2} and relations $x_1^3 = 1$, $x_i^2 = 1$ $(i \neq 1)$, $(x_ix_{i+1})^3 = 1$ and $(x_ix_j)^2 = 1$ for $|i - j| \geq 2$. Show that A is isomorphic to the alternating group A_n.

9.7 Show directly that the Coxeter group of the diagram



is the wreath product $\mathbb{Z}_2 \wr S_m$.

9.8 Let W be the Coxeter group with generating set R and relations $(rs)^{m_{rs}} = 1$ for $r, s \in R$. Let V be the real vector space with basis $\{ e_r \mid r \in R \}$. Define a symmetric bilinear form β on V by

$$\beta(e_r, e_s) := -\cos(\pi/m_{rs}).$$

For $r \in R$, let $\sigma_r(v) = v - 2\beta(v, e_r)e_r$.

(i) Show that for r, $s \in R$ such that $r \neq s$, the matrix of the restriction of $(\sigma_r\sigma_s)^k$ to the subspace spanned by e_r and e_s is

$$\frac{1}{\sin(\pi/m_{rs})} \begin{pmatrix} \sin((2k+1)\pi/m_{rs}) & -\sin(2k\pi/m_{rs}) \\ \sin(2k\pi/m_{rs}) & -\sin((2k-1)\pi/m_{rs}) \end{pmatrix}.$$

(ii) Deduce from (i) that rs has order m_{rs} and that the map $r \mapsto \sigma_r$ extends to a homomorphism from W to the group of transformations of V preserving β.

9.9 Consider the symmetric group S_n as a Coxeter group with generators $w_i := (i, i+1)$ for $1 \leq i < n$. For $w \in S_n$ show that, in the notation of Theorem 9.28,

$$D(w) = \{ (i,j) \mid i < j \text{ and } w^{-1}(i) > w^{-1}(j) \}.$$

(Use (9.22) and induction on $\ell(w)$.)

9.10 The Weyl group of the group of isometries of a polar geometry is $W := \mathbb{Z}_2 \wr S_m$.

(i) Show that this group can be identified with the group of all permutations w of the set $\Omega := \{\pm 1, \pm 2, \ldots, \pm m\}$ such that, for all $i \in \Omega$, $w(-i) = -w(i)$.

(ii) The distinguished generators of $\mathbb{Z}_2 \wr S_m$ are the permutations $(1,2)(-1,-2), \ldots, (m-1,m)(-m+1,-m)$ and $(m,-m)$. Show that the set T defined in (9.20) has m^2 elements and consists of two conjugacy classes of W.

(*iii*) Let V be the Euclidean vector space with orthonormal basis $\{ e_i \mid 1 \le i \le m \}$ and inner product (u, v). Show that there is an action of W on V such that for all $w \in W$ and $1 \le i \le m$,

$$w.e_i = \begin{cases} e_{w(i)} & \text{if } w(i) > 0 \\ -e_{-w(i)} & \text{if } w(i) < 0. \end{cases}$$

Let $\Sigma^+ := \{ e_i \mid 1 \le i \le m \} \cup \{ e_i - e_j, e_i + e_j \mid 1 \le i < j \le m \}$ and put $\Sigma^- := \{ -v \mid v \in \Sigma^+ \}$. Show that for each $t \in T$ there is a unique vector $a_t \in \Sigma^+$ such that for all $v \in V$,

$$t.v = v - \frac{2(v, a_t)}{(a_t, a_t)} a_t.$$

(*iv*) Show that $t \in D(w)$ if and only if $w^{-1}(a_t) \in \Sigma^-$.

9.11 Let (W, R) be a Coxeter system. Using Lemma 9.34, or otherwise, show that for $J, K, L \subseteq R$

(*i*) $W_J \cap W_K W_L = (W_J \cap W_K)(W_J \cap W_L)$, and

(*ii*) $W_J(W_K \cap W_L) = W_J W_K \cap W_J W_L$.

9.12 Let (W, R) be a Coxeter system and suppose $J, K \subseteq R$.

(*i*) Show that each coset $W_J w$ contains a unique element d of shortest length and that for all $v \in W_J$, $\ell(vd) = \ell(v) + \ell(d)$.

(*ii*) Let D_J be the set of representatives of shortest length for the cosets $W_J w$. Using Theorem 9.28 show that

$$D_J = \{ d \in W \mid J \cap D(d) = \emptyset \}.$$

(*iii*) Let $D_{J,K} := D_j \cap D_K^{-1}$ and show that for all $w \in W$,

$$D_{J,K} \cap W_J w W_K = \{d\},$$

where d is the unique element of shortest length in $W_J w W_K$. Show that every element of $W_J w W_K$ can be written in the form udv, where $u \in W_J$, $v \in W_K$ and

$$\ell(udv) = \ell(u) + \ell(d) + \ell(v).$$

(*iv*) Show that for all $d \in W_J$, $W_J \cap (W_K d W_L) = W_{J \cap K} d W_{J \cap L}$.

9.13 In the notation of the previous exercise show that for $d \in D_{J,K}$

(i) $d^{-1}W_J d \cap W_K = W_{d^{-1}Jd\cap K}$,

(ii) $W_J \cap dW_K d^{-1} = W_{J\cap dKd^{-1}}$, and

(iii) If $u \in W_J \cap D^{-1}_{J\cap dKd^{-1}}$, $v \in W_{J\cap dKd^{-1}}$ and $t \in W_K \cap D_{d^{-1}Jd\cap K}$, then

$$\ell(uvdt) = \ell(u) + \ell(v) + \ell(d) + \ell(t).$$

Moreover, every element of $W_J dW_K$ can be written uniquely in the form $uvdt$.

9.14 (The Bruhat Order) Let (W, R) be a Coxeter system and let T be its set of reflections. For w, $w' \in W$, define $w \leq w'$ if $w = w'$ or if there is a sequence of elements t_1, t_2, … t_k of T such that $w' = t_k t_{k-1} \cdots t_1 w$ and

$$\ell(w) < \ell(t_1 w) < \ell(t_1 t_2 w) < \cdots < \ell(w').$$

(i) Observe that $t \in T \setminus D(w)$ implies $w < tw$ and prove that if $w \leq w'$ and $r \in R$, then $rw \leq w'$ or $rw \leq rw'$.

(ii) Deduce that if $r_1 r_2 \cdots r_\ell$ is a reduced expression for w', then $w \leq w'$ if and only if $w = r_{i_1} \cdots r_{i_k}$ for some sequence $1 \leq i_1 < \cdots < i_k \leq \ell$.

(iii) Show that $D(w_1) \subseteq D(w_2)$ if and only if $w_2 = w_1 v$ for some element v of length $\ell(w_2) - \ell(w_1)$. Deduce that $D(w_1) \subseteq D(w_2)$ implies $w_1 \leq w_2$.

9.15 Suppose that (W, R) is a Coxeter system and that $|W|$ is finite. Let w_R be an element of maximum length in W and let T be the set of reflections.

(i) Show that $D(w_R) = T$.

(ii) Show that for all $w \in W$, $D(ww_R) = T \setminus D(w)$. Deduce that w_R is the unique element of maximum length in W and that $w_R^2 = 1$.

(iii) If $w_R = w_1 w_2$, show that $\ell(w_R) = \ell(w_1) + \ell(w_2)$. Regarding the elements of W as the chambers of the Coxeter complex $\Sigma(W, R)$, deduce that every element of W is part of a gallery from 1 to w_R.

(iv) Show that $D(w_R w w_R) = w_R D(w) w_R$. In particular, $w_R R w_R = R$ and so conjugation by w_R induces an automorphism of the Coxeter-Dynkin diagram. Interpret this for the Coxeter groups S_n and $\mathbb{Z}_2 \wr S_m$ by describing w_R explicitly.

(v) For $J \subseteq R$, let w_J be the element of maximal length in W_J. Show that $wRw^{-1} \setminus D(w) = J$ if and only if $w = w_J w_R$.

9.16 Let Δ be a building whose apartments (isomorphic to the Coxeter complex $\Sigma(W, R)$) are finite. Chambers C and D are said to be *opposite* if $d(C, D) = |T|$, where T is the set of reflections in W.

(*i*) Show that every apartment contains a pair of opposite chambers and that for each pair of opposite chambers of Δ there is a unique apartment which contains them.

(*ii*) Let Δ be the building of totally isotropic flags of a polar geometry V. Show that the chambers

$$V_1 \subset V_2 \subset \ldots \subset V_m \quad \text{and} \quad V_1' \subset V_2' \subset \ldots \subset V_m'$$

are opposite if and only if $V_i^\perp \oplus V_i' = V$ for all i.

(*iii*) Describe the pairs of opposite chambers in the building of $SL(V)$.

9.17 Let Σ be an apartment of the building Δ and let C be a chamber of Σ. Let ρ_C^Σ be the retraction onto Σ centred at C. Show that for all chambers D,

$$d(C, \rho_C^\Sigma(D)) = d(C, D).$$

9.18 Let Δ be a building of type M. Show that the rank 2 residues of the associated diagram geometry are generalized m_{ij}-gons and that the flags of the geometry are the simplexes of Δ.

9.19 Let $S := A \times B$, where A and B are non-empty sets and declare the sets $\{a\} \times B$ for $a \in A$ and $A \times \{b\}$ for $b \in B$ to be the lines of S. Show that S is a generalized quadrangle.

9.20 Consider a generalized n-gon with a finite number of points. Assume that each point lies on at least three lines and that each line contains at least three points. Show that there are integers s and t such that every point is on $t + 1$ lines and every line contains $s + 1$ points.

10

Unitary Groups

Let π be a unitary polarity of the projective geometry $\mathcal{P}(V)$, where V is a vector space of dimension n over the field $\mathbb{F}$ and suppose that π is induced by a σ-hermitian form β, where σ is an automorphism of $\mathbb{F}$ of order 2. The elements $f \in GL(V)$ such that

$$\beta(f(u), f(v)) = \beta(u, v) \quad \text{for all } u,\ v \text{ in } V$$

form the *unitary group* $U(V)$. The *full* unitary group $\mathit{\Gamma} U(V)$ consists of those τ-semilinear transformations f that induce a collineation of $\mathcal{P}(V)$ which commutes with π. That is, for some $a \in \mathbb{F}$ such that $a = \sigma(a)$ we have

$$\beta(f(u), f(v)) = a\tau\beta(u, v) \quad \text{for all } u,\ v \text{ in } V.$$

The *general unitary group* is $GU(V) = \mathit{\Gamma} U(V) \cap GL(V)$.

In contrast to the symplectic groups, these groups depend on the form as well as on the dimension of V and therefore our notation is ambiguous. However, from now on we shall be paying increasing attention to groups over *finite* fields and in this case we shall see that, up to isomorphism, there is just one group $U(V)$ in each dimension.

Our goal is to study the normal subgroups of $U(V)$, determine their action on the projective space, and determine whether they can be generated by transvections. Unlike the symplectic groups of Chapter 8 and the linear groups of Chapter 4, a unitary group $U(V)$ need not contain transvections. We shall see that the existence of transvections in $U(V)$ is equivalent to the existence of isotropic points in $\mathcal{P}(V)$ and for most of the chapter we shall consider only those groups that contain transvections.

Matrices

Let $e_1, e_2, \ldots, e_n$ be a basis for V and let $J := (\beta(e_i, e_j))$ be the matrix of β. If A is the matrix of $f \in GL(V)$ with respect to this basis, then by (7.6), $f \in U(V)$ if and only if

$$A^t J\sigma(A) = J.$$

Taking determinants, we see that $\det A \det \sigma(A) = 1$. The form β is not alternating and therefore there exists $v \in V$ such that $\beta(v,v) \neq 0$. The linear transformation that sends v to av and acts as the identity on $\langle v \rangle^\perp$ has determinant a and belongs to $U(V)$ if and only if $a\sigma(a) = 1$. It follows that the map

$$\det : U(V) \to \{ a \in \mathbb{F}^\times \mid a\sigma(a) = 1 \}$$

is onto. The *special unitary group* is defined to be the kernel of this map, namely

$$SU(V) := \{ f \in U(V) \mid \det f = 1 \}.$$

The scalar transformation $a\mathbf{1}$ is in $U(V)$ if and only if $a\sigma(a) = 1$ Thus the group $PU(V)$ of collineations of $\mathcal{P}(V)$ induced by $U(V)$ is isomorphic to

$$U(V)/\{ a\mathbf{1} \mid a\sigma(a) = 1 \}.$$

Similarly, the group $PSU(V)$ of collineations of $\mathcal{P}(V)$ induced by $SU(V)$ is isomorphic to

$$SU(V)/\{ a\mathbf{1} \mid a\sigma(a) = 1 \text{ and } a^n = 1 \}.$$

The Field $\mathbb{F}$

From now on we shall write $\bar{a}$ instead of $\sigma(a)$ whenever convenient and we let $\mathbb{F}_0$ denote the fixed field of σ, i.e.,

$$\mathbb{F}_0 := \{ a \in \mathbb{F} \mid a = \bar{a} \}.$$

Thus $\mathbb{F}_0$ is a subfield of $\mathbb{F}$ and $\dim_{\mathbb{F}_0} \mathbb{F} = 2$. The functions

$$Tr : \mathbb{F} \to \mathbb{F}_0 : a \mapsto a + \bar{a}$$

and

$$N : \mathbb{F}^\times \to \mathbb{F}_0^\times : a \mapsto a\bar{a}$$

are called the *trace* and *norm* respectively.

10.1 Lemma.

(i) $Tr : \mathbb{F} \to \mathbb{F}_0$ *is an* $\mathbb{F}_0$*-linear map onto* $\mathbb{F}_0$.

(ii) $Tr(a) = 0$ *if and only if* $a = b - \bar{b}$ *for some* $b \in \mathbb{F}$.

(iii) $N : \mathbb{F}^\times \to \mathbb{F}_0^\times$ *is a homomorphism.*

(iv) $N(a) = 1$ *if and only if* $a = b/\bar{b}$ *for some* $b \in \mathbb{F}^\times$.

(v) *If* $\mathbb{F}$ *is finite, then* N *is onto.*

Proof. (i) Certainly Tr is $\mathbb{F}_0$-linear and as Tr does not map every element of $\mathbb{F}$ to 0, it must map $\mathbb{F}$ onto $\mathbb{F}_0$. Thus $\mathbb{F}_0 = \ker(\mathbf{1} - \sigma) = \operatorname{im}(\mathbf{1} + \sigma)$.

(ii) We have $\operatorname{im}(\mathbf{1}-\sigma) \subseteq \ker(\mathbf{1}+\sigma)$, whence equality holds; both spaces being of $\mathbb{F}_0$-dimension 1.

(iii) This is clear.

(iv) If $a \neq -1$ and $N(a) = 1$, put $b = 1 + a$. If $a = -1$, choose b such that $Tr(b) = 0$.

(v) If $|\mathbb{F}_0| = q$, then $|\mathbb{F}| = q^2$ and $\sigma(a) = a^q$ for all $a \in \mathbb{F}$. Then

$$\ker N = \{\, a \in \mathbb{F}^\times \mid a^{q+1} = 1 \,\}.$$

The multiplicative group of a finite field is cyclic and therefore $\ker N$ is a cyclic group of order $q + 1$. It follows that $\operatorname{im} N = \mathbb{F}_0^\times$. □

Hyperbolic Pairs

10.2 Lemma. *If* $\dim V \geq 2$ *and the norm map is onto, then* V *contains isotropic vectors.*

Proof. Suppose that $v \neq 0$ is not isotropic and let $b = \beta(v, v)$. Then $b = \bar{b}$ and for $a \in \mathbb{F}$, $u \in \langle v \rangle^\perp$ we have

$$\beta(u + av, u + av) = \beta(u, u) + a\bar{a}b.$$

Since $-b^{-1}\beta(u, u) \in \mathbb{F}_0$ and the norm map is onto, either u is isotropic or there exists $a \in \mathbb{F}^\times$ such that $u + av$ is isotropic. □

10.3 Corollary. *If* $\dim V \geq 2$ *and* $\mathbb{F}$ *is finite, then* V *contains isotropic vectors.* □

If V contains an isotropic vector u, then by Lemma 7.3 V contains a hyperbolic pair (u, v). The isotropic points of $\mathcal{P}(\langle u, v \rangle)$ are $\langle v \rangle$ and $\langle u + bv \rangle$, where $b + \bar{b} = 0$.

If L_1 is a hyperbolic line in V, then $V = L_1 \perp L_1^\perp$ and if $L_1^\perp$ also contains an isotropic vector, we can continue this process, splitting off hyperbolic lines, until we find that

$$V = L_1 \perp L_2 \perp \ldots \perp L_m \perp W,$$

where m is the Witt index of V and W does not contain any isotropic vectors. By Witt's Theorem, V is determined up to isomorphism by m and W.

Let (e_i, f_i) be a hyperbolic pair and a basis for L_i. If $\mathbb{F}$ is finite, then $\dim W = 0$ or 1. When $\dim W = 1$, Lemma 10.1 *(iv)* shows that we may suppose that $W = \langle w \rangle$, where $\beta(w,w) = 1$. Then V has a basis e_1, f_1 $\ldots$, e_m, f_m or e_1, f_1 $\ldots$, e_m, f_m, w and the group $U(V)$ acts regularly on the bases of this type.

Instead of working with hyperbolic pairs we can consider bases of non-isotropic vectors. Given a non-isotropic vector $u_1 \neq 0$, we have $V = \langle u_1 \rangle \perp \langle u_1 \rangle^{\perp}$ and by induction we can write

$$V = \langle u_1 \rangle \perp \langle u_2 \rangle \perp \ldots \perp \langle u_n \rangle,$$

where the u_i are non-isotropic. A set of n mutually orthogonal non-isotropic points is called an *orthogonal frame* of $\mathcal{P}(V)$.

Order Formulae

If $\mathbb{F}_0$ is the finite field $\mathbb{F}_q$, we write $U(n,q)$, $SU(n,q)$, etc. instead of $U(V)$, $SU(V)$, etc. The first step in computing the orders of these finite groups is to determine the number of isotropic vectors and hyperbolic pairs in V.

10.4 Lemma. *V contains $(q^{n-1} - (-1)^{n-1})(q^n - (-1)^n)$ isotropic vectors.*

Proof. Let ι_n be the number of isotropic points in $\mathcal{P}(V)$. Then the number of isotropic vectors in V is $(q^2 - 1)\iota_n$. We shall establish a recurrence relation for ι_n. First observe that $\iota_1 = 0$ and, from the description of the isotropic points of a hyperbolic line given above, $\iota_2 = q + 1$.

Suppose that $n > 2$ and that P and Q are isotropic points such that $P+Q$ is non-degenerate. For each isotropic point $R \in (P+Q)^{\perp}$ there are q^2 isotropic points on $P+R$ other than P. As every line of $P^{\perp}$ through P meets $(P+Q)^{\perp}$, it follows that there are $1 + q^2\iota_{n-2}$ isotropic points in $P^{\perp}$.

If $R \notin P^{\perp}$, then $P + R$ is a hyperbolic line which contains $q + 1$ isotropic points. There are q^{2n-4} lines through P not in $P^{\perp}$ and hence q^{2n-3} isotropic points not in $P^{\perp}$. It follows that

$$\iota_n = q^2\iota_{n-2} + q^{2n-3} + 1 \tag{10.5}$$

It is easy to check that the solution to this recurrence relation is

$$\iota_n = (q^{n-1} - (-1)^{n-1})(q^n - (-1)^n)/(q^2 - 1). \qquad \square$$

10.6 Corollary. *The number of hyperbolic pairs in the vector space V is $q^{2n-3}(q^{n-1} - (-1)^{n-1})(q^n - (-1)^n)$.*

Proof. If u is an isotropic vector, it was shown above that the number of isotropic points not in $\langle u \rangle^{\perp}$ is q^{2n-3}. for each isotropic point $Q \notin \langle u \rangle^{\perp}$

there is exactly one choice of $v \in Q$ such that $\beta(u, v) = 1$. Thus the number of hyperbolic pairs in V is $(q^2 - 1)\iota_n q^{2n-3}$. □

It is now an easy matter to calculate the order of $U(n, q)$. We have shown that V has a basis of the form e_1, f_1, …, e_m, f_m, or e_1, f_1, …, e_m, f_m, w, where the (e_i, f_i) are mutually orthogonal hyperbolic pairs and $\beta(w, w) = 1$. The number of ways of choosing e_1, f_1, …, e_m, f_m, is

$$q^{2n-3}(q^2 - 1)\iota_n q^{2n-7}(q^2 - 1)\iota_{n-2} \cdots$$

and when n is odd, the number of choices for w is $q + 1$. The group $U(n, q)$ acts regularly on these bases and therefore

$$|U(n, q)| = q^{\frac{1}{2}n(n-1)} \prod_{i=1}^{n} (q^i - (-1)^i).$$

The special unitary group has index $q+1$ in $U(n, q)$ and the scalar matrices in $U(n, q)$ form a subgroup of order $q + 1$. Therefore

$$|PU(n, q)| = |SU(n, q)| = q^{\frac{1}{2}n(n-1)} \prod_{i=2}^{n} (q^i - (-1)^i).$$

The scalar matrices in $SU(n, q)$ form a subgroup of order d, where d is the greatest common divisor of n and $q + 1$. Therefore

$$|PSU(n, q)| = d^{-1} q^{\frac{1}{2}n(n-1)} \prod_{i=2}^{n} (q^i - (-1)^i).$$

Unitary Transvections

10.7 Lemma. *If* $f \in U(V)$, *then* $\ker(\mathbf{1} - f)^\perp = \operatorname{im}(\mathbf{1} - f)$.

Proof. Suppose that $v \in \ker(\mathbf{1} - f)$. Then $f(v) = v$ and so for all $u \in V$, $\beta(u - f(u), v) = \beta(u, v) - \beta(f(u), v) = 0$. Thus $\operatorname{im}(\mathbf{1} - f) \subseteq \ker(1 - f)^\perp$. On comparing dimensions we see that equality holds. □

If $\varphi \in V^*$, $u \in V$ and $\varphi(u) = 0$, the transvection $t_{\varphi,u}$ is the linear transformation

$$t_{\varphi,u}(v) = v + \varphi(v)u.$$

If $t_{\varphi,u} \neq \mathbf{1}$, then $\ker(\mathbf{1} - t_{\varphi,u}) = \ker\varphi$ and $\operatorname{im}(\mathbf{1} - t_{\varphi,u}) = \langle u \rangle$. If $t_{\varphi,u} \in U(V)$, it follows from the lemma that $\ker\varphi = \langle u \rangle^\perp$ and hence u is isotropic. Furthermore $t_{\varphi,u}$ preserves β and therefore, for $v, w \in V$,

$$\varphi(v)\beta(u,w) + \overline{\varphi(w)}\beta(v,u) = 0.$$

(Compare this with similar calculations for symplectic transvections.) It follows that the unitary transvections have the form

$$t(v) = v + a\beta(v,u)u, \tag{10.8}$$

where u is isotropic and $a \in \mathbb{F}$ satisfies $a + \bar{a} = 0$. Conversely, every transvection of this form is in $U(V)$. In particular, V contains isotropic vectors if and only if $SU(V)$ contains transvections.

In general, if V is a unitary geometry over an infinite field there is no guarantee that V will contain any isotropic vectors. For example, let V be a vector space over the field $\mathbb{C}$ of complex numbers and take $\bar{a}$ to be the usual complex conjugate of a. Suppose that $e_1, e_2, \ldots, e_n$ is a basis for V and let β be the hermitian form such that $\beta(e_i, e_j) = \delta_{ij}$. It is easy to see that V does not have any isotropic vectors.

The *root group* $X_{P,P^\perp}$ consists of those transvections of the form (10.8) for which $P = \langle u \rangle$. By Witt's Theorem, $U(V)$ is transitive on isotropic points and therefore $U(V)$ has a single conjugacy class of root groups of the form $X_{P,P^\perp}$. In what follows we use $T(V)$ to denote the subgroup of $SU(V)$ generated by the transvections.

Hyperbolic Lines

10.9 Theorem. *If L is a hyperbolic line, then $SU(L) = SL(2, \mathbb{F}_0)$.*

Proof. We may suppose that $L = \langle e, f \rangle$, where (e, f) is a hyperbolic pair. Then the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ represents an element of $SU(L)$ if and only if $ad - bc = 1$, $a\bar{c} + \bar{a}c = b\bar{d} + \bar{b}d = 0$ and $a\bar{d} + \bar{b}c = 1$. This is the case if and only if $a, d \in \mathbb{F}_0$, $b + \bar{b} = c + \bar{c} = 0$ and $ad - bc = 1$. Choose $s \in \mathbb{F}^\times$ so that $s + \bar{s} = 0$ and define $A' = \begin{pmatrix} a & sb \\ s^{-1}c & d \end{pmatrix}$. Then the mapping $A \mapsto A'$ is an isomorphism of $SU(L)$ onto $SL(2, \mathbb{F}_0)$. □

Notice that $e + af$ is isotropic if and only if $a + \bar{a} = 0$ and that the action of $PSU(L)$ on the isotropic points of $\mathcal{P}(L)$ corresponds to the action of $PSL(2, \mathbb{F}_0)$ on the projective line over $\mathbb{F}_0$. Moreover, the transvections of $SU(L)$ correspond to the transvections of $SL(2, \mathbb{F}_0)$.

10.10 Corollary.

(i) *$T(L) = SU(L)$; i.e., $SU(L)$ is generated by transvections.*

(ii) *$T(L) \subseteq SU(L)'$ except when $\mathbb{F}$ is $\mathbb{F}_4$ or $\mathbb{F}_9$.*

Proof. By the preceding remarks, (i) follows from Theorem 4.3 (ii) and (ii) follows from Theorem 4.4. □

10.11 Lemma. *For all $a \in \mathbb{F}_0 \setminus \{0\}$, the action of $SU(L)$ on the set $\{\, v \mid \beta(v,v) = a \,\}$ is regular.*

Proof. Suppose that $\beta(v,v) \neq 0$. Then $f \in SU(L)$ fixes v if and only if $f = \mathbf{1}$. On the other hand, by Witt's Theorem, $U(L)$ is transitive on the set $\{\, v \mid \beta(v,v) = a \,\}$. If $g \in U(L)$ sends w to v and $d = \det(g)$, let $\langle\, w \,\rangle^\perp = \langle\, u \,\rangle$ and then define f by $f(u) = d^{-1}u$, $f(v) = v$. It follows that $fg \in SU(L)$ and fg takes w to v. □

Our description of $SU(L)$ shows that even though it is doubly transitive on the isotropic points (see Theorem 4.1), it cannot be transitive on the isotropic vectors.

The Action of $PSU(V)$ on Isotropic Points.

Throughout this section we assume that the Witt index of V is at least 1; i.e., V contains isotropic vectors.

10.12 Theorem.

(i) *If the Witt index of V is 1, then $PSU(V)$ is a faithful doubly transitive group on the isotropic points of $\mathcal{P}(V)$.*

(ii) *If the Witt index of V is at least 2, then $PSU(V)$ is a primitive rank 3 group on the isotropic points of $\mathcal{P}(V)$.*

Proof. We shall use the fact that if L is a hyperbolic line, then $PSU(L)$ is doubly transitive on the isotropic points. Also, if E is a totally isotropic subspace of V, then every element of $GL(E)$ extends to an element of $SU(V)$ (Exercise 10.9). If an element of $PSU(V)$ fixes every isotropic point, it fixes every hyperbolic line pointwise and hence fixes every point of $\mathcal{P}(V)$. Thus $PSU(V)$ acts faithfully and transitively on the isotropic points.

Suppose that P is isotropic and that Q_1 and Q_2 are isotropic points not in $P^\perp$. By Witt's Theorem there exists $g \in U(V)$ such that $g(P) = P$ and $g(Q_2) = Q_1$. Let $d = \det(g)$. Then $d\bar{d} = 1$ and by Lemma 10.1 (iv), we have $d = a/\bar{a}$ for some a. We may suppose that $P = \langle\, e \,\rangle$ and $Q_1 = \langle\, f \,\rangle$, where $\beta(e,f) = 1$. Define $g' \in U(P + Q_1)$ by $g'(e) = \bar{a}e$ and $g'(f) = a^{-1}f$. Extend g' to V by defining it to be the identity on $(P + Q_1)^\perp$. Then $g'g \in SU(V)$

and $g'g$ fixes P and takes Q_2 to Q_1. It follows that $PSU(V)_P$ is transitive on the isotropic points not in $P^\perp$. In particular, if the Witt index of V is 1, $PSU(V)$ is doubly transitive.

Next suppose that the Witt index is at least 2 and that Q_1 and Q_2 are isotropic points of $P^\perp$. If $E := \langle P, Q_1, Q_2 \rangle$ is totally isotropic, then there is an element $g \in SL(E)$ which fixes P and takes Q_2 to Q_1. As noted above, g extends to an element of $SU(V)$. If E is not totally isotropic, then $L := Q_1 + Q_2$ is hyperbolic and there is an element $g \in SU(L)$ that takes Q_2 to Q_1. But then g extends to an element of $SU(V)$ that fixes P. It follows that $PSU(V)_P$ is transitive on the isotropic points of $P^\perp \setminus P$. This proves that $PSU(V)$ is a permutation group of rank 3. The proof of Theorem 8.3 carries over without change to show that $PSU(V)$ is primitive. □

Three-Dimensional Unitary Geometries

Throughout this section suppose that V is a unitary geometry of dimension 3 and Witt index 1. Let Ω be the set of isotropic points of $\mathcal{P}(V)$. By Theorem 10.12, $PSU(V)$ is doubly transitive on Ω.

Let (e, f) be a hyperbolic pair in V and let w be a non-zero element of $\langle e, f \rangle^\perp$. Replace the form β by $\beta(w, w)^{-1}\beta$. This does not change the group $U(V)$ but it does allow us to suppose that $\beta(w, w) = 1$. (If the field is finite, or (more generally) if the norm map is onto, we may choose w so that $\beta(w, w) = 1$ without changing the form.)

The vectors e, w, f form a basis for V such that the matrix of the form β is

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

If $g \in U(V)$ fixes $\langle e \rangle$, then g fixes $\langle e \rangle^\perp = \langle e, w \rangle$ and g is represented by an upper triangular matrix. If g fixes $\langle f \rangle$ as well, it must be diagonal and of the form

$$H(k, l) := \begin{pmatrix} k & 0 & 0 \\ 0 & l & 0 \\ 0 & 0 & \bar{k}^{-1} \end{pmatrix}$$

where $l\bar{l} = 1$. A straightforward calculation shows that every element of $U(V)_{\langle e \rangle}$ can be written as a product of a diagonal matrix $H(k, l)$ and a matrix of the form

$$Q(a, b) := \begin{pmatrix} 1 & -\bar{a} & b \\ 0 & 1 & a \\ 0 & 0 & 1 \end{pmatrix}$$

where $a\bar{a} + b + \bar{b} = 0$. The isotropic points of $\mathcal{P}(V)$, apart from $\langle e \rangle$, are of the form $\langle be + aw + f \rangle$, where $a\bar{a} + b + \bar{b} = 0$ and therefore the group

$$Q := \{ Q(a,b) \mid a\bar{a} + b + \bar{b} = 0 \}$$

acts regularly on them. The group $H := \{ H(k,l) \mid l\bar{l} = 1 \}$ is a complement to Q in $U(V)_{\langle e \rangle}$. If $\mathbb{F} = \mathbb{F}_{q^2}$, then Q is a Sylow subgroup of $U(3,q)$ (of order q^3).

10.13 Lemma. *Q' is the group $X_{\langle e \rangle, \langle e \rangle^\perp}$ and we have $T(V) \subseteq SU(V)'$.*

Proof. Multiplication in Q is given by

$$Q(a_1,b_1)Q(a_2,b_2) = Q(a_1 + a_2, b_1 + b_2 - \bar{a}_1 a_2)$$

and the commutator of $Q(a_1,b_1)$ and $Q(a_2,b_2)$ is $Q(0, a_1\bar{a}_2 - \bar{a}_1 a_2)$. It follows from Lemma 10.1 (ii) that

$$Q' = \{ Q(0,b) \mid b + \bar{b} = 0 \}.$$

This is the root group $X_{\langle e \rangle, \langle e \rangle^\perp}$ and it follows that all transvections belong to $SU(V)'$. □

10.14 Lemma. *If $\mathbb{F}$ is finite, $T(V)$ is transitive on $\{ v \mid \beta(v,v) = 1 \}$ except when $\mathbb{F} = \mathbb{F}_4$.*

Proof. Suppose that $\beta(v,v) = \beta(w,w) = 1$. If $\langle v, w \rangle$ is non-degenerate, then by Corollary 10.3, $\langle v, w \rangle$ is a hyperbolic line and by Lemma 10.11 there is an element of $T(V)$ that takes v to w. So suppose that $\langle v, w \rangle^\perp = \langle e \rangle \subset \langle v, w \rangle$, where e is chosen so that $v = e + aw$. Choose $f \in \langle w \rangle^\perp$ so that $\beta(e,f) = 1$. Then e, w, f is a basis for V and $\langle v \rangle^\perp = \langle e, w - \bar{a}f \rangle$.

Now choose $u \in \langle v \rangle^\perp$ so that $\langle u, v \rangle$ and $\langle u, w \rangle$ are non-degenerate. This amounts to putting $u = be + w - \bar{a}f$, where b is chosen so that $\beta(u,u) \neq 0$ and $\beta(be - \bar{a}f, be - \bar{a}f) \neq 0$. That is, we require $ab + \bar{a}\bar{b} \notin \{0, 1\}$. This can always be achieved provided $|\mathbb{F}| > 4$. It is now the case that the subspaces $\langle u, v \rangle$ and $\langle u, w \rangle$ are hyperbolic lines and so by Lemma 10.11 there is an element of $T(V)$ that takes v to w. □

10.15 Theorem. *The groups $PSU(3,q)$ are simple except for $PSU(3,2)$.*

Proof. Suppose that $q \neq 2$ and let v be a vector such that $\beta(v,v) = 1$. We have $SU(3,q)_v = SU(2,q)$ and by Corollary 10.10(i) this group is generated by transvections. But we have just shown that when $q \neq 2$, $T(V)$ is transitive on vectors v such that $\beta(v,v) = 1$. Hence $T(V) = SU(3,q)$. It now follows from Lemma 10.13 that $SU(3,q)' = SU(3,q)$. By Theorem 10.12 (i),

$SU(3,q)$ is primitive in its action on the isotropic points. If P is an isotropic point, the root group $X_{P,P^\perp}$ is an abelian normal subgroup of $SU(3,q)_P$ and so by Iwasawa's criterion, $PSU(3,q)$ is simple. □

In the next section we study the exceptional case $PSU(3,2)$. To prepare for this and for its own interest we now introduce a new combinatorial structure.

Block Designs

By definition a *block design* consists of a set X of v points and a set $\mathcal{B}$ of k-element subsets of X (called *blocks*) such that every 2-element subset of X belongs to exactly λ blocks. We also refer to X as a 2-(v,k,λ) design.

If $\mathbb{F} = \mathbb{F}_{q^2}$, every pair of distinct isotropic points belongs to a unique hyperbolic line and every hyperbolic line contains $q+1$ isotropic points. Thus the set Ω of isotropic points together with the intersections of Ω with the hyperbolic lines of $\mathcal{P}(V)$ is a 2-$(q^3+1, q+1, 1)$ design. The automorphism group of this design is $P\Gamma U(3,q)$. (See O'Nan (1973) and Taylor (1974).)

Each point is in q^2 blocks and there are $q^2(q^2-q+1)$ blocks. In $U(3,q)$ the stabilizer of a block is isomorphic to $U(2,q)$ and in $SU(3,q)$ it is $SU(2,q)$.

The Group $PSU(3,2)$

One of the goals of this chapter is to show that if the Witt index is at least 1, then except for $PSU(3,2)$, the groups $PSU(V)$ are generated by transvections. In this section we deal with the exceptional case. For reference we list the order of $PSU(3,2)$ and related groups.

group	order
$P\Gamma U(3,2)$	$2^4.3^3$
$PU(3,2)$	$2^3.3^3$
$PSU(3,2)$	$2^3.3^2$

group	order
$\Gamma U(3,2)$	$2^4.3^4$
$U(3,2)$	$2^3.3^4$
$SU(3,2)$	$2^3.3^3$

In the case we are now considering, the block design described at the end of the last section has 9 points, 12 blocks, and each block has 3 points.

10.16 Lemma. *Every 2-$(9,3,1)$ design can be identified with the points and lines of the affine plane over $\mathbb{F}_3$.*

Proof. Choose a point and label it 0. For any other point P, let $-P$ be the third point of the block containing 0 and P. If 0, P and Q are distinct points, define $P+Q$ to be the third point of the block containing $-P$ and

$-Q$. We leave it as an exercise to check that this addition is associative. By construction, the set of points becomes a vector space of dimension 2 over $\mathbb{F}_3$. The blocks correspond to the affine lines. □

Let E_9 denote the additive group of the vector space of dimension 2 over $\mathbb{F}_3$. In other words, $E_9 = \mathbb{F}_3 \oplus \mathbb{F}_3$. The affine group of E_9 is the semidirect product $E_9.GL(2,3)$—of order $2^4.3^3$. On comparing orders we find that

$$\begin{aligned} P\varGamma U(3,2) &\simeq E_9.GL(2,3), \\ PU(3,2) &\simeq E_9.SL(2,3), \quad \text{and} \\ PSU(3,2) &\simeq E_9.Q_8, \end{aligned}$$

where Q_8 is the quaternion group of order 8. If S is the Sylow 3-subgroup of $SU(3,2)$, then $SU(3,2) = S.Q_8$. We know that $T(V) \subseteq SU(3,2)'$ and therefore $T(V) \neq SU(3,2)$. Each transvection fixes a unique isotropic point and consequently $SU(3,2)$ contains 9 transvections. It follows that $T(V) = SU(3,2)' = S.\mathbb{Z}_2$.

Let us look at the geometry in somewhat more detail. The projective plane $\mathcal{P}(V)$ has 9 isotropic points and 12 non-isotropic points. Each hyperbolic line has 3 isotropic points and a pair of orthogonal non-isotropic points. Every other line is of the form $P^\perp$ for some isotropic point P. The 4 points of $P^\perp$ other than P are non-isotropic and no two are orthogonal.

If A is non-isotropic, then $A^\perp$ is a hyperbolic line which contains just two (orthogonal) non-isotropic points. Each non-isotropic point belongs to a unique orthogonal frame and so the 12 non-isotropic points form 4 disjoint orthogonal frames. Moreover, the transvection that fixes the isotropic point P fixes exactly one point in each of the 4 frames (namely, the points of $P^\perp \setminus P$) and interchanges the other two. It follows that the 4 orthogonal frames are the orbits of $T(V)$ on non-isotropic points. The 9 isotropic points form a single orbit for $T(V)$. Note that $\varGamma U(3,2)$ acts as S_4 on the orthogonal frames and that $T(V)$ is the kernel of this action; i.e., $\varGamma U(3,2)/T(V) \simeq S_4$.

If $v \neq 0$ is a non-isotropic vector, then $\beta(v,v) = 1$ and it follows that for every hyperbolic line L, $T(L)$ is transitive on the 6 non-zero non-isotropic vectors of L. Thus if $\{A, B, C\}$ is an orthogonal frame, then $T(V)$ is transitive on the 9 vectors representing A, B and C.

We conclude this section with a description of the Sylow 3-subgroup S of $SU(3,2)$. Suppose that $\mathbb{F}_4 = \mathbb{F}_2[\theta]$, where $\theta^2+\theta+1 = 0$. With respect to the basis e, w, f introduced at the beginning of this section it is easy to check that S contains

$$\begin{pmatrix} \theta^2 & 1 & 1 \\ \theta & 0 & 1 \\ \theta & \theta & \theta^2 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

These elements have order 3 and their commutator is θI. Thus $S' = Z(S) \simeq \mathbb{Z}_3$ and $S/Z(S) \simeq E_9$. Every element of S has order 3 (Exercise 10.8).

The Group $SU(4,2)$

In this section V denotes a four-dimensional unitary geometry over the field $\mathbb{F}_4 = \mathbb{F}_2[\theta]$. It follows from Lemma 10.4 that $\mathcal{P}(V)$ contains

45	isotropic points,
40	non-isotropic points,
27	totally isotropic lines, and
240	hyperbolic lines.

Each totally isotropic line contains 5 points. Each isotropic point is on 16 hyperbolic lines and on 3 totally isotropic lines.

The polar geometry of isotropic points and totally isotropic lines is a generalized quadrangle. (The definition is at the end of Chapter 8.) It is the purpose of this section to construct a generalized quadrangle from the 40 non-isotropic points and then identify this with the generalized quadrangle of the symplectic geometry of dimension 4 over $\mathbb{F}_3$.

From the order formula given earlier in this chapter we have $|U(4,2)| = 2^6.3^5.5$. The only scalar transformation in $SU(4,2)$ is the identity and therefore $PU(4,2) = SU(4,2) = PSU(4,2)$—a group of order $2^6.3^4.5$.

Let Ω be the set of non-isotropic points of $\mathcal{P}(V)$ and let $\mathcal{L}$ be the set of orthogonal frames.

10.17 Lemma. *The sets Ω and $\mathcal{L}$ are the points and lines of a generalized quadrangle. The group $PSU(4,2)$ acts as a primitive rank 3 group on both Ω and $\mathcal{L}$. The stabilizer of a point is isomorphic to $U(3,2)$ and the stabilizer of a line is isomorphic to the semidirect product $E_{27}.S_4$, where E_{27} is the group $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$.*

Proof. If P_1 and P_2 are orthogonal points of Ω, then $(P_1 + P_2)^\perp$ is a hyperbolic line. We know that every hyperbolic line contains two points of Ω and that these points are orthogonal. Thus each pair of distinct points is in at most one orthogonal frame.

If $P \in \Omega$, then $PSU(V)_P \simeq U(3,2)$ and (by Witt's Theorem) the orbits of $PSU(V)_P$ on Ω are

$$\{P\}, \quad \{\, Q \in \Omega \mid Q \in P^\perp \,\} \quad \text{and} \quad \{\, Q \in \Omega \mid Q \neq P,\ Q \notin P^\perp \,\}.$$

The lengths of these orbits are 1, 12 and 27 respectively. It follows that the action of $PSU(V)$ on Ω is primitive. (See Exercise 1.16.)

The number of orthogonal frames in $\mathcal{P}(V)$ is $40.12.2.1/4.3.2.1 = 40$ and each point of Ω belongs to 4 frames. If $\Delta = \{P_1, P_2, P_3, P_4\} \in \mathcal{L}$, there are precisely 12 elements of $\mathcal{L}$ with just one point in common with Δ. Three points of Ω are pairwise orthogonal if and only if they belong to an element of $\mathcal{L}$. Hence the 12 elements of $\mathcal{L}$ that have one point in common with Δ contain all the points of Ω. Therefore if $Q \in \Omega$ is not in Δ, there is a *unique* point of Δ that is in an orthogonal frame containing Q. This proves that Ω and $\mathcal{L}$ form the points and lines of a generalized quadrangle.

Suppose that $P_i = \langle e_i \rangle$ $(1 \leq i \leq 4)$. It is clear that $PSU(V)$ is transitive on $\mathcal{L}$ and that the stabilizer of Δ in $PSU(V)$ is $(\mathbb{F}_4^\times)^3.S_4$, where S_4 corresponds to the group of 4×4 permutation matrices with respect to the basis e_1, e_2, e_3, e_4 and the group $(\mathbb{F}_4^\times)^3$ consists of the diagonal matrices of determinant 1.

In the previous section we saw that for $P \in \Omega$, $PSU(V)_P \simeq U(3,2)$ acts as A_4 on the orthogonal frames containing P. Thus $PSU(V)_\Delta$ is transitive on the 12 frames that meet Δ in one point. The only diagonal matrix of determinant 1 that fixes an orthogonal frame disjoint from Δ is the identity. Thus the group $(\mathbb{F}_4^\times)^3$ of diagonal matrices in $PSU(V)_\Delta$ acts regularly on the frames disjoint from Δ. In particular, $PSU(V)$ is a rank 3 group on $\mathcal{L}$ and, by Exercise 1.16, it is primitive. □

Even though Ω and $\mathcal{L}$ both contain 40 points, the stabilizer in $PSU(4,2)$ of a point of Ω is not isomorphic to the stabilizer of a point of $\mathcal{L}$ and therefore the permutation representations of $PSU(4,2)$ on these sets are not equivalent.

10.18 Theorem. *The generalized quadrangle $(\Omega, \mathcal{L})$ is isomorphic to the generalized quadrangle of points and totally isotropic lines of a 4-dimensional symplectic geometry over $\mathbb{F}_3$.*

Proof. We make the set Ω into a projective geometry by defining 'lines' and 'planes' as follows. If P and Q are orthogonal points of Ω, the line PQ containing P and Q is defined to be the unique orthogonal frame containing P and Q. If P and Q are distinct but not orthogonal, the line PQ is defined to be $(P + Q) \cap \Omega$. In both cases PQ contains 4 points.

For each $R \in \Omega$, define $\pi(R) := (\{R\} \cup R^\perp) \cap \Omega$ and declare the sets $\pi(R)$ to be the 'planes' of Ω. If P and Q are distinct points of $\pi(R)$, then PQ is contained in $\pi(R)$. Thus the plane $\pi(R)$ has 13 points, 13 lines and each line has 4 points. The points not on a given line of $\pi(R)$ form a 2-$(9,3,1)$ design and by Lemma 10.13 this can be identified with the affine plane over $\mathbb{F}_3$. It follows that $\pi(R)$ can be identified with the projective plane over $\mathbb{F}_3$.

Any 3 non-collinear points of Ω are in at most one 'plane'. Counting 4-tuples (A, B, C, R) such that A, B, C are non-collinear points of $\pi(R)$ shows that every triple of non-collinear points is in exactly one 'plane'. Another simple counting argument shows that the lines of $\pi(R)$ together with the lines that meet $\pi(R)$ in one point account for all the lines of the geometry. That is, every line meets every plane. Similarly, every pair of distinct planes have a common line.

The method of Lemma 10.13 can now be used to show that the set $\Gamma = \Omega \setminus \pi(R)$ of the 27 points not on a given plane $\pi(R)$ form a three-dimensional geometry over $\mathbb{F}_3$. That is, choose a point $0 \notin \pi(R)$ and for every point $P \notin \pi(R) \cup \{0\}$ define $-P$ to be the third point of Γ on the line through 0 and P. If 0, P and Q are distinct points of Γ, define $P + Q$ to be the third point of Γ on the line through $-P$ and $-Q$. It follows from Lemma 10.13 that addition is associative and that Γ is a vector space over $\mathbb{F}_3$. The geometry of Ω is uniquely determined by Γ (the points of $\pi(R)$ correspond to the parallel classes of lines of Γ) and therefore it is the projective geometry of a vector space of dimension 4 over $\mathbb{F}_3$.

The correlation $R \mapsto \pi(R)$ is a symplectic polarity and the elements of $\mathcal{L}$ correspond to the isotropic lines. □

10.19 Corollary. *$PSU(4,2)$ is isomorphic to $PSp(4,3)$ and it is a simple group generated by unitary transvections.*

Proof. It follows from the proof of the theorem that every element of $P\Gamma U(4,2)$ induces a collineation of the projective geometry which commutes with the symplectic polarity. The groups $P\Gamma U(4,2)$ and $P\Gamma Sp(4,3)$ both have order 51840 and therefore they are isomorphic. It follows from Theorem 8.8 that the simple group $PSp(4,3)$ is the unique subgroup of index 2 in $P\Gamma Sp(4,3)$. Thus $PSU(4,2)$ is isomorphic to $PSp(4,3)$. □

This isomorphism provides an example of a group with two different BN-pair structures. In the BN-pair arising from the action on the unitary geometry, the group B is the normalizer of a Sylow 2-subgroup whereas in the BN-pair arising from the symplectic geometry the group B is the normalizer of a Sylow 3-subgroup.

The Simplicity of $PSU(V)$

10.20 Theorem. *For $n \geq 2$, the groups $PSU(n,q)$ are simple except for $PSU(2,2)$, $PSU(2,3)$ and $PSU(3,2)$.*

Proof. Theorems 10.9, 4.5 and 10.15 show that the result holds for $n = 2$ and $n = 3$. To prove the general case we first show that except for $PSU(3,2)$, the groups $PSU(n,q)$ are generated by transvections.

Suppose that v and w are vectors such that $\beta(v,v) = \beta(w,w) = 1$. If $n > 3$, choose a non-isotropic vector $u \in \langle v, w \rangle^\perp \setminus \langle v, w \rangle$. Then $\langle u, v \rangle$ and $\langle u, w \rangle$

are hyperbolic lines and by Lemma 10.11 there is a product of transvections that takes v to w. Thus the subgroup T generated by the transvections is transitive on the vectors v such that $\beta(v, v) = 1$. (When $n = 2$ or 3 this result was proved as Lemmas 10.11 and 10.14 respectively.)

Suppose that $n > 3$. Then we have $SU(n, q)_v \simeq SU(n-1, q)$ and so by induction or by Theorem 10.15 or by Corollary 10.19, $SU(n, q)_v \subseteq T$. It follows from the transitivity of T that $SU(n, q) = T$. In particular, if P is an isotropic point, then $SU(n, q)$ is generated by the conjugates of the root group $X_{P,P^\perp}$. Also, from Lemma 10.13 we have $SU(n, q)' = SU(n, q)$ and from Theorem 10.12 (*ii*), the action of $SU(n, q)$ on isotropic points is primitive. The root group $X_{P,P^\perp}$ is an abelian normal subgroup of $SU(n, q)_P$ and so by Iwasawa's criterion, $PSU(n, q)$ is simple. □

This proof also holds for infinite fields provided the Witt index is at least 1 and provided that we can show that the subgroup $T(V)$ generated by the transvections is transitive on vectors v such that $\beta(v, v) = a$ for some $a \in \mathbb{F}_0^\times$. In general, a more elaborate proof is needed.

In preparation for the general proof suppose from now on that V is a unitary geometry of dimension $n \geq 2$ over $\mathbb{F}$ and that V contains isotropic vectors.

It follows from Lemma 10.1 (*i*) that if L is a hyperbolic line, then for all $a \in \mathbb{F}_0$ there exists $v \in L$ such that $\beta(v, v) = a$.

10.21 Lemma. *Except when $n = 3$ and $\mathbb{F} = \mathbb{F}_4$, the subgroup $T(V)$ generated by the transvections is transitive on the set $\{ v \mid \beta(v, v) = e \}$, where $e \in \mathbb{F}_0^\times$.*

Proof. If $\dim V = 2$ or if the field is finite we have already proved this result. So suppose that $\dim V > 2$ and that $|\mathbb{F}| > 4$. Let x and x' be distinct vectors such that $\beta(x, x) = \beta(x', x') = e$.

Suppose at first that $x' - x$ is not isotropic. By Witt's Theorem there is a hyperbolic line L which contains $x' - x$ and we may write $x = y + z$ and $x' = y' + z'$, where $y, y' \in L$ and $z, z' \in L^\perp$. Then

$$z - z' = x - x' - y + y' \in L \cap L^\perp = \{0\}$$

and so $z = z'$ and $\beta(y, y) = \beta(y', y')$. If y is not isotropic, then we know that there is a product of transvections h in $SU(L)$ that takes y' to y. But then h extends to an element of $T(V)$ that takes x' to x. Thus we may suppose that y is isotropic for all choices of hyperbolic line L containing $x' - x$. We shall show that this leads to a contradiction.

Write $x = u + v$, where $u \in \langle x' - x \rangle$ and $v \in \langle x' - x \rangle^\perp$. Then $u \neq 0$, otherwise $x' - x$ would be isotropic, contrary to assumption. Now $\beta(u, y) =$

$\beta(u,x) = \beta(u,u)$ and therefore $\beta(u, y-u) = 0$. Put $w = y-u$. Then $\langle\, u\,\rangle \neq \langle\, w\,\rangle$ and $v = w+z$. Thus $\beta(v,w) = \beta(w,w) = -\beta(u,u)$.

If we can choose $\lambda, \mu \in \mathbb{F}$ so that $u+\lambda v+\mu w$ is isotropic, then the line $L_1 := \langle\, u, u+\lambda v+\mu w\,\rangle$ will be hyperbolic, because $\beta(u, u+\lambda v+\mu w) = \beta(u,u) \neq 0$. The vector $u+\lambda v+\mu w$ is isotropic if and only if

$$b = \lambda\bar{\lambda} + (\bar{\lambda}\mu + \lambda\bar{\mu})b + \mu\bar{\mu}b, \tag{10.22}$$

where $a = \beta(v,v)$ and $b = -\beta(u,u)$.

We can find values of λ and μ satisfying (10.22) as follows. Set $c = b^{-1}(b-a)$ and choose $\eta \in \mathbb{F}$ such that $Tr(\eta) = 1+c$ but $\eta \neq 0, c$. (This is possible because $\mathbb{F}_0 \neq \mathbb{F}_2$.) Next, since there are at most two elements of $\mathbb{F}$ with a given norm and trace, we can choose ξ so that $N(\xi) = N(\eta)$ but $Tr(\xi) \neq 1+c$. Then (10.22) is satisfied by $\lambda = (\eta-c)^{-1}$ and $\mu = (\xi-1)(\eta-c)^{-1}$ and in addition we have $\mu\bar{\mu} \neq 1$.

Now that we know that L_1 exists we can write $x = y_1+(x-y_1)$, where $y_1 \in L_1$ is isotropic and $x-y_1 \in L_1^{\perp}$. Since $y_1 \in L_1$, we have $\gamma y_1 = u + d(\lambda v+\mu w)$ for some d and some $\gamma \neq 0$. It follows that $d\bar{d} = 1$ and

$$0 = \beta(x-y_1, \gamma y_1) = \beta(u+v, u+d(\lambda v+\mu w)) = -b + d(\lambda a+\mu b).$$

Thus $b^2 = (\lambda a+\mu b)(\bar{\lambda}a+\bar{\mu}b)$ and from (10.22) we obtain $(\bar{\mu}\mu-1)(a-b) = 0$, whence $a = b$. But now we have $\beta(v,v) = -\beta(u,u)$ and hence $\beta(x,x) = 0$ —a contradiction.

From now on we may suppose that $x'-x$ is isotropic. Set $a = \beta(x,x')$ so that $a+\bar{a} = 2\beta(x,x)$. If $a \neq 0$ and λ is an element of $\mathbb{F}$ such that $\lambda\bar{\lambda} = 1$, then

$$\beta(x'-\lambda x, x'-\lambda x) = a+\bar{a}-\lambda a-\bar{\lambda}\bar{a}.$$

If $x'-\lambda x$ were isotropic for all λ then we would have $\lambda^2 a - (a+\bar{a})\lambda + \bar{a} = 0$ for all such λ. But this is not possible because $|\mathbb{F}_0| > 2$. Therefore $x'-\lambda x$ is not isotropic for some λ of norm 1. Since $\beta(\lambda x, \lambda x) = \beta(x,x)$, the first part of the proof shows that there exists $h_1 \in T(V)$ such that $h_1(x') = \lambda x$. But we also have $\beta(x-\lambda x, x-\lambda x) = (1-\lambda)(1-\bar{\lambda})\beta(x,x) \neq 0$ so the same argument shows that $h_2(x) = \lambda x$ for some $h_2 \in T(V)$. Thus $h_2^{-1}h_1(x') = x$ and this finishes the case $a \neq 0$.

Now suppose that $a = 0$. It follows that $\mathbb{F}$ has characteristic 2. We are assuming that $|\mathbb{F}_0| > 2$ and therefore there exists $\eta \in \mathbb{F}$ such that $Tr(\eta) = 0$ but $1+\eta \neq 0$. Let $\lambda = \eta(1+\eta)^{-1}$ and $\mu = (1+\eta)^{-1}$. Then $\lambda\bar{\lambda} + \mu\bar{\mu} = 1$ and if $y = \lambda x + \mu x'$, then $\beta(y,y) = \beta(x,x) = \beta(x',x')$. Since $\beta(y,x) \neq 0$ and $\beta(y,x') \neq 0$, there are elements $h_1, h_2 \in T(V)$ such that $h_1(x') = y$ and $h_2(y) = x$. Then $h_2h_1(x') = x$ and this completes the proof. □

10.23 Theorem. *If* $\dim V \geq 2$ *and if the Witt index of* V *is at least 1, then except for* $PSU(2,2)$, $PSU(2,3)$ *and* $PSU(3,2)$, *the group* $PSU(V)$ *is simple.*

Proof. Let x be a non-isotropic vector such that $\langle x \rangle^\perp$ contains an isotropic vector. By Lemma 10.21, $T(V)$ is transitive on $\{ v \mid \beta(v,v) = \beta(x,x) \}$, except when $n = 3$ and $\mathbb{F} = \mathbb{F}_4$. As in Theorem 10.20 we can show by induction that $SU(V) = T(V)$ and hence $SU(V) = SU(V)'$ except for $SU(2,2)$, $SU(2,3)$ and $SU(3,2)$. It follows from Iwasawa's criterion and Theorem 10.12 that $PSU(V)$ is simple except for $PSU(2,2)$, $PSU(2,3)$ and $PSU(3,2)$. □

An Example

So far we have left the unitary geometries of Witt index 0 out of our deliberations. These geometries generally behave quite differently to those of positive index and we shall give an example (due to J. Dieudonné) to illustrate this. Let $\mathbb{F}$ be the field of formal power series $\xi = \sum_{k=t}^{\infty} a_k x^k$, where $a_k \in \mathbb{C}$ for all k and $t \in \mathbb{Z}$. Define $\bar{\xi} = \sum_{k=t}^{\infty} \bar{a}_k x^k$. Let V be a vector space with basis $e_1, e_2, \ldots, e_n$ over $\mathbb{F}$ and define

$$\beta(\xi_1 e_1 + \cdots + \xi_n e_n, \eta_1 e_1 + \cdots + \eta_n e_n) = \xi_1 \bar{\eta}_1 + \cdots + \xi_n \bar{\eta}_n.$$

Then β is a non-degenerate hermitian form and V has no isotropic vectors. The *order* of $\xi = \sum_{k=t}^{\infty} a_k x^k$ is the least integer h such that $a_h \neq 0$. If $A = (\xi_{ij})$ is an element of $U(V)$, then $\sum_{i=1}^{n} \xi_{ij} \bar{\xi}_{ij} = 1$ for all j and it follows that the order of ξ_{ij} is ≥ 0 for all i and j. Let G_m be the set of unitary matrices of the form $I + x^m B$, where every entry of B has order ≥ 0. Then G_m is a normal subgroup of $U(V)$ and

$$\bigcap_{m \geq 0} G_m = \{I\}.$$

It turns out that the factor groups G_m / G_{m+1} are abelian.

Unitary BN-pairs

The construction of a BN-pair for a strongly transitive group acting on the building of a polar geometry has been given in Chapter 9. This construction applies to any unitary geometry V that contains isotropic vectors—the group N is the stabilizer of a polar frame

$$\mathcal{F} := \{ P_i, P_i^* \mid 1 \leq i \leq m \},$$

and the group B is the stabilizer of the chamber

$$M := \{\langle P_1, \dots, P_i \rangle \mid 1 \leq i \leq m\}.$$

It follows by induction from Theorem 10.12 that $SU(V)$ is strongly transitive and therefore, by Theorem 9.8, the stabilizers of M and $\mathcal{F}$ in $SU(V)$ form a BN-pair for $SU(V)$. This is also a consequence of the proof Theorem 8.9, which depends only on the existence of the polarity π and not on the precise nature of the underlying geometry.

It was shown on p. 85 that the Weyl group $N/B \cap N$ is isomorphic to $\mathbb{Z}_2 \wr S_m$ and generated by elements $n_i B \cap N$. In the case of a unitary geometry the elements n_i can be chosen in $SU(V)$. In fact, for $i < m$, the definition of $\hat{n}_i$ in the section 'Symplectic BN-pairs' of Chapter 8 carries over to the unitary case and produces an element of the required type. The remaining generator n_m is defined by putting $n_m(e_m) := \lambda f_m$ and $n_m(f_m) := -\dfrac{1}{\lambda} e_m$, where $\lambda + \bar{\lambda} = 0$, and requiring n_m to be the identity on $\langle e_m, f_m \rangle^\perp$.

EXERCISES

10.1 Show that $Sp(2m, q)$ is contained in $SU(2m, q)$.

10.2 If a unitary geometry has isotropic vectors, show that it has a basis of isotropic vectors.

10.3 Suppose that $\dim V \geq 2$ and that V contains isotropic vectors. Show that the centralizer of $SU(V)$ in $GL(V)$ is the group of scalar transformations.

10.4 If V is the unitary geometry of dimension n over $\mathbb{F}_{q^2}$, show that the number of totally isotropic subspaces of dimension k in V is

$$\prod_{i=n+1-2k}^{n} (q^i - (-1)^i) \Big/ \prod_{j=1}^{k} (q^{2j} - 1).$$

10.5 Find all the normal subgroups of $U(2, q)$.

10.6 Construct a three-dimensional unitary geometry of Witt index 1 that contains a two-dimensional subspace without any isotropic points.

10.7 If $n \geq 4$, show that $PSU(n, 2)$ is a rank 3 group on the non-isotropic points. Show that the stabilizer of a point has orbits of length 1, $\frac{1}{3} 2^{n-2}(2^{n-1} - (-1)^{n-1})$ and $2^{2n-3} - (-1)^n 2^{n-2} - 1$.

10.8 Let S be the Sylow 3-subgroup of $SU(3,2)$. Show that $x^3 = 1$ for all $x \in S$.

10.9 Let E be a totally isotropic subspace of the unitary geometry V. Show that every element of $SL(E)$ extends to an element of $SU(V)$. Then show that $SU(V)$ is transitive on the flags of totally isotropic subspaces of a given type.

10.10 (*i*) Let $G := \langle\, x, y \mid x^2 = y^5 = (xy)^3 = 1 \,\rangle$. Set $x_1 = x^{-1}y^{-2}xy^2$, $x_2 = y^2xy^{-2}$ and $x_3 = x$. Show that x_1, x_2 and x_3 satisfy the relations of Exercise 9.6. Deduce that $G \simeq A_5$.

(*ii*) Suppose that 5 divides $q - 1$ and let a be an element of order 5 in $\mathbb{F}_q$. Let $b = (a - a^{-1})^{-1}$ and set

$$A = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} b & c \\ d & -b \end{pmatrix},$$

where c and d are chosen so that $\det B = 1$. Show that the image of $\langle\, A, B \,\rangle$ in $PSL(2,q)$ is isomorphic to A_5.

(*iii*) Suppose that 5 divides $q + 1$ and let a be an element of order 5 in $\mathbb{F}_{q^2}$. Let $b = (a - \bar{a})^{-1}$ and choose c so that $b\bar{b} + c\bar{c} = 1$. If

$$A = \begin{pmatrix} a & 0 \\ 0 & \bar{a} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} b & c \\ -\bar{c} & \bar{b} \end{pmatrix},$$

show that $\langle\, A, B \,\rangle \subseteq SU(2,q)$ and that the image of $\langle\, A, B \,\rangle$ in $PSU(2,q)$ is isomorphic to A_5. (Take the matrix of the hermitian form to be the identity.)

(*iv*) Show that A_5 is contained in $PSL(2,q)$ if and only if 5 divides $|PSL(2,q)|$.

(*v*) Deduce from (*iv*) that $PSL(2,9) \simeq A_6$ and then show that $PGL(2,9)$ is *not* isomorphic to S_6.

10.11 (Cayley's parametrization) Suppose that A is an $n \times n$ matrix over $\mathbb{F}$ such that $I + A \in GL(n,\mathbb{F})$. Let $S = 2(I + A)^{-1} - I$ and then show that

$$S = (I - A)(I + A)^{-1} \qquad \text{and}$$
$$A = (I - S)(I + S)^{-1}.$$

If J is any $n \times n$ matrix and $a \mapsto \bar{a}$ is an automorphism of $\mathbb{F}$ of order 2, show that $A^tJ\bar{A} = J$ if and only if $S^tJ + J\bar{S} = 0$.

10.12 Let V be the unitary geometry over the field of formal power series with complex entries introduced at the end of this chapter. Let G_m be the subgroup consisting of the matrices of the form $I+x^mB$, where every element of B has order ≥ 0. If $A \in G_m$, show that $I+A$ is non-singular and $A = (I - x^mT)(I + x^mT)^{-1}$, where every element of T has order ≥ 0 and $\bar{T} + T^t = 0$. Show that G_m/G_{m+1} is isomorphic to the additive group of matrices T such that $\bar{T} + T^t = 0$.

10.13 Let V be a unitary geometry of dimension n over the field $\mathbb{F}_{q^2}$ defined by the non-degenerate hermitian form β and let W be a totally isotropic subspace of dimension k.

(i) Show that $\bar{\beta}(u+W, v+W) = \beta(u,v)$ defines a non-degenerate hermitian form on $W^\perp/W$.

(ii) Let U_W denote the subgroup of unitary transformations that fixes W and $W^\perp/W$ pointwise. If $e_1, e_2, \ldots, e_k$ is a basis for W, show that U_W acts regularly on the set Ω of k-tuples $(f_1, f_2, \ldots, f_k)$ such that $(e_1, f_1), \ldots, (e_k, f_k)$ are mutually orthogonal hyperbolic pairs. Deduce that U_W is a normal subgroup of $U(V)_W$ of order q^{2kn-3k^2} and that $U(V)_W$ is the semidirect product $U_W.U(V)_{W,W'}$, where $W' := \langle f_1, f_2, \ldots, f_k \rangle$ and $(f_1, f_2, \ldots, f_k) \in \Omega$.

(iii) Show that $U(V)_{W,W'} \simeq GL(k, q^2) \times U(n-2k, q)$.

(iv) For $i \neq j$, let $\hat{X}_{ij}$ be the group of transformations

$$v \mapsto v + a\beta(v, f_j)e_i - \bar{a}\beta(v, e_i)f_j.$$

Show that $\langle \hat{X}_{ij} \mid i \neq j \rangle \subset U(n,q)$ is isomorphic to $SL(W)$.

(v) Show that $U'_W = Z(U_W)$ is the subgroup of $U(V)$ that fixes $W^\perp$ pointwise and that its order is q^{k^2}.

(vi) Show that U_W/U'_W is isomorphic to the direct sum of $k(n-2k)$ copies of the additive group of $\mathbb{F}_{q^2}$.

10.14 Let β be a skew hermitian form on the vector space V of dimension m over $\mathbb{F}_{q^2}$. Show that $Tr\,\beta$ is an alternating form on V considered as a vector space of dimension $2m$ over $\mathbb{F}_q$. Deduce that $U(m,q) \subseteq Sp(2m,q)$.

10.15 Let V be a unitary geometry over $\mathbb{F}_{q^2}$. For each isotropic point $P = \langle u \rangle$, let $X_{P,P^\perp}$ be the root group of transvections

$$v \mapsto v + a\beta(v,u)u,$$

where $a + \bar{a} = 0$. Show that $\langle X_{P,P^\perp}, X_{Q,Q^\perp} \rangle$ is isomorphic to $\mathbb{F}_q$, $\mathbb{F}_q^+ \oplus \mathbb{F}_q^+$ or $SL(2,q)$ according to whether $P = Q$, or $P + Q$ is a totally isotropic or a hyperbolic line, respectively.

10.16 Let V be the *symplectic* geometry of dimension $2m$ over $\mathbb{F}_q$, where q is odd. Let g be an element of $Sp(V)$ such that $g^2 = -1$ and let C be the centralizer of g in $Sp(V)$.

(i) If 4 divides $q - 1$, show that the eigenspaces of g are maximal totally isotropic subspaces and that $C \simeq GL(m,q)$.

(ii) If 4 divides $q + 1$, show that it is possible to write V in the form $L_1 \perp L_2 \perp \ldots \perp L_m$, where each L_i is a hyperbolic line spanned by a hyperbolic pair (e_i, f_i) such that $g(e_i) = f_i$ and $g(f_i) = -e_i$. Regard V as a vector space of dimension m over $\mathbb{F}_{q^2} = \mathbb{F}_q[\omega]$, where $\omega^2 = -1$ and where multiplication by $a + b\,\omega$ is defined by

$$(a + b\,\omega)v := av + bg(v).$$

Show that the elements of C act as linear transformations on this space and that they preserve the hermitian form $\hat{\beta}$ defined by $\hat{\beta}(e_i, e_j) := \delta_{ij}$. Deduce that $C \simeq U(m,q)$.

(iii) If $g' \in Sp(V)$ and $g'^2 = -1$, show that g' is conjugate to g.

10.17 Let V be a unitary geometry defined by a non-degenerate hermitian form β and let π be the corresponding polarity. Then π induces an automorphism $\hat{\pi} : f \mapsto f^\perp$ of $SL(V)$, where $f^\perp$ is defined by (7.7).

(i) Show that every transvection $t \in SL(V)$ can be written in the form $t(x) = x + \beta(x,v)u$ for some u, $v \in V$ such that $\beta(u,v) = 0$ and show that $t^\perp = x - \beta(x,u)v$.

(*ii*) Let X be a root group of $SL(V)$ and let X_π be the subgroup of fixed elements of $\hat{\pi}$ in $\langle X, \hat{\pi}(X) \rangle$. Show that one of the following occurs.

(a) $X = \hat{\pi}(X)$ and X_π consists of unitary transvections.

(b) There is a totally isotropic line $L = \langle u, v \rangle$ such that X_π consists of the transformations

$$x \mapsto x + a\beta(x,v)u - \bar{a}\beta(x,u)v, \qquad a \in \mathbb{F}.$$

(c) There is an isotropic point $P = \langle u \rangle$ and a non-isotropic vector $v \in P^\perp$ such that X_π consists of the transformations

$$x \mapsto x + a\beta(x,v)u - \bar{a}\beta(x,u)v + b\beta(x,u)u,$$

where $a, b \in \mathbb{F}$ and $a\bar{a}+b+\bar{b} = 0$. Show that these are essentially the transformations $Q(a,b)$ of p. 121.

(d) There is a non-degenerate subspace L of dimension 2 such that X and $\hat{\pi}(X)$ fix $L^\perp$ pointwise and $X_\pi = SU(L)$.

Note that in cases (a), (b) and (c) there is an isotropic point $P = \langle u \rangle$ such that the restrictions of the elements of X_π to $P^\perp$ have the form $x \mapsto x + a\beta(x,v)u$, for some $v \in P^\perp$.

10.18 Suppose that q is a power of a prime and that p is an odd prime that does not divide q. Let f be the least positive integer such that p divides $(-q)^f - 1$.

(*i*) Show that the Sylow p-subgroup of $U(f,q)$ is cyclic.

(*ii*) Given n, let $\ell = \lfloor n/f \rfloor$ and let D be the direct product of ℓ copies of $U(f,q)$. As in Exercises 4.11 and 8.7, show that S_ℓ acts as a group of automorphisms of D and that a Sylow p-subgroup of the semidirect product of D by S_ℓ is isomorphic to a Sylow p-subgroup of $U(n,q)$.

(*iii*) Show that the order of a Sylow p-subgroup of $PSU(n,q)$ is less than $(\sqrt{3}\,(q+1))^n$.

(*iv*) State and prove the result corresponding to (*ii*) when $p = 2$; then show that the order of a Sylow 2-subgroup of $PSU(n,q)$ is at most $(2(q+1))^n$. (For help, see Carter and Fong (1964).)

(*v*) Show that, except for $PSU(3,2)$, $PSU(3,3)$ and $PSU(4,2)$, the largest order of a Sylow subgroup of $PSU(n,q)$, where $n > 2$, is $q^{n(n-1)/2}$.

11

Orthogonal Groups

In Chapters 8 and 10 we considered the groups associated with symplectic and unitary polarities of a projective geometry $\mathcal{P}(V)$, where V is a finite-dimensional vector space over a field $\mathbb{F}$. It follows from Theorem 7.1 that if π is a polarity of $\mathcal{P}(V)$ which is neither of symplectic nor of unitary type, then π is of orthogonal type; i.e., it is induced by a symmetric bilinear form β. In this chapter we deal with the groups which preserve such a form. However, in order to include the orthogonal groups over fields of characteristic 2 we shall assume that the orthogonal geometry is defined by a non-degenerate quadratic form (see p. 54)

$$Q : V \to \mathbb{F}$$

whose polar form is $\beta(u, v) := Q(u + v) - Q(u) - Q(v)$.

In general we have $\beta(v, v) = 2Q(v)$ and therefore β is an alternating form when the characteristic of $\mathbb{F}$ is 2, and Q is completely determined by β when the characteristic of $\mathbb{F}$ is not 2.

By definition, Q is non-degenerate if $V^\perp$ has no singular vectors (see p. 56). When the characteristic of $\mathbb{F}$ is not 2 this corresponds to β being non-degenerate.

The *orthogonal group* associated with V and Q is

$$O(V, Q) := \{ f \in GL(V) \mid Q(f(v)) = Q(v) \text{ for all } v \in V \}.$$

The *full* orthogonal group $\Gamma O(V, Q)$ consists of the σ-semilinear transformations f of V such that for some $a \in \mathbb{F}$

$$Q(f(v)) = a\sigma Q(v) \quad \text{for all } v \in V,$$

the *general* orthogonal group is $GO(V, Q) := \Gamma O(V, Q) \cap GL(V)$ and, except for fields of characteristic 2, the *special* orthogonal group is $SO(V, Q) := O(V, Q) \cap SL(V)$. The derived subgroup $O(V, Q)'$ of $O(V, Q)$ is denoted by $\Omega(V, Q)$.

In general it is possible to have several types of orthogonal geometries with the same underlying vector space and therefore we make the form explicit in the notation for the group. If it is clear from the context which geometry is intended, we shall abbreviate $O(V, Q)$, $\Omega(V, Q)$, etc. to $O(V)$, $\Omega(V)$, etc.

As in the previous chapters our goal is to study the action of $O(V)$ on the projective space $\mathcal{P}(V)$ and ultimately to determine its normal subgroups. If the orthogonal groups behaved analogously to the symplectic and unitary groups we could expect the group $PSO(V)$ induced by $SO(V)$ on $\mathcal{P}(V)$ to be simple whenever $\mathcal{P}(V)$ contained singular points except perhaps for spaces of small dimension. But this is not the case, and in general $\Omega(V)$ is a proper normal subgroup of $SO(V)$. Moreover, except for fields of characteristic 2, there are no orthogonal transvections. Thus to elucidate the structure of these groups we need a substitute for the transvections used in previous chapters. We shall use certain transformations first defined by C. L. Siegel (1938) and later used by M. Eichler (1952) and T. Tamagawa (1958) in their study of orthogonal groups.

The structure of orthogonal geometries is influenced to a large extent by the arithmetic of the underlying field, particularly when the geometry has no singular points. But if there are singular points and the dimension of V is not 4 we can give a uniform proof of the simplicity of $P\Omega(V)$ and we can construct a BN-pair and the associated polar building along the lines indicated in Chapter 9. For these reasons we generally restrict our attention to the geometries with singular points.

Throughout this chapter a *hyperbolic pair* will mean a pair (e, f) of vectors such that $Q(e) = Q(f) = 0$ and $\beta(e, f) = 1$.

Matrices

Suppose that $\mathbb{F}$ is not of characteristic 2. Let $e_1, e_2, \ldots, e_n$ be a basis for V and let $J := (\beta(e_i, e_j))$ be the matrix of β. If A is the matrix of $f \in GL(V)$ with respect to this basis, then by (7.6), $f \in O(V)$ if and only if

$$A^t J A = J.$$

Taking determinants, we see that $(\det A)^2 = 1$. If v is a non-singular vector, the transformation that sends v to $-v$ and acts as the identity on $\langle v \rangle^\perp$ belongs to $O(V)$ and has determinant -1. It follows that $SO(V)$ is a subgroup of index 2 in $O(V)$.

If the field has characteristic 2, then the elements of $O(V)$ preserve the alternating form β (and act as the identity on $V^\perp$). Hence by Corollary 8.6 every element of $O(V)$ has determinant 1.

The scalar transformation $a\mathbf{1}$ belongs to $O(V)$ if and only if $a^2 = 1$ Thus the group $PO(V)$ of collineations of $\mathcal{P}(V)$ induced by $O(V)$ is isomorphic to

$$O(V)/\{\pm\mathbf{1}\}.$$

As usual, if X is the symbol for a group of transformations of V, we use PX to denote the corresponding projective group.

Finite Fields

11.1 Lemma. *If $\mathbb{F}$ is a finite field and $a, b \in \mathbb{F}^\times$, then for all $c \in \mathbb{F}$, there exist x, $y \in \mathbb{F}$ such that $ax^2 + by^2 = c$.*

Proof. If the characteristic of $\mathbb{F}$ is 2, every element of $\mathbb{F}$ is a square and the result is obvious.

If $q = |\mathbb{F}|$ is odd, then the sets $\{\, ax^2 \mid x \in \mathbb{F} \,\}$ and $\{\, c - by^2 \mid y \in \mathbb{F} \,\}$ both contain $\frac{1}{2}(q+1)$ elements and hence they have an element in common. □

11.2 Theorem. *If $\mathbb{F}$ is finite and $\dim V \geq 3$, then V contains a singular vector.*

Proof. Suppose at first that the characteristic of $\mathbb{F}$ is 2 and that $0 \neq u \in V$. By hypothesis, $\dim V \geq 3$ and so $\dim\langle\, u \,\rangle^\perp \geq 2$. Therefore we can choose $v \in \langle\, u \,\rangle^\perp \setminus \langle\, u \,\rangle$ and consequently

$$Q(xu + yv) = x^2 Q(u) + y^2 Q(v).$$

Every element of $\mathbb{F}$ is a square and so there exist $x, y \in \mathbb{F}$ such that $xu+yv \neq 0$ and $Q(xu + yv) = 0$.

Next suppose that the characteristic of $\mathbb{F}$ is odd. Choose non-zero vectors u, v and w such that $v \in \langle\, u \,\rangle^\perp$ and $w \in \langle\, u, v \,\rangle^\perp$. We may suppose that u, v and w are non-singular, and then by Lemma 11.1 we can find x and y such that $x^2 Q(u) + y^2 Q(v) = -Q(w)$. Then $Q(xu + yv + w) = 0$. □

Suppose that V is an orthogonal geometry over $\mathbb{F}_q$. If V contains a singular vector u, then by Lemma 7.3 $\mathcal{P}(V)$ contains a hyperbolic line $L_1 := \langle\, u, v \,\rangle$ and we have $V = L_1 \perp L_1^\perp$. It follows that we can write

$$V = L_1 \perp L_2 \perp \ldots \perp L_m \perp W, \tag{11.3}$$

where L_1, L_2, $\ldots$, L_m are hyperbolic lines and W does not contain any singular vectors. Then m is the Witt index of V and by Theorem 11.2 the dimension of W is 0, 1 or 2. By Witt's Theorem, V is determined up to isomorphism by m and W.

For each i we may suppose that $L_i := \langle\, e_i, f_i \,\rangle$, where (e_i, f_i) is a hyperbolic pair. We have $Q(xe_i + yf_i) = xy$ and therefore $\langle\, e_i \,\rangle$ and $\langle\, f_i \,\rangle$ are the only singular points of L_i.

We have several cases to consider, depending on the dimension of W.

I. If $W = 0$, then $\dim V = 2m$ and we write $O^+(2m, q)$, $\Omega^+(2m, q)$, etc. to denote $O(V)$, $\Omega(V)$, etc. In this case the quadratic form is

$$Q\big(\sum_{i=1}^{m}(x_i e_i + y_i f_i)\big) = \sum_{i=1}^{m} x_i y_i.$$

II. If $W = \langle\, w \,\rangle$, the geometry depends on the value of $Q(w) \pmod{\mathbb{F}_q^2}$, where $\mathbb{F}_q^2 := \{\, a^2 \mid a \in \mathbb{F}_q^\times \,\}$. This is because $Q(aw) = a^2 Q(w)$.

If q is even, there is just one type of geometry. If q is odd there are two types of geometry: those for which $Q(w)$ is a square and those for which it is a non-square. This distinction is important when V occurs as a subgeometry of some larger geometry but it does not affect the group of isometries because the quadratic forms can be interchanged by multiplying by a non-square. Thus we may unambiguously write $O(2m{+}1, q)$ to denote the group. The quadratic form is

$$Q\big(\sum_{i=1}^{m}(x_i e_i + y_i f_i) + zw\big) = \sum_{i=1}^{m} x_i y_i + Q(w) z^2.$$

III. In this case $\dim W = 2$ and it follows from Lemma 11.1 that it is possible to write $W = \langle\, e, f \,\rangle$, where $Q(e) = 1$ and $\beta(e, f) = 1$.

Then $Q(xe + yf) = x^2 + xy + ay^2$, where $a = Q(f)$. The polynomial $X^2 + X + a$ is irreducible over $\mathbb{F}_q$, otherwise W would contain singular vectors.

Thus $\mathbb{F}_{q^2} = \mathbb{F}_q[\omega]$, where $\omega^2 + \omega + a = 0$ and we identify W with $\mathbb{F}_{q^2}$ by mapping $xe + yf$ to $x - y\omega$. The field automorphism σ of $\mathbb{F}_{q^2}$ whose fixed field is $\mathbb{F}_q$ sends ω to $-1 - \omega$, and therefore the quadratic form Q corresponds to the norm map

$$N : \mathbb{F}_{q^2} \to \mathbb{F}_q : a \mapsto a\sigma(a).$$

This shows that up to isomorphism there is just one orthogonal geometry of dimension $2m{+}2$ and Witt index m. In this case we write $O^-(2m{+}2, q)$, $\Omega^-(2m + 2, q)$, etc. to denote $O(V)$, $\Omega(V)$, etc. The quadratic form is

$$Q\big(\sum_{i=1}^{m}(x_i e_i + y_i f_i) + xe + yf\big) = \sum_{i=1}^{m} x_i y_i + x^2 + xy + ay^2.$$

11.4 Theorem. *For $\varepsilon = \pm 1$, $O^\varepsilon(2, q)$ is a dihedral group of order $2(q - \varepsilon)$.*

Proof. For $O^+(2, q)$ we may choose a hyperbolic pair (e, f) and use matrices with respect to the basis e, f. Then $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is in $O^+(2, q)$ if and only if $ac = bd = 0$ and $ad + bc = 1$. There are two possibilities: either $d = a^{-1}$ and $b = c = 0$, or $c = b^{-1}$ and $a = d = 0$. Thus $O^+(2,q)$ is the group

$$\left\{ \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \mid a \in \mathbb{F}_q^\times \right\}$$

extended by the element $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ of order 2; i.e., a dihedral group of order $2(q-1)$.

For $O^-(2,q)$ we may identify the orthogonal geometry with $\mathbb{F}_{q^2} = \mathbb{F}_q[\omega]$, where $\omega^2 + \omega + a = 0$ for some $a \in \mathbb{F}_q$ and take the quadratic form to be the norm map $N : \mathbb{F}_{q^2} \to \mathbb{F}_q$.

The field automorphism σ whose fixed field is $\mathbb{F}_q$ is an element of $O^-(2,q)$. Similarly, for all $b \in \mathbb{F}_{q^2}$ such that $N(b) = 1$, the transformation obtained by multiplying by b is also in $O^-(2,q)$. We shall show that $O^-(2,q)$ is the cyclic group $\{ b \in \mathbb{F}_{q^2} \mid N(b) = 1 \}$ extended by $\langle \sigma \rangle$. Since $\sigma(b) = b^{-1}$ whenever $N(b) = 1$, this is a dihedral group of order $2(q+1)$.

If $g \in O^-(2,q)$, then $b := g(1)$ has norm 1 and so by replacing g with $b^{-1}g$ we may suppose that $g(1) = 1$. We must show that $g \in \langle \sigma \rangle$. It suffices to show that $d := g(\omega)$ is either ω or $\sigma(\omega)$. We have $g(1+\omega) = 1+d$ and so both d and $1+d$ are elements of norm a. That is, $d\sigma(d) = a$ and $(1+d)(1+\sigma(d)) = a$, whence $d^2 + d + a = 0$. Therefore d is either ω or $\sigma(\omega)$, as required. □

Order Formulae, I

Let V be an orthogonal geometry of dimension n and Witt index m over $\mathbb{F}_q$. Put $\varepsilon = 2m - n + 1$ so that ε is 1, 0 or -1 according to whether $O(V)$ is $O^+(2m,q)$, $O(2m+1,q)$ or $O^-(2m+2,q)$, respectively. Let σ_m^ε be the number of singular vectors in V.

11.5 Theorem. $\sigma_m^\varepsilon = (q^{m-\varepsilon} + 1)(q^m - 1)$.

Proof. Suppose that P and Q are singular points such that $P + Q$ is a hyperbolic line. For each singular point $R \in (P+Q)^\perp$ there are q singular points on $P + R$ other than P. Every line of $P^\perp$ through P meets $(P+Q)^\perp$, and as the geometry of $(P+Q)^\perp$ is the same type as V, it follows that there are $1 + q(q-1)^{-1}\sigma_{m-1}^\varepsilon$ singular points in $P^\perp$.

If $R \notin P^\perp$, then $P+R$ is a hyperbolic line and therefore has just one singular point other than P. There are q^{n-2} lines through P not in $P^\perp$ and hence q^{n-2} singular points of $\mathcal{P}$ not in $P^\perp$. It follows that

$$\sigma_m^\varepsilon = q\sigma_{m-1}^\varepsilon + q^{n-1} - q^{n-2} + q - 1.$$

For all ε we have $\sigma_0^\varepsilon = 0$, and therefore

$$\sigma_m^\varepsilon = (q^{m-\varepsilon}+1)(q^m-1). \qquad \square$$

The number of hyperbolic pairs in V is $q^{n-2}\sigma_m^\varepsilon$, and therefore the number of ordered choices of m mutually orthogonal hyperbolic pairs is

$$q^{2m-1-\varepsilon}\sigma_m^\varepsilon q^{2m-3-\varepsilon}\sigma_{m-1}^\varepsilon \cdots q^{1-\varepsilon}\sigma_1^\varepsilon.$$

From (11.3) we see that the order of $O(V)$ is the product of the above expression with $|O(W)|$. We have $O(1,q) = \{\pm\mathbf{1}\}$ and $|O^-(2,q)| = 2(q+1)$, therefore

$$\begin{aligned}
|O^+(2m,q)| &= 2q^{m(m-1)}(q^m-1)\prod_{i=1}^{m-1}(q^{2i}-1)\\
|O^-(2m+2,q)| &= 2q^{m(m+1)}(q^{m+1}+1)\prod_{i=1}^{m}(q^{2i}-1)\\
|O(2m+1,q)| &= \begin{cases} q^{m^2}\displaystyle\prod_{i=1}^{m}(q^{2i}-1) & q \text{ even}\\[2ex] 2q^{m^2}\displaystyle\prod_{i=1}^{m}(q^{2i}-1) & q \text{ odd}\end{cases}
\end{aligned}$$

As a curiosity, notice that except when q is even and n is odd,

$$|O(V)| = 2q^{\lfloor\frac{n}{2}\rfloor(\lfloor\frac{n}{2}\rfloor-1)}(q^{\lfloor\frac{n}{2}\rfloor}-\varepsilon)\prod_{i=1}^{\lfloor\frac{n-1}{2}\rfloor}(q^{2i}-1).$$

Notice too that for odd q, $|SO(2m+1,q)| = |Sp(2m,q)|$. When q is even we *define* $SO(2m+1,q)$ to be $O(2m+1,q)$. We shall see below that when q is even the groups $SO(2m+1,q)$ and $Sp(2m,q)$ are isomorphic but that this is not the case when q is odd.

Order formulae for the finite groups $\Omega(V)$ will be given in a later section after we have a better hold on the structure of these groups.

Three-Dimensional Orthogonal Groups

11.6 Theorem. *If V is an orthogonal geometry of dimension 3 and Witt index 1 over $\mathbb{F}$, then $O(V) \simeq \{\pm\mathbf{1}\} \times SO(V)$, $SO(V) \simeq PGL(2, \mathbb{F})$ and $SO(V)$ acts triply transitively on the set of singular points of $\mathcal{P}(V)$.*

Proof. Let (e, f) be a hyperbolic pair in V and let w be a non-zero element of $\langle\, e, f \,\rangle^\perp$. By replacing Q by a scalar multiple we may suppose that $Q(w) = -1$.

If $g \in SO(V)$ fixes $\langle\, e \,\rangle$, then g fixes $\langle\, e \,\rangle^\perp = \langle\, e, w \,\rangle$ and the matrix of g with respect to the basis e, w, f is upper triangular. If g fixes $\langle\, w \,\rangle$ as well, its matrix has the form

$$H(a) := \begin{pmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a^{-1} \end{pmatrix}.$$

A straightforward calculation (Exercise 11.5) shows that every element of $SO(V)_{\langle\, e \,\rangle}$ can be written as a product of a diagonal matrix $H(a)$ and a matrix of the form

$$S(b) := \begin{pmatrix} 1 & 2b & b^2 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}. \tag{11.7}$$

The singular points of $\mathcal{P}(V)$ have the form $\langle\, xe + yf + zw \,\rangle$, where $xy = z^2$. Thus the set of singular points is

$$\Omega := \{\langle\, e \,\rangle\} \cup \{\, \langle\, b^2e + bw + f \,\rangle \mid b \in \mathbb{F} \,\}.$$

The group $S := \{\, S(b) \mid b \in \mathbb{F} \,\}$ fixes $\langle\, e \,\rangle$ and acts regularly on the remaining points of Ω. Similarly $H := \{\, H(a) \mid a \in \mathbb{F}^+ \,\}$ fixes $\langle\, e \,\rangle$ and $\langle\, f \,\rangle$ and acts regularly on the remaining points.

The linear transformation t such that $t(e) = f$, $t(f) = e$, and $t(w) = -w$ belongs to $SO(V)$ and interchanges $\langle\, e \,\rangle$ and $\langle\, f \,\rangle$. Thus $SO(V)$ acts triply transitively on Ω.

We may identify Ω with the projective line $\mathbb{F} \cup \{\infty\}$ (see p. 23) by assigning $\langle\, e \,\rangle$ to ∞ and $\langle\, z^2e + zw + f \,\rangle$ to $z \in \mathbb{F}$.

The element with matrix $S(b)$ induces the transformation $z \mapsto z + b$, $H(a)$ induces $z \mapsto az$ and t induces $z \mapsto -z^{-1}$. These transformations generate $PGL(2, \mathbb{F})$ and therefore $SO(V) \simeq PGL(2, \mathbb{F})$. □

11.8 Corollary. *If $\dim V = 3$ and the Witt index is 1, then $\Omega(V) \simeq PSL(2, \mathbb{F})$, except that $\Omega(3, 2) \simeq \mathbb{Z}_3$.* □

With this theorem we see hints of things to come. For example, when q is odd, $SO(3,q) \simeq PGL(2,q)$ has a simple subgroup $PSL(2,q)$ of index 2 and when q is even, $O(3,q) = SO(3,q) \simeq PGL(2,q) = Sp(2,q)$. Moreover, the elements of $SO(V)$ with matrices of the form $S(b)$ correspond to transvections in $PSL(2,\mathbb{F})$.

Degenerate Polar Forms and the Group $O(2m+1,2^k)$

Throughout this section we suppose that the map $\sigma : x \mapsto x^2$ is an automorphism of $\mathbb{F}$; for example this is true when $\mathbb{F}$ is a finite field of characteristic 2. Such fields are said to be *perfect*.

Let V be an orthogonal geometry over $\mathbb{F}$ defined by the quadratic form Q whose polar form is β. Let $V^\perp$ be the radical of V with respect to β. For u, $v \in V^\perp$ and $a \in \mathbb{F}$ we have $Q(u+v) = Q(u) + Q(v)$ and $Q(au) = a^2 Q(u)$. Thus $Q : V^\perp \to \mathbb{F}$ is a σ-semilinear transformation onto $\mathbb{F}$. We shall suppose that β is degenerate but that Q is non-degenerate. That is, $V^\perp \neq \{0\}$ and Q does not vanish on the non-zero vectors of $V^\perp$. These assumptions force $\dim V^\perp$ to be 1 and so we may write $V^\perp = \langle\, w \,\rangle$, where $Q(w) = 1$. See Dieudonné (1971), §16 for information about what happens when the field is not perfect.

For $u \in V$, let $\overline{u} := u + V^\perp$ be the image of u in $\overline{V} := V/V^\perp$. Then $\overline{\beta}(\overline{u},\overline{v}) := \beta(u,v)$ defines a non-degenerate alternating form on $\overline{V}$. In particular, $\dim \overline{V} = 2m$ is even and $\dim V = 2m+1$.

Let S be the set of singular vectors of V together with 0. Then for all $v \in V$, $Q(v + \lambda w) = Q(v) + \lambda^2$ and so there is a unique element of S in $\overline{v}$. Thus $S \to \overline{V} : v \mapsto \overline{v}$ is a bijection. Moreover, $W \mapsto \overline{W} := \{\, \overline{v} \mid v \in W \,\}$ is an isomorphism from the partially ordered set of totally singular subspaces of V to the partially ordered set of totally isotropic subspaces of $\overline{V}$. (See Exercise 11.6.)

Similarly, the map $S \to \overline{V}$ induces a bijection between the set of polar frames of $\mathcal{P}(V)$ and the symplectic frames of $\mathcal{P}(\overline{V})$ (see pp. 84 and 69). In other words, the polar building of $\mathcal{P}(V)$ is isomorphic to the polar building of $\mathcal{P}(\overline{V})$.

11.9 Theorem. *$O(V) \simeq Sp(\overline{V})$ and in particular, $O(2m+1,2^k)$ is isomorphic to $Sp(2m,2^k)$.*

Proof. For $f \in O(V)$, let f' be the restriction of f to S and let $\overline{f}$ be the transformation of $\overline{V}$ defined by $\overline{f}(\overline{v}) := \overline{f(v)}$. Then $\overline{f}$ is well-defined and belongs to $Sp(\overline{V})$. Moreover, f' corresponds to $\overline{f}$ via the bijection $S \to \overline{V}$.

We have $f(V^\perp) = V^\perp$ and therefore $f(w) = \lambda w$ for some λ. Applying Q we find that $\lambda^2 = 1$ and hence $\lambda = 1$. Now every element of V can be written uniquely in the form $v + aw$, where $v \in S$ and we have $f(v + aw) = f'(v) + aw$. Thus f is completely determined by its restriction to S and therefore $O(V) \to Sp(\overline{V}) : f \mapsto \overline{f}$ is one-to-one.

Conversely, if we are given $\overline{f} \in Sp(\overline{V})$, then for $v \in S$ define $f(v)$ to be the unique element of S in $\overline{f}(\overline{v})$ and then for $a \in F$ define

$$f(v + aw) := f(v) + aw.$$

For all $u \in V$ we have $\overline{f(u)} = \overline{f}(\overline{u})$, $Q(f(u)) = Q(u)$ and $f(au) = af(u)$. It remains to show that $f(v_1 + v_2) = f(v_1) + f(v_2)$ for all v_1, $v_2 \in S$. We have

$$f(v_1 + v_2) = f(v_1) + f(v_2) + \lambda w$$

for some $\lambda \in \mathbb{F}$. Applying Q we find that $\beta(v_1, v_2) = \beta(f(v_1), f(v_2)) + \lambda^2$ and hence $\overline{\beta}(\overline{v}_1, \overline{v}_2) = \overline{\beta}(\overline{f}(\overline{v}_1), \overline{f}(\overline{v}_2)) + \lambda^2$. But $\overline{f}$ preserves $\overline{\beta}$ and therefore $\lambda = 0$, as required. □

To complete this section we describe the elements of $O(V)$ that correspond to the transvections of $Sp(\overline{V})$. So suppose that $t \in O(V)$ and that $\bar{t}$ is a transvection. Then for some $u \in V$ and $a \in \mathbb{F}$ we have

$$\bar{t}(\overline{v}) = \overline{v} + a\overline{\beta}(\overline{v}, \overline{u})\overline{u}.$$

We may suppose that $u \in S$ and then for all $v \in S$ we have

$$t(v) = v + a\beta(v, u)u + \lambda w,$$

where λ is chosen so that $Q(t(v)) = 0$. That is $a\beta(v, u)^2 + \lambda^2 = 0$. We may write $a = b^2$ so that $\lambda = b\beta(v, u)$. Putting $u' := bu + w$ we find that for all $v \in V$,

$$t(v) = v + \beta(v, u')u'$$

and $Q(u') = 1$. Thus t is a transvection in $O(V)$. By definition a transvection of $O(V)$ fixes every element of a hyperplane of V and this hyperplane contains $V^\perp$. Thus every transvection in $O(V)$ corresponds to a transvection in $Sp(\overline{V})$.

Reflections

From now on let V be an orthogonal geometry over $\mathbb{F}$ defined by a quadratic form Q and suppose that the polar form β of Q is non-degenerate. In this section we determine the orthogonal transformations of V which fix every element of some hyperplane of V. It turns out that, except when the characteristic of $\mathbb{F}$ is 2, these transformations are never transvections. In the next section we shall define the transformations that play the rôle of transvections in orthogonal groups.

11.10 Lemma. *For all $f \in O(V)$ and all integers $k \geq 1$,*

$$\ker(\mathbf{1} - f)^k = \big(\operatorname{im}(\mathbf{1} - f)^k\big)^\perp.$$

Proof. For all u, $v \in V$ we have $\beta((\mathbf{1} - f)u, v) = \beta(f(u), (f - \mathbf{1})v)$ and by induction $\beta((\mathbf{1} - f)^k u, v) = \beta(f^k(u), (f - \mathbf{1})^k v)$. If $v \in \ker(\mathbf{1} - f)^k$, then $\beta((\mathbf{1} - f)^k u, v) = 0$ and therefore $\ker(\mathbf{1} - f)^k \subseteq \big(\operatorname{im}(\mathbf{1} - f)^k\big)^\perp$. These subspaces have the same dimension and consequently they coincide. □

11.11 Theorem. *If $t \in O(V)$ fixes every vector in a hyperplane of V, then t is either the identity or there is a non-singular vector v such that for all $w \in V$,*

$$t(w) = w - Q(v)^{-1}\beta(w, v)v.$$

Proof. If $t \in GL(V)$ and $\ker(\mathbf{1} - t)$ is a hyperplane, then we know (from p. 20) that t has the form $t(w) = w + \varphi(w)v$, for some $v \in V$, $\varphi \in V^*$. As β is non-degenerate, the map $V \to V^* : u \mapsto \beta(-, u)$ is an isomorphism and we may suppose that $\varphi := \beta(-, u)$. Then t can be written

$$t(w) = w + \beta(w, u)v \tag{11.12}$$

for some u, $v \in V$. Thus $\ker(\mathbf{1} - t) = \langle\, u \,\rangle^\perp$ and $\operatorname{im}(1 - t) = \langle\, v \,\rangle$. If $t \in O(V)$, then from the lemma, $\langle\, u \,\rangle = \langle\, v \,\rangle$, and so $u = av$ for some a. Moreover t preserves Q and therefore

$$a\beta(w, v)^2 + a^2\beta(w, v)^2 Q(v) = 0$$

for all $w \in V$. Thus v is non-singular and $a = -Q(v)^{-1}$. Hence

$$t(w) = w - Q(v)^{-1}\beta(w, v)v. \qquad \square$$

We call $t_P := t$ the *reflection* in the hyperplane $P^\perp$, where $P := \langle\, v \,\rangle$ and we note that it depends only on P, not v.

The determinant of t_P is -1 and $t_P^2 = \mathbf{1}$. If $f \in O(V)$, then $f t_P f^{-1} = t_{f(P)}$. The reflection t_P is a transvection if and only if $\beta(v, v) = 0$ and this is the case if and only if the characteristic of $\mathbb{F}$ is 2. Thus in general orthogonal groups do not contain transvections.

The calculations of this section show that $f \in O(V)$ is a reflection if and only if $\dim(\operatorname{im}(\mathbf{1} - f)) = 1$. More generally, we have

11.13 Theorem. *If t is the reflection in $\langle v \rangle^\perp$ and $f \in O(V)$, then*

(i) $v \in \operatorname{im}(\mathbf{1}-f)$ *implies* $\operatorname{im}(\mathbf{1}-tf) = \operatorname{im}(\mathbf{1}-f) \cap \langle u \rangle^\perp$, *where* $v = (\mathbf{1}-f)u$.

(ii) $v \notin \operatorname{im}(\mathbf{1}-f)$ *implies* $\operatorname{im}(\mathbf{1}-tf) = \operatorname{im}(\mathbf{1}-f) \oplus \langle v \rangle$.

Proof. (i) Suppose that $v = u - f(u)$. Then $Q(v) = \beta(u,v)$ and therefore $t(u) = u - v = f(u)$. Thus $u \in \ker(\mathbf{1}-tf)$ and it follows from Lemma 11.10 that $\operatorname{im}(\mathbf{1}-tf) \subseteq \langle u \rangle^\perp$.

Writing $\mathbf{1} - tf = (\mathbf{1}-t)f + (\mathbf{1}-f)$ we see that $\operatorname{im}(\mathbf{1}-tf) \subseteq \operatorname{im}(\mathbf{1}-f)$. As $u \notin \ker(\mathbf{1}-f)$ it follows that $\operatorname{im}(\mathbf{1}-f) \not\subseteq \langle u \rangle^\perp$ and therefore $\operatorname{im}(\mathbf{1}-tf) = \operatorname{im}(\mathbf{1}-f) \cap \langle u \rangle^\perp$.

(ii) We have $w \in \ker(\mathbf{1}-tf)$ if and only if $t(w) = f(w)$ and this is the case if and only if $(\mathbf{1}-f)w = Q(v)^{-1}\beta(w,v)v$. But $v \notin \operatorname{im}(\mathbf{1}-f)$ and therefore this last condition is equivalent to $w \in \ker(\mathbf{1}-f) \cap \langle v \rangle^\perp$. Thus

$$\ker(\mathbf{1}-tf) = \ker(\mathbf{1}-f) \cap \langle v \rangle^\perp$$

and by Lemma 11.10 we have $\operatorname{im}(\mathbf{1}-tf) = \operatorname{im}(\mathbf{1}-f) \oplus \langle v \rangle$. □

11.14 Corollary. (i) $\dim(\operatorname{im}(\mathbf{1}-tf)) = \dim(\operatorname{im}(\mathbf{1}-f)) \pm 1$.

(ii) *If f is a product of k reflections then,* $\dim(\operatorname{im}(\mathbf{1}-f)) \equiv k \pmod 2$. □

In the semidirect product $V.O(V)$ we have $C_V(f) = \ker(\mathbf{1}-f)$ and $[V,f] = \operatorname{im}(\mathbf{1}-f)$. This is the notation used by Aschbacher (1986) and we shall use it too from time to time.

Root Groups

In order to apply Iwasawa's criterion to $P\Omega(V)$ later in this chapter we need, for each singular point P, an abelian normal subgroup of $P\Omega(V)_P$ whose conjugates generate $P\Omega(V)$. In the case of the symplectic and unitary groups, root groups of transvections played this rôle. But we have just seen that, except in characteristic 2, the orthogonal groups do not contain transvections. The goal of this section is to describe the root groups of $O(V)$. We assume that the Witt index of V is at least 1.

Suppose at first that the characteristic of $\mathbb{F}$ is not 2. From the symmetric form β we obtain an automorphism $\hat{\pi}$ of $GL(V)$, where $f^\perp := \hat{\pi}(f)$ is determined by (7.7), i.e.,

$$\beta(f^\perp(u), f(v)) = \beta(u,v) \tag{11.15}$$

for all u, $v \in V$. In particular, $f \in O(V)$ if and only if $f = f^\perp$.

The transformation t defined by equation (11.12) is a transvection whenever $\beta(u,v)=0$. Using (11.15) it can be verified that

$$t^{\perp}(x) = x - \beta(x,v)u$$

and therefore, for the root group

$$X_{u,v} := \{\, x \mapsto x + a\beta(x,u)v \mid a \in \mathbb{F}\,\} \tag{11.16}$$

of $SL(V)$, we have $\hat{\pi}(X_{u,v}) = X_{v,u}$. We shall examine the possibilities for $\langle\, X_{u,v}, X_{v,u}\,\rangle \cap O(V)$. (See Exercises 8.3 and 10.17 for the analogous results for $Sp(V)$ and $U(V)$.)

The root group $X_{u,v}$ of (11.16) is well-defined and consists of transvections even for fields of characteristic 2, and so from now on we allow $\mathbb{F}$ to be an arbitrary field. We put $\widehat{X}_{u,v} := \langle\, X_{u,v}, X_{v,u}\,\rangle \cap O(V)$, where $\beta(u,v)=0$.

I. $u = v$. In this case the elements of $\widehat{X}_{u,u}$ are transvections and therefore $\widehat{X}_{u,u} = 1$ except when the characteristic of $\mathbb{F}$ is 2.

II. $\langle\, u,v\,\rangle$ *is totally isotropic with respect to* β.
Suppose that $t_1(x) := x + a\beta(x,v)u$ and $t_2(x) := x + b\beta(x,u)v$. Then

$$t_1t_2(x) = t_2t_1(x) = x + a\beta(x,v)u + b\beta(x,u)v$$

and this transformation belongs to $O(V)$ if and only if

$$a^2\beta(x,v)^2Q(u) + b^2\beta(x,u)^2Q(v) + (a+b)\beta(x,u)\beta(x,v) = 0$$

for all $u,\ v \in V$. Choosing x in $\langle\, u\,\rangle^{\perp} \setminus \langle\, u,v\,\rangle^{\perp}$, $\langle\, v\,\rangle^{\perp} \setminus \langle\, u,v\,\rangle^{\perp}$, and $V \setminus (\langle\, u\,\rangle^{\perp} \cup \langle\, v\,\rangle^{\perp})$ in succession, we find that $b^2Q(v) = a^2Q(u) = a+b = 0$. Thus in order that $\widehat{X}_{u,v}$ be non-trivial, u and v must be singular and $b = -a$. That is,

$$\widehat{X}_{u,v} = \{\, x \mapsto x + a\beta(x,v)u - a\beta(x,u)v \mid a \in \mathbb{F}\,\}.$$

It is easily checked (Exercise 11.7) that $\widehat{X}_{u,v}$ depends only on the totally singular subspace $\langle\, u,v\,\rangle$ and not on the particular choice of basis $u,\ v$.

III. u *is singular,* $v \in \langle\, u\,\rangle^{\perp}$ *and* $\beta(v,v) \neq 0$. Note that this implies that the characteristic of $\mathbb{F}$ is not 2. By Lemma 7.3 we may choose $w \in \langle\, v\,\rangle^{\perp}$ such that $\beta(u,w) = 1$ and $Q(w) = 0$. Then $\langle\, u,v,w\,\rangle$ is non-degenerate and the elements of $X_{u,v}$ and $X_{v,u}$ leave $\langle\, u,v,w\,\rangle$ invariant and fix every vector of $\langle\, u,v,w\,\rangle^{\perp}$. Moreover, with respect to the basis $u,\ v,\ w$, the matrices of $\langle\, X_{u,v}, X_{v,u}\,\rangle$ (acting on $\langle\, u,v,w\,\rangle$) are upper triangular with 1's on the diagonal. A short calculation, similar to the one leading to (11.7), shows that the matrices have the form

$$\begin{pmatrix} 1 & 2aQ(v) & -a^2Q(v) \\ 0 & 1 & -a \\ 0 & 0 & 1 \end{pmatrix}.$$

In other words,

$$\widehat{X}_{u,v} = \{\, x \mapsto x + a\beta(x,v)u - a\beta(x,u)v - a^2Q(v)\beta(x,u)u \mid a \in \mathbb{F}\,\}.$$

IV. $\beta(u,u) \neq 0$, $\beta(v,v) \neq 0$, *and* $\beta(u,v) = 0$. Again $\mathbb{F}$ cannot be of characteristic 2, and in this case $\widehat{X}_{u,v}$ is isomorphic to $O(\langle\, u, v\,\rangle)$.

The transformations given in II and III above are the appropriate generalizations of transvections to orthogonal groups and in these cases we say that the groups $\widehat{X}_{u,v}$ are *root groups* for $O(V)$.

Siegel Transformations

If u is singular and $v \in \langle\, u\,\rangle^\perp$, we put

$$\rho_{u,v}(x) := x + \beta(x,v)u - \beta(x,u)v - Q(v)\beta(x,u)u \tag{11.17}$$

and note that $\rho_{u,v}$ is a well-defined element of $SO(V)$ for fields of all characteristics. Moreover, if v is singular, we obtain the transformations of case II whereas if v is non-singular, we obtain those of case III of the previous section. Following Tamagawa (1958) and Higman (1978) we call $\rho_{u,v}$ a *Siegel transformation.* (In Hahn (1979) it is called an Eichler transformation.)

11.18 Theorem. *If u is singular and $v \in \langle\, u\,\rangle^\perp$, then $\rho_{u,v}$ is the unique element of $O(V)$ whose restriction to $\langle\, u\,\rangle^\perp$ has the form $x \mapsto x + \beta(x,v)u$.*

Proof. Certainly the restriction of $\rho_{u,v}$ to $\langle\, u\,\rangle^\perp$ has this form. If f is another element of $O(V)$ with the same restriction to $\langle\, u\,\rangle^\perp$, then $f^{-1}\rho_{u,v}$ fixes every element of $\langle\, u\,\rangle^\perp$. In Theorem 11.11 we showed that such a transformation is either the identity or the reflection in $\langle\, u\,\rangle^\perp$. But u is singular and there is no reflection in $\langle\, u\,\rangle^\perp$. Thus $f^{-1}\rho_{u,v} = 1$, as required. □

The following is the analogue of Theorem 4.2 for Siegel transformations.

11.19 Theorem. *If u is singular, v, v_1 and $v_2 \in \langle\, u\,\rangle^\perp$ and $f \in O(V)$, then*

(i) $\rho_{au,v} = \rho_{u,av}$ *for all* $a \in \mathbb{F}^\times$.

(ii) $\rho_{u,v_1+v_2} = \rho_{u,v_1}\rho_{u,v_2}$.

(iii) $f\rho_{u,v}f^{-1} = \rho_{f(u),f(v)}$.

Proof. In all cases it suffices to consider the restriction to $\langle\, u\,\rangle^\perp$ where the calculations are straightforward (and follow from Theorem 4.2). □

Now let $P := \langle\, u\,\rangle$ be a singular point of $\mathcal{P}(V)$, and put

$$X_P := \{\, \rho_{u,v} \mid v \in \langle\, u\,\rangle^\perp \,\}.$$

11.20 Theorem. *If $P := \langle\, u\,\rangle$ is singular, then X_P is an abelian normal subgroup of $O(V)_P$ and isomorphic to the additive group of $P^\perp/P$.*

Proof. We use Theorem 11.19. It follows from (i) and (iii) that X_P is a normal subgroup of $O(V)_P$. And it follows from (ii) that the map $P^\perp \to X_P : v \mapsto \rho_{u,v}$ is a homomorphism onto X_P. If $\rho_{u,v} = \mathbf{1}$, then for all $x \in P^\perp$ we have $\beta(x,v) = 0$ and therefore $v \in P^{\perp\perp} = P$. Thus $P^\perp/P \simeq X_P$. □

11.21 Lemma. *The orthogonal geometry V is spanned by its non-singular vectors, except in the case of the hyperbolic line over $\mathbb{F}_2$.*

Proof. Suppose that u is singular. If $\dim V > 2$, then $\langle\, u\,\rangle^\perp$ contains a non-singular vector v. Then $u = (u+v) - v$, and both $u+v$ and v are non-singular. Thus if V is not spanned by its non-singular vectors, it is a hyperbolic line. But a hyperbolic line has only two singular points and it is clear that it is spanned by its non-singular vectors except when the field is $\mathbb{F}_2$ (in which case there is just one non-singular vector). □

11.22 Corollary. *For all singular points $P := \langle\, u\,\rangle$, the group X_P is generated by the Siegel transformations $\rho_{u,v}$, where v is non-singular, except when V is the orthogonal geometry of dimension 4 and Witt index 2 over $\mathbb{F}_2$.* □

It is clear from (11.17) that $\operatorname{im}(\mathbf{1} - \rho_{u,v}) = \langle\, u, v\,\rangle$ and by Theorem 11.13, if v is non-singular, then $\rho_{u,v}$ is the product of two reflections. The next theorem gives an explicit description of these reflections. Recall that t_P denotes the reflection in $P^\perp$.

11.23 Theorem. *If u is singular and $v \in \langle\, u\,\rangle^\perp$ is non-singular, then*

$$\rho_{u,v} = t_{\langle\, v\,\rangle} t_{\langle\, Q(v)u - v\,\rangle}.$$

Proof. Choose w so that $v = (\mathbf{1} - \rho_{u,v})w$. Then from (11.17) we have $\beta(w, u) = 1$ and $\beta(w, Q(v)u - v) = 0$. From Theorem 11.13,

$$\operatorname{im}(\mathbf{1} - t_{\langle v \rangle}\rho_{u,v}) = \langle u, v \rangle \cap \langle w \rangle^{\perp} = \langle Q(v)u - v \rangle.$$

Thus $t_{\langle v \rangle}\rho_{u,v} = t_{\langle Q(v)u-v \rangle}$, as required. □

11.24 Theorem. *Except in the case of $\Omega^+(4, 2)$, all Siegel transformations belong to $\Omega(V)$.*

Proof. If $\rho_{u,v}$ is a Siegel transformation with v non-singular, then by the previous theorem, $\rho_{u,v} = t_{\langle v \rangle}t_{\langle v' \rangle}$, where $v' := Q(v)u - v$. As $Q(v') = Q(v)$ it follows from Witt's theorem that $f(v) = v'$ for some $f \in O(V)$. Then $ft_{\langle v \rangle}f^{-1} = t_{\langle v' \rangle}$ and so

$$\rho_{u,v} = t_{\langle v \rangle}ft_{\langle v \rangle}f^{-1} \in O(V)' = \Omega(V).$$

It now follows from Corollary 11.22 that, except for $\Omega^+(4, 2)$, every Siegel transformation belongs to $\Omega(V)$. □

To conclude this section we shall show that $\Omega^+(4, 2)$ is indeed an exception to Theorem 11.24. So suppose that V is the orthogonal geometry of dimension 4 and Witt index 2 over $\mathbb{F}_2$. We may write $V = L_1 \perp L_2$, where $L_1 := \langle u_1, v_1 \rangle$ and $L_2 := \langle u_2, v_2 \rangle$ are hyperbolic lines. There are 6 non-singular vectors in V, namely the non-zero elements of the subspaces

$$\begin{aligned} V_1 &:= \{0, u_1 + v_1, u_1 + u_2 + v_2, u_2 + v_1 + v_2\}, \quad \text{and} \\ V_2 &:= \{0, u_2 + v_2, u_1 + u_2 + v_1, u_1 + v_1 + v_2\} \end{aligned}$$

We have $V = V_1 \perp V_2$ and so the elements of $O(V)$ either interchange V_1 and V_2, or fix them. From Theorem 11.4 (or directly), $O(V_1) \simeq O(V_2) \simeq S_3$ and therefore $O(V) \simeq (S_3 \times S_3)\langle \rho \rangle$, where ρ is a transformation that interchanges V_1 and V_2. In fact we may take ρ to be ρ_{u_1,u_2} and this shows that $\rho_{u_1,u_2} \notin \Omega^+(4, 2)$.

The reflections of $O^+(4, 2)$ are the elements of order 2 in $O(V_1)$ and $O(V_2)$ and thus the group they generate is $O(V_1) \times O(V_2)$. The group $\Omega^+(4, 2) = O^+(4, 2)'$ is a subgroup of index 2 in $O(V_1) \times O(V_2)$ (Exercise 11.8).

The Action of $P\Omega(V)$ on Singular Points

In this section we assume, in addition to the polar form β being non-degenerate, that the Witt index of V is at least 1 and the dimension of V is at least 3.

11.25 Lemma. *The group $PO(V)$ acts faithfully on the set of singular points of $\mathcal{P}(V)$.*

Proof. Suppose that $f \in O(V)$ fixes every singular point of $\mathcal{P}(V)$. If u is non-singular, then $u \in L$ for some hyperbolic line L. As $\dim V \geq 3$ there is a non-singular vector $w \in L^\perp$. Then $W := \langle L, w \rangle$ is non-degenerate and the calculations in the proof of Theorem 11.6 show that f fixes every point of $\mathcal{P}(W)$. Thus f fixes every point of $\mathcal{P}(V)$ and so $f = \pm\mathbf{1}$. This proves that $PO(V) := O(V)/\{\pm\mathbf{1}\}$ acts faithfully on the set of singular points. □

11.26 Lemma. *For each singular point P, the group X_P acts regularly on the set of singular points not orthogonal to P.*

Proof. Suppose that $P := \langle u \rangle$, and that $Q := \langle v \rangle$ and $R := \langle w \rangle$ are singular points not orthogonal to P. We may choose v and w so that $\beta(u,v) = \beta(u,w) = 1$. Now $V = (P+Q) \perp (P+Q)^\perp$ and writing $w = au + bv + x$ for some $x \in (P+Q)^\perp$ we see that $b = 1$ and $a = -Q(x)$. Thus $w = -Q(x)u + v + x$ and $\rho_{u,-x}(v) = w$.

If for some $y \in P^\perp$, $\rho_{u,y}(Q) = Q$, then from (11.17) we see that $y \in P + Q$. In fact $y \in P$, otherwise we would have $v \in P^\perp$, contrary to our assumption. Thus $\rho_{u,y} = \mathbf{1}$. □

Eventually we shall prove that, except for $\Omega^+(4,2)$, $\Omega(V)$ is generated by the Siegel transformations. But for the moment we don't assume this.

11.27 Lemma. *For all orthogonal geometries V of dimension at least 3 the subgroup of $O(V)$ generated by the Siegel transformations is transitive on the singular points of $\mathcal{P}(V)$.*

Proof. Given singular points $P := \langle u \rangle$ and $Q := \langle v \rangle$ we shall produce a Siegel transformation that takes P to Q.

If $\beta(u,v) = 0$, choose a singular vector w such that $\beta(u,w) = \beta(v,w) = 1$. (By Lemma 7.5 there are vectors u' and v' such that (u,u') and (v,v') are orthogonal hyperbolic pairs. Put $w := u' + v'$.) Now use the previous lemma to obtain an element of $X_{\langle w \rangle}$ that takes P to Q.

If $\beta(u,v) \neq 0$, then $P + Q$ is a hyperbolic line and we may suppose that $\beta(u,v) = 1$. Taking our cue from the previous lemma, let w be a non-singular vector in $(P+Q)^\perp$ (which is non-zero because $\dim V \geq 3$) and let $x := u - Q(w)v + w$. Then x is singular, $\beta(u,x) = -Q(w)$, and $\beta(v,x) = 1$. By the previous lemma there is an element of $X_{\langle x \rangle}$ that takes P to Q. □

11.28 Theorem. *If the Witt index is 1 and $\dim V \geq 3$, then $P\Omega(V)$ acts doubly transitively on the singular points of $\mathcal{P}(V)$.*

Proof. By Theorem 11.24, $P\Omega(V)$ contains all the Siegel transformations and so the result follows from the lemmas just proved. □

In an earlier section we showed that when $\dim V = 3$ and the Witt index is 1, then $\Omega(V) \simeq PSL(2, \mathbb{F})$, and so this corollary should come as no surprise. The exceptional case of $\Omega(3, 2)$ does not arise here because we have excluded odd dimensional geometries over fields of characteristic 2.

11.29 Lemma. *Suppose the dimension of V is at least 5 and the Witt index of V is at least 2. Then for all singular points P, Q and R such that Q, $R \in P^\perp \setminus \{P\}$ there is a product of Siegel transformations which fixes P and takes Q to R.*

Proof. (We follow Higman (1978), p. 50.) If Q is not orthogonal to R, then $Q + R$ is a hyperbolic line and $P \in (Q + R)^\perp$. Let L be a hyperbolic line of $(Q + R)^\perp$ that contains P. Then $Q + R \subseteq L^\perp$ and $\dim L^\perp \geq 3$, therefore by Lemma 11.27 the group generated by the Siegel transformations of $L^\perp$ acts transitively on the singular points of $\mathcal{P}(L^\perp)$. A Siegel transformation of $L^\perp$ extends to a Siegel transformation of V which fixes L pointwise, and hence there is a product of Siegel transformations in $\Omega(V)_P$ which takes Q to R.

If Q is orthogonal to R, then by Lemma 7.5 there is a singular point $S \in P^\perp$ not orthogonal to Q or to R. Now we can use Siegel transformations in $\Omega(V)_P$ to move Q to S, and then S to R. □

11.30 Theorem. *If the Witt index of V is at least 2 and the dimension of V is at least 5, then $P\Omega(V)$ is a primitive rank 3 group on the singular points of $\mathcal{P}(V)$.*

Proof. Lemmas 11.26, 11.27 and 11.29 show that $P\Omega(V)$ is a rank 3 group on the singular points. To show that it is primitive we repeat the argument of Theorem 8.3:

Suppose that B is a block of imprimitivity such that $|B| > 1$ and choose $P \in B$. If $B \cap P^\perp$ contains a point other than P, then by the theorem $P^\perp \subseteq B$. In this case, if R is a singular point not in $P^\perp$, choose a singular point $Q \in (P + R)^\perp$. Then $Q \in B$ and $R \in Q^\perp$, hence $R \in B$. Thus B consists of all the singular points.

Now suppose that B contains a singular point not in $P^\perp$. Then by Lemma 11.26, B contains all the singular points not in $P^\perp$. Suppose that R is a singular point of $P^\perp$, $R \neq P$ and (using Lemma 7.5) choose a singular point $Q \notin P^\perp \cup R^\perp$. Then $Q \in B$, and since $R \notin Q^\perp$, it follows that $R \in B$. Again B consists of all the singular points of $\mathcal{P}(V)$, and this proves that $P\Omega(V)$ is primitive. □

Wall's Parametrization of $O(V)$

Ultimately we shall show that every orthogonal group except $O^+(4,2)$ is generated by its reflections. This is a relatively straightforward matter except for fields of characteristic 2, particularly $\mathbb{F}_2$. But it turns out that there is a uniform approach to this question using a description of the elements of $O(V)$ due to G. E. Wall (1959 and 1963). Using this method, for each $f \in O(V)$, we can completely determine the shortest length of any expression for f as a product of reflections. In fact this parametrization of $O(V)$ is interesting in its own right and later we shall use it to obtain a particularly elegant criterion to determine which elements of $O(V)$ belong to $\Omega(V)$.

From now on, for $f \in O(V)$, we use $[V, f]$ to denote $\operatorname{im}(\mathbf{1} - f)$. As usual we assume that the orthogonal geometry is defined by a quadratic form Q with non-degenerate polar form β. We make no assumptions about the Witt index at this stage.

For $f \in O(V)$ and u, $v \in [V, f]$ we define the *Wall form* of f to be

$$\chi_f(u, v) := \beta(w, v), \tag{11.31}$$

where w is some vector such that $u = w - f(w)$.

11.32 Theorem. *χ_f is a well-defined non-degenerate bilinear form on $[V, f]$ such that $\chi_f(u, u) = Q(u)$ for all $u \in [V, f]$.*

Proof. If $u = w - f(w) = w' - f(w')$, then $w - w' \in \ker(\mathbf{1} - f) = [V, f]^\perp$ (from Lemma 11.10). Thus for $v \in [V, f]$, $\beta(w, v) = \beta(w', v)$ and therefore $\chi_f(u, v)$ is well-defined.

If $\chi_f(u, v) = 0$ for all $u \in [V, f]$, then $\beta(w, v) = 0$ for all $w \in V$, and so $v = 0$. Thus χ_f is non-degenerate. On putting $f' := \mathbf{1} - f$, the equation $Q(f(w)) = Q(w)$ becomes $Q(f'(w)) = \beta(w, f'(w))$ and therefore $\chi_f(u, u) = Q(u)$ for all $u \in [V, f]$. $\square$

Note that χ_f is not necessarily a reflexive form.

11.33 Theorem. *The assignment $f \mapsto ([V, f], \chi_f)$ is a one-to-one correspondence between $O(V)$ and the set of pairs (I, χ), where I is a subspace of V and χ is a non-degenerate bilinear form on I such that $\chi(u, u) = Q(u)$ for all $u \in I$.*

Proof. Suppose that $[V, f] = [V, g]$ and $\chi_f = \chi_g$. Then for $w \in V$ and $v \in [V, f]$, (11.31) implies $\chi_f((\mathbf{1} - f)w, v) = \beta(w, v) = \chi_g((\mathbf{1} - g)w, v)$. Since $\chi_f = \chi_g$ is non-degenerate, $(\mathbf{1} - f)w = (\mathbf{1} - g)w$ for all $w \in V$, and hence $f = g$.

Conversely, suppose that I is a subspace of V and that χ is a non-degenerate bilinear form on I such that $\chi(u,u) = Q(u)$ for all $u \in I$. As χ is non-degenerate there is a unique linear transformation $f' : V \to V$ such that

$$\chi(f'(u), v) = \beta(u, v)$$

for all $u \in V$ and $v \in I$. Putting $f := \mathbf{1} - f'$ we find that, for all $u \in V$, $Q(f(u)) = Q(u)$, and therefore $f \in O(V)$. From the definition of f', $f'(u) = 0$ if and only if $\beta(u,v) = 0$ for all $v \in I$. Thus $I = (\ker f')^\perp = [V,f]$ and $\chi_f = \chi$. □

The bijection $f \leftrightarrow ([V,f], \chi_f)$ is Wall's parametrization of $O(V)$. There are similar parametrizations of the symplectic and unitary groups. (See Wall (1963) for applications.)

11.34 Lemma. *For all f, $g \in O(V)$ and all u, $v \in [V,f]$*

(*i*) $\beta(u,v) = \chi_f(u,v) + \chi_f(v,u)$.

(*ii*) $\chi_f(f(u),v) = -\chi_f(v,u)$.

(*iii*) $[V,f] = [V,f^{-1}]$ *and* $\chi_{f^{-1}}(u,v) = \chi_f(v,u)$.

(*iv*) $[V,gfg^{-1}] = g([V,f])$ *and* $\chi_{gfg^{-1}}(g(u),g(v)) = \chi_f(u,v)$.

Proof. (*i*) is obtained from the equation $Q(u) = \chi_f(u,u)$ by replacing u by $u+v$ and (*ii*) follows from (*i*) and the definition of χ_f. In (*iii*) it is clear that $[V,f] = [V,f^{-1}]$ and then the rest of (*iii*) as well as (*iv*) follows directly from (11.31). □

Because χ_f is not necessarily reflexive we must be more careful than usual when dealing with orthogonal complements. We continue to use $X^\perp$ to denote the orthogonal complement of X with respect to β and if W is a subspace of $[V,f]$ we define its *left* and *right* orthogonal complements (with respect to the form χ_f) to be

$$\begin{aligned} {}^{\triangleleft}W &:= \{\, v \in [V,f] \mid \chi_f(v,w) = 0 \text{ for all } w \in W \,\}, \quad \text{and} \\ W^{\triangleright} &:= \{\, v \in [V,f] \mid \chi_f(w,v) = 0 \text{ for all } w \in W \,\}, \end{aligned}$$

respectively.

11.35 Lemma. (*i*) ${}^{\triangleleft}(W^{\triangleright}) = W = ({}^{\triangleleft}W)^{\triangleright}$.

(*ii*) $\dim {}^{\triangleleft}W = \dim W^{\triangleright} = \dim[V,f] - \dim W$.

(*iii*) ${}^{\triangleleft}W = f(W)^{\triangleright} = f(W^{\triangleright})$.

(*iv*) *If* $f(W) = W$, *then* ${}^{\triangleleft}W = W^{\triangleright} = W^\perp \cap [V,f]$.

Proof. (*i*) and (*ii*) are immediate consequences of the non-degeneracy of χ_f. Parts (*ii*) and (*i*) of Lemma 11.34 imply (*iii*) and (*iv*) respectively. □

Factorization Theorems

We shall use Wall's parametrization to obtain factorizations of the elements of $O(V)$. It is always the case that $[V, f_1 f_2] \subseteq [V, f_1]+[V, f_2]$ but the following theorems show that under certain circumstances much more can be said.

11.36 Theorem. *For $f \in O(V)$, suppose that the restriction χ_1 of χ_f to a subspace I_1 of $[V, f]$ is non-degenerate, and let χ_2 be the restriction of χ_f to $I_2 := I_1^{\triangleright}$. If f_1 and f_2 are the elements of $O(V)$ corresponding to (I_1, χ_1) and (I_2, χ_2), then*

(*i*) $[V, f] = I_1 \oplus I_2$,

(*ii*) $f = f_1 f_2$, *and*

(*iii*) $f_1 f_2 = f_2 f_1$ *if and only if* $[V, f] = I_1 \perp I_2$. *In this case* f_1 *coincides with* f *on* $I_2^{\perp}$, *and* f_2 *coincides with* f *on* $I_1^{\perp}$.

Proof. (*i*) The form χ_f is non-degenerate, and therefore $[V, f] = I_1 \oplus I_1^{\triangleright}$.

(*ii*) If $u_1 \in I_1$ $u_2 \in I_2$, then $\chi_f(u_1, u_2) = 0$ and from Lemma 11.34 (*i*), $\chi_f(u_2, u_1) = \beta(u_2, u_1)$. From (11.31) we have $\chi_f((\mathbf{1} - f_1)w, u_1) = \beta(w, u_1)$ and $\chi_f((\mathbf{1} - f_2)w, u_2) = \beta(w, u_2)$. Writing $\mathbf{1} - f_1 f_2$ as $(\mathbf{1} - f_1)f_2 + (\mathbf{1} - f_2)$ we see that

$$\begin{aligned}\chi_f((\mathbf{1} - f_1 f_2)w, u_1 + u_2) &= \beta(f_2(w), u_1) + \beta((\mathbf{1} - f_2)w, u_1) + \beta(w, u_2)\\ &= \beta(w, u_1 + u_2)\\ &= \chi_f((\mathbf{1} - f)w, u_1 + u_2)\end{aligned}$$

This holds for all $u_1 \in I_1$, $u_2 \in I_2$, and therefore $f = f_1 f_2$.

(*iii*) If $f_1 f_2 = f_2 f_1$, then from Lemma 11.34 (*iv*) f fixes I_1. Consequently ${}^{\triangleleft}I_1 = I_1^{\triangleright}$ and $[V, f] = I_1 \perp I_2$. Conversely, if $I = I_1 \perp I_2$, then reversing the rôles of f_1 and f_2 in (*ii*) leads to the conclusion $f = f_2 f_1$. Finally, if $x \in I_2^{\perp}$, then $f_2(x) = x$ and therefore $f(x) = f_1(x)$. Similarly, if $x \in I_1^{\perp}$, then $f(x) = f_2(x)$. ◻

It follows from Lemma 11.10 that the restriction of $\mathbf{1} - f$ to $[V, f]$ is invertible if and only if $[V, f]$ is non-degenerate (with respect to β). In this case we say that f is *regular*; and we define f to be *unipotent* if $\mathbf{1} - f$ is *nilpotent*; i.e., if $(\mathbf{1} - f)^k = 0$ for some k.

For example, the Siegel transformation $\rho_{u,v}$ satisfies $(\mathbf{1} - \rho_{u,v})^3 = 0$ and therefore it is unipotent. (If the characteristic is 2, then $(\mathbf{1} - \rho_{u,v})^2 = 0$.)

11.37 Theorem. *Every $f \in O(V)$ has a factorization of the form $f = f_r f_u = f_u f_r$, where f_r is regular and f_u is unipotent. Moreover, $[V, f] = [V, f_r] + [V, f_u]$ and $g \in O(V)$ commutes with f if and only if it commutes with f_r and f_u.*

Proof. The chain of subspaces $\operatorname{im}(\mathbf{1} - f) \supseteq \operatorname{im}(\mathbf{1} - f)^2 \supseteq \ldots$ cannot decrease indefinitely and therefore there exists k such that $\operatorname{im}(\mathbf{1} - f)^k = \operatorname{im}(\mathbf{1} - f)^{k+1}$. We put $I_r := \operatorname{im}(\mathbf{1} - f)^k$ and we note that by Lemma 11.10 $I_r^\perp = \ker(\mathbf{1} - f)^k$.

If $v \in I_r \cap I_r^\perp$, then $v = (\mathbf{1} - f)^k w$ for some $w \in V$ and so $(\mathbf{1} - f)^{2k} w = 0$. Thus $w \in \ker(\mathbf{1} - f)^{2k} = \ker(\mathbf{1} - f)^k$, whence $v = 0$. This shows that $I_r \cap I_r^\perp = \{0\}$ and therefore $V = I_r \perp I_r^\perp$.

Put $I_u := [V, f] \cap I_r^\perp$ and observe that, as f fixes I_r, Lemma 11.35 implies $I_u = I_r^{\triangleright}$. It follows from the previous theorem that $f = f_r f_u = f_u f_r$, where $[V, f_r] = I_r$, $[V, f_u] = I_u$ and χ_{f_r} and χ_{f_u} are restrictions of χ_f. By the theorem, f_u coincides with f on $I_r^\perp = \ker(\mathbf{1} - f)^k$ and as $I_r \subseteq I_u^\perp = \ker(\mathbf{1} - f_u)$ it follows that f_u is nilpotent.

Finally, if g commutes with f, then g fixes I_r and I_u and by Lemma 11.34 (*iv*) $\chi_f(g(u), g(v)) = \chi_f(u, v)$. But χ_{f_r} and χ_{f_u} are obtained by restricting χ_f and thus another application of Lemma 11.34 shows that g commutes with f_r and f_u. □

We call f_r and f_u the *regular* and *unipotent parts* of f. But note that it is possible to have $f = f_1 f_2 = f_2 f_1$ with f_1 regular, f_2 unipotent, and $f_1 \neq f_r$ (cf. Exercise 11.9).

The Generation of $O(V)$ by Reflections

Cartan showed that when V is an orthogonal geometry of dimension n over the real or complex field, every element of $O(V)$ is a product of at most n reflections. This was subsequently proved by Dieudonné (1948) for arbitrary fields of characteristic $\neq 2$. Then Scherk (1950) for fields of characteristic $\neq 2$, and Dieudonné (1955) for fields of characteristic 2, proved that with certain exceptions every element $f \in O(V)$ is a product of $\dim[V, f]$ reflections. An error in the paper of Dieudonné (1955) led to several exceptions being overlooked. This was pointed out, and the error corrected, by Callan (1976).

We shall obtain these results using Theorem 11.13 and Wall's parametrization of $O(V)$.

First observe that if $u_1, u_2, \ldots, u_h$ are non-singular vectors of V and if $f := t_{\langle u_1 \rangle} t_{\langle u_2 \rangle} \cdots t_{\langle u_h \rangle}$, where $t_{\langle u_i \rangle}$ is the reflection in $\langle u_i \rangle^\perp$, then by Theorem 11.13, $\dim[V, f] \leq h$. Moreover, if equality holds, then $u_1, u_2, \ldots, u_h$ is a basis for $[V, f]$. Therefore, if $f \in O(V)$ and $r := \dim[V, f]$, it is not possible to write f as a product of fewer than r reflections. If $[V, f]$ is totally singular, even r reflections will not suffice.

11.38 Lemma. *Suppose that χ is a non-degenerate bilinear form on a finite-dimensional vector space W over $\mathbb{F}$. If $\mathbb{F} \neq \mathbb{F}_2$ and if χ is not alternating, then W has a basis $e_1, e_2, \ldots, e_m$ such that $\chi(e_i, e_i) \neq 0$ for $i = 1, 2, \ldots, m$ and $\chi(e_i, e_j) = 0$ for $i < j$.*

Proof. Since χ is not alternating, there is a vector $u \in W$ such that $\chi(u, u) \neq 0$. Thus the result is true when $\dim W = 1$, and henceforth we may suppose that $\dim W \geq 2$. In this case $W = \langle u \rangle \oplus \langle u \rangle^{\triangleright}$ and the restriction of χ to $\langle u \rangle^{\triangleright}$ is non-degenerate.

By induction we may suppose that χ restricts to an alternating form on $\langle u \rangle^{\triangleright}$. This forces $\dim\langle u \rangle^{\triangleright}$ to be even. Choose a non-zero vector $v \in \langle u \rangle^{\triangleright}$ and consider $\chi(u + av, u + av) = \chi(u, u) + a\chi(v, u)$. This cannot be 0 for all choices of $a \neq 0$, otherwise $\chi(u, u)$ would be 0. (It is here that we use the assumption that $\mathbb{F} \neq \mathbb{F}_2$.) Thus we may suppose v has been chosen so that $c := \chi(u + v, u + v) \neq 0$. Now choose $w \in \langle u \rangle^{\triangleright}$ so that $\chi(v, w) = 1$. Then for all b, $\chi(u + v, u + bv - cw) = 0$ and

$$\chi(u + bv - cw, u + bv - cw) = \chi(u, u) + b\chi(v, u) - c\chi(w, u).$$

If for some b this quantity does not vanish, the restriction of χ to $\langle u + v \rangle^{\triangleright}$ is non-degenerate and not alternating. Thus we may set $e_1 := u + v$ and obtain a basis of the desired form by induction.

If the above quantity vanishes for all b, then $\chi(v, u) = 0$ and $\chi(w, u) = 1$. We repeat the argument of the previous paragraphs with w in place of v. That is, replacing w by a suitable multiple we may suppose that $c' := \chi(u+w, u+w) \neq 0$. Next we choose $w' \in \langle u \rangle^{\triangleright}$ so that $\chi(w, w') = 1$. Since $\chi(w, u) \neq 0$, the previous argument shows that the restriction of χ to $\langle u + w \rangle^{\triangleright}$ is non-degenerate and not alternating. We put $e_1 := u+w$ and once again we obtain a suitable basis by induction. □

11.39 Theorem. *Suppose that V is an orthogonal geometry over a field $\mathbb{F}$, where $\mathbb{F} \neq \mathbb{F}_2$. Then every element $f \in O(V)$ is a product of $\dim[V, f]$ reflections except when $[V, f]$ is totally singular in which case f is a product of $\dim[V, f] + 2$ reflections. In particular, f is a product of at most $\dim V$ reflections.*

Proof. Let χ be the Wall form of f. If $[V, f]$ is not totally singular, then by Theorem 11.33, χ is not alternating and by Lemma 11.38 $[V, f]$ has a basis $e_1, e_2, \ldots, e_r$ of non-singular vectors such that $\chi(e_i, e_j) = 0$ for $i < j$. By Theorem 11.36 we have $f = t_1 t_2 \cdots t_r$, where $[V, t_i] = \langle e_i \rangle$. Thus for all i, t_i is the reflection in $\langle e_i \rangle^{\perp}$.

If $[V, f]$ is totally singular and u is any non-singular vector not in $[V, f]$, then by Theorem 11.13 (*ii*), $[V, t_{\langle u \rangle} f] = [V, f] \oplus \langle u \rangle$. But now $[V, t_{\langle u \rangle} f]$ is not totally singular and by the previous paragraph, $t_{\langle u \rangle} f$ is a product of $\dim[V, f]+1$ reflections. Thus f is a product of $\dim[V, f]+2$ reflections. □

In order to deal with orthogonal groups over $\mathbb{F}_2$ we first prove a somewhat weaker version of Lemma 11.38.

11.40 Lemma. *Suppose that W is a vector space over $\mathbb{F}_2$ and that χ is a non-degenerate symmetric bilinear form on W. If there is a vector $w \in W$ such that $\chi(w, w) = 1$ and such that the restriction of χ to $\langle w \rangle^{\triangleright}$ is alternating, then W has a basis $e_1, e_2, \ldots, e_m$ such that $\chi(e_i, e_j) = 0$ for $i < j$ and $\chi(e_i, e_i) = 1$ for all i.*

Proof. Choose a symplectic basis $u_1, v_1, u_2, v_2, \ldots, u_k, v_k$ for $\langle w \rangle^{\perp}$, then set $e_1 := w + u_1$ and $e_2 := w + u_1 + v_1$. Then $\chi(e_1, e_1) = \chi(e_2, e_2) = 1$ and $\chi(e_1, e_2) = 0$. The subspace

$$\langle e_1, e_2 \rangle^{\perp} = \langle w + v_1, u_2, v_2, \ldots, u_k, v_k \rangle$$

satisfies the same conditions as W and so the result follows by induction.

□

11.41 Theorem. *Suppose that V is an orthogonal geometry over $\mathbb{F}_2$ and for $f \in O(V)$ let $r := \dim[V, f]$.*

(*i*) *If $O(V) \neq O^{+}(4, 2)$, then f is a product of r, $r+2$ or $r+4$ reflections. Moreover, f is a product of $r+4$ reflections (and no fewer) if and only if $[V, f]$ is totally singular and $\dim V = 2r$.*

(*ii*) *If $O(V) = O^{+}(4, 2)$ and if f is a product of reflections, then f is a product of r reflections.*

Proof. (*i*) We first show that f is a product of r, $r+2$ or $r+4$ reflections. If $u \in [V, f]$ is non-singular, then by Theorem 11.13 (*i*), $\dim[V, t_{\langle u \rangle} f] = r-1$. Thus, by induction on r, we may assume that $U := [V, f]$ is totally singular. In this case the Wall form χ of f is alternating and by Lemma 7.5 we may write $V = U^{\perp} \oplus U'$, where U' is totally singular.

Suppose at first that $\dim V > 2r$ and choose $w \in (U \oplus U')^{\perp}$ such that $Q(w) = 1$. Then $[V, t_{\langle w \rangle} f] = U \perp \langle w \rangle$ and by Theorem 11.36 (*ii*) the Wall form χ' of $t_{\langle w \rangle} f$ restricts to χ on U. Thus χ' satisfies the conditions of Lemma 11.40. It follows that $U \perp \langle w \rangle$ has a basis $e_1, e_2, \ldots, e_{r+1}$ of non-singular vectors such that $\chi'(e_i, e_j) = 0$ for $i \neq j$. By Theorem 11.36 we have $t_{\langle w \rangle} f = t_{\langle e_1 \rangle} \cdots t_{\langle e_{r+1} \rangle}$ and so f is a product of $r+2$ reflections.

Now suppose that $\dim V = 2r$. By assumption $r \neq 2$ and, since the alternating form χ is non-degenerate, r must be even. If $L \subseteq U$ is a hyperbolic line for χ, then $U = L \oplus L^{\triangleright}$ and by Theorem 11.36 we have $f = f_1 f_2$, where $[V, f_1] = L$ and $[V, f_2] = L^{\triangleright}$. It follows from the previous paragraph that f_1 is the product of 4 reflections and f_2 is the product of r reflections. Thus f is the product of $r+4$ reflections and this completes the proof that every element of $O(V)$ is a product of r, $r+2$ or $r+4$ reflections. At the same time we have shown that, except when $[V, f]$ is totally singular and $\dim V = 2r$, f is the product of r or $r+2$ reflections.

We still must show that when $[V, f]$ is totally singular and $\dim V = 2r$, f is not the product of fewer than $r+4$ reflections. If $f = t_{\langle w \rangle} f'$, then $w \notin [V, f]$ and by Theorem 11.13, $[V, f'] = [V, f] \oplus \langle w \rangle$. As before we may write $V = U \oplus U'$, where $U := [V, f]$ and U' is totally singular. Write $w = u_1 + u_1'$, where $u_1 \in U$ and $u_1' \in U'$, and put $v_1 := (\mathbf{1} - f)u_1'$. Then $\chi(v_1, u_1) = \beta(u_1', u_1) = 1$ and therefore $L := \langle u_1, v_1 \rangle$ is a hyperbolic line with respect to χ. Let L' be the orthogonal complement of L in U with respect to χ. Then for $v' \in L'$, $\beta(u_1', v') = \chi(v_1, v') = 0$ and $\beta(u_1', v_1) = \beta(v_1, v_1) = 0$. Consequently

$$Q(aw + bu_1 + cv_1 + v') = a + ab,$$

and therefore the non-singular elements of $U \oplus \langle w \rangle$ are all of the form $w + cv_1 + v'$. The subspace generated by the non-singular elements of $U \oplus \langle w \rangle$ is therefore $\langle w, v_1 \rangle \oplus L' \neq U \oplus \langle w \rangle$. On the other hand, if f' were the product of $r+1$ reflections $t_{\langle e_1 \rangle} \cdots t_{\langle e_{r+1} \rangle}$, then $e_1, e_2, \ldots, e_{r+1}$ would generate $U \oplus \langle w \rangle$. This contradiction combined with Corollary 11.14 shows that f cannot be written as the product of fewer than $r+4$ reflections.

(ii) Suppose that $f \in O^+(4, 2)$ and that $U := [V, f]$ is totally singular. Then $\dim U = 2$ and there is a unique non-degenerate alternating form on U. That is, there is only one possibility for the Wall form of f and therefore f is uniquely determined by U. Thus f is a Siegel transformation ρ_{u_1, u_2}, where $\langle u_1, u_2 \rangle = U$ and we have seen in the last part of the section 'Siegel Transformations' that these elements are not in the subgroup R generated by reflections. Therefore, if $f \in R$, then $[V, f]$ is not totally singular and it follows as in the first part of (i) that f is a product of r reflections. $\square$

If $[V, f]$ is totally singular, then this theorem shows that f is a product of no fewer than $[V, f] + 2$ reflections. Examples show (Exercise 11.11) that the converse is not true.

11.42 Corollary. *Every orthogonal group except $O^+(4, 2)$ is generated by reflections.* $\square$

Dickson's Invariant

For $f \in O(V)$ define the *Dickson invariant* of f to be

$$D(f) := \dim[V, f] \pmod 2$$

and regard $D(f)$ as an element of the (additive) group $\mathbb{Z}_2$ of order 2. The following theorem can be found in Dye (1977).

11.43 Theorem. *The map $D : O(V) \to \mathbb{Z}_2$ is a homomorphism.*

Proof. If $O(V) \neq O^+(4,2)$, then $O(V)$ is generated by its reflections. If $t_1, t_2, \ldots, t_s$ are reflections, then from Corollary 11.14 (*ii*), $D(t_1 t_2 \cdots d_s) \equiv s \pmod 2$. In particular, D is a homomorphism.

We may regard $O^+(4,2)$ as a subgroup of $O^+(4,4)$ and then $D : O^+(4,2) \to \mathbb{Z}_2$ is the restriction of $D : O^+(4,4) \to \mathbb{Z}_2$, which we have just shown to be a homomorphism. □

If the characteristic of $\mathbb{F}$ is not 2 and t is a reflection, then $\det(t) = -1$ and $D(t) = 1$. Hence for all $f \in O(V)$, $D(f) = 0$ if and only if $\det(f) = 1$. Therefore $SO(V) = \ker D$ in this case.

If the characteristic of $\mathbb{F}$ is 2, we *define* $SO(V)$ to be $\ker D$. Thus whenever the polar form of the orthogonal geometry is non-degenerate it is the case that $|O(V) : SO(V)| = 2$.

11.44 Theorem. *If $O(V) \neq O^+(4,2)$, then $f \in O(V)$ belongs to $SO(V)$ if and only if it is a product of an even number of reflections.* □

The Simplicity of $P\Omega(V)$

We continue to assume that V is an orthogonal geometry defined by a quadratic form whose polar form is non-degenerate, and that the Witt index of V is at least 1. We shall use Iwasawa's criterion to prove that in this case most of the groups $P\Omega(V)$ are simple. In order to do this we need to show that $\Omega(V)$ is generated by Siegel transformations and that $\Omega(V)' = \Omega(V)$.

11.45 Theorem. *If $\dim V \geq 3$ and $O(V) \neq O^+(4,2)$, then $\Omega(V) = SO(V)'$ and $\Omega(V)$ is generated by the commutators $[t_1, t_2]$, where t_1 and t_2 are reflections.*

Proof. If H is the subgroup of $O(V)$ generated by all $[t_1, t_2] = (t_1 t_2)^2$, where t_1 and t_2 are reflections, then $H \subseteq \Omega(V) := O(V)'$ and $O(V)/H$ is abelian. Thus $\Omega(V) \subseteq H$; whence $H = \Omega(V)$.

Suppose that u and v are non-singular. If $U = \langle u, v \rangle$ and $w \in U^\perp$ is non-singular, then $t_{\langle w \rangle}$ commutes with $t_{\langle u \rangle}$ and $t_{\langle v \rangle}$. Hence

$$[t_{\langle u \rangle}, t_{\langle v \rangle}] = [t_{\langle u \rangle} t_{\langle w \rangle}, t_{\langle v \rangle} t_{\langle w \rangle}] \in SO(V)'.$$

If $U^\perp$ is totally singular, then $U^\perp \subseteq U$. Since U contains non-singular vectors we cannot have $U^\perp = U$ and so $\dim V = 3$ and the Witt index of V is 1. It follows from Theorem 11.6 that $[t_{\langle u \rangle}, t_{\langle v \rangle}] \in SO(V)'$. Since $O(V)$ is generated by reflections, $\Omega(V) \subseteq SO(V)'$. On the other hand, $SO(V)' \subseteq O(V)' = \Omega(V)$ and therefore $\Omega(V) = SO(V)'$. □

11.46 Theorem. *If* $\dim V \geq 3$, *the Witt index of* V *is at least 1, and* $\Omega(V) \neq \Omega^+(4,2)$, *then* $\Omega(V)$ *is generated by the Siegel transformations of* V.

Proof. If H is the subgroup of $O(V)$ generated by the Siegel transformations, then from Theorem 11.24, $H \subseteq \Omega(V)$.

Let L be a hyperbolic line. If $u \in V$ is non-singular, then there exists $u' \in L$ such that $Q(u') = Q(u)$ and hence there exists $f \in O(V)$ such that $f(u') = u$. By Lemmas 11.26 and 11.27, there exists $g \in H$ such that $f(L) = g(L)$ and therefore $t_{\langle u \rangle} = g t_{\langle w \rangle} g^{-1}$, where $w := g^{-1}(u) \in L$. As H is a normal subgroup of $O(V)$ it follows that

$$t_{\langle u \rangle} = t_{\langle w \rangle}[t_{\langle w \rangle}, g] \in O(L)H,$$

where we regard $O(L)$ as a subgroup of $O(V)$ (acting trivially on $L^\perp$). Now $O(V)$ is generated by reflections and therefore $O(V) = O(L)H$. Moreover, $H \subseteq SO(V)$ and hence $SO(V) = SO(L)H$. Therefore

$$SO(V)/H \simeq SO(L)/H \cap SO(L),$$

and as $SO(L)$ is abelian (Exercise 11.2), $\Omega(V) = SO(V)' \subseteq H$. It follows that $\Omega(V) = H$. □

11.47 Theorem. *If* $\dim V \geq 3$ *and* V *contains singular vectors, then* $\Omega(V)' = \Omega(V)$ *except for* $\Omega(3,3)$, $\Omega^+(4,2)$ *and* $\Omega^+(4,3)$.

Proof. We shall show that, except in the excluded cases, every Siegel transformation $\rho_{u,v}$ belongs to $\Omega(V)'$, for all singular points $\langle u \rangle$ and all $v \in \langle u \rangle^\perp \setminus \langle u \rangle$. So suppose that $\Omega(V)$ is not $\Omega(3,3)$, $\Omega^+(4,2)$ or $\Omega^+(4,3)$ and let (u, w) be a hyperbolic pair in $\langle v \rangle^\perp$ through u. Put $L := \langle u, w \rangle$.

For $a \in \mathbb{F}^\times$ there exists $g \in O(L)$ such that $g(u) = au$ and $g(w) = a^{-1}w$. Extend g to V by making it act as the identity on $L^\perp$. By Theorem 11.45, $g^2 \in \Omega(V)$. If $|\mathbb{F}| > 3$, we may choose a so that $a^2 \neq 1$, and then (by Theorem 11.19)

$$\rho_{u,v} = [g^2, \rho_{u,(a^2-1)^{-1}v}] \in \Omega(V)'.$$

Thus $\Omega(V) = \Omega(V)'$ except possibly when $\mathbb{F}$ is $\mathbb{F}_2$ or $\mathbb{F}_3$. Corollary 11.22 shows that to prove $\Omega(V) = \Omega(V)'$ it suffices to show that $\rho_{u,v} \in \Omega(V)'$ whenever v is non-singular. Thus from now on we suppose that $Q(v) \neq 0$.

Suppose that $\mathbb{F} = \mathbb{F}_2$. By Lemma 11.21 not every non-singular vector of $L^\perp$ belongs to $\langle v \rangle^\perp$ and therefore there exists $x \in L^\perp$ such that $\langle x, v \rangle$ has Witt index 0. Define $g \in O(V)$ by putting $g(x) := x + v$, $g(v) := x$, and $g(z) := z$ for all $z \in \langle x, v \rangle^\perp$. Then $g^3 = \mathbf{1}$, and so (by Theorem 11.45) $g \in \Omega(V)$. It follows that

$$\rho_{u,v} = [g, \rho_{u,x}] \in \Omega(V)'.$$

Suppose that $\mathbb{F} = \mathbb{F}_3$. We shall show that there exists $x \in L^\perp \cap \langle v \rangle^\perp$ such that $Q(x) = Q(v)$. If $\dim V = 4$, then $L^\perp = \langle v \rangle \perp \langle x \rangle$ for some x. In this case, Q(x) = Q(v), otherwise $Q(x) = -Q(v)$ and then $x + v$ would be singular, contrary to assumption. If $\dim V > 4$, we may write $L^\perp = \langle v \rangle \perp \langle x \rangle \perp \langle y \rangle \perp L_1$. If $Q(x) = Q(y) = -Q(v)$, then $Q(x + y) = Q(v)$, and so we can always choose the notation so that $Q(x) = Q(v)$.

Now choose $g \in O(V)$ so that $g(u) = u$, $g(x) = -v$, and $g(v) = x$. Then $g^2 \in \Omega(V)$ and

$$\rho_{u,v} = [g^2, \rho_{u,v}] \in \Omega(V)'.$$

In all cases, $\Omega(V) = \Omega(V)'$. □

11.48 Theorem. *Let V be an orthogonal geometry defined by a quadratic form of Witt index at least 1 whose polar form is non-degenerate. If* $\dim V \geq 3$*, then $P\Omega(V)$ is a simple group except for $P\Omega(3, 3)$, and except when* $\dim V = 4$ *and the Witt index is 2.*

Proof. By Theorem 11.20, for each singular point P, the group X_P of Siegel transformations fixing P is an abelian normal subgroup of $\Omega(V)_P$. Moreover, by Theorem 11.46, the groups X_P generate $\Omega(V)$. By Theorems 11.28 and 11.30 the action of $\Omega(V)$ on the singular points of $\mathcal{P}(V)$ is primitive and, by Theorem 11.47, $\Omega(V) = \Omega(V)'$. The simplicity of $P\Omega(V)$ now follows from Iwasawa's criterion (Theorem 1.2). □

It is a consequence of Theorem 11.6 that $P\Omega(3, q) \simeq PSL(2, q)$. In particular, $P\Omega(3, 3) \simeq A_4$, which is not simple. When $\dim V = 4$ and the Witt index is 2, it turns out that the action of $P\Omega(V)$ on the singular points of $\mathcal{P}(V)$ is imprimitive; thus Iwasawa's criterion does not apply in this case. On the other hand, the action of $O(V)$ on the singular points *is* primitive (Exercise 11.15). We shall see later that the groups $P\Omega^+(4, \mathbb{F})$ are not simple.

The Spinor Norm

Even though we now know that $P\Omega(V)$ is simple when V contains singular vectors and the dimension of V is large enough, we have not yet determined the index of $\Omega(V)$ in $O(V)$, nor do we know when $-\mathbf{1}$ belongs to $\Omega(V)$. The usual approach to these questions is via Clifford algebras and the spinor norm: see, for example, Artin (1957), Aschbacher (1986), Dieudonné (1971), or Higman (1978). We shall use the spinor norm, but instead of Clifford algebras we use Wall's parametrization of $O(V)$. For further details and the connection between this approach and Clifford algebras see the articles by 'Lipschitz' (1959), Zassenhaus (1962) and Hahn (1979).

We begin by reviewing the definition of the discriminant of a bilinear form. If χ is a bilinear form defined on a vector space W with basis $e_1, e_2, \ldots, e_m$, then χ is non-degenerate if and only if the determinant of $X := \big(\chi(e_i, e_j)\big)$ is non-zero. Suppose that χ is non-degenerate. If $e'_1, e'_2, \ldots, e'_m$ is another basis of W, then $e'_j = \sum_{i=1}^m a_{ij} e_i$ for some non-singular matrix $A := \big(a_{ij}\big)$, and if $X' := \big(\chi(e'_i, e'_j)\big)$, then $X' = A^t X A$. Thus $\det X' = (\det A)^2 \det X$, and

$$\operatorname{disc}(\chi) := (\det X)\mathbb{F}^2$$

is a well-defined element of $\mathbb{F}^\times/\mathbb{F}^2$, where $\mathbb{F}^2 := \{\, a^2 \mid a \in \mathbb{F}^\times \,\}$. We call $\operatorname{disc}(\chi)$ the *discriminant* of χ.

Now suppose that V is an orthogonal geometry defined by a quadratic form Q whose polar form β is non-degenerate. For $f \in O(V)$, the Wall form χ_f is a non-degenerate bilinear form on $[V, f]$ and we define the *spinor norm* of f to be

$$\theta(f) := \operatorname{disc}(\chi_f).$$

The spinor norm of a reflection is easy to calculate. Indeed, if u is non-singular and if t is the reflection in $\langle\, u \,\rangle^\perp$, then $[V, t] = \langle\, u \,\rangle$ and, from Theorem 11.32, $\theta(t) = Q(u)\mathbb{F}^2$.

Next, consider the Siegel transformation $\rho_{u,v}$. Then $[V, \rho_{u,v}] = \langle\, u, v \,\rangle$ and from Theorem 11.32 $\chi_{\rho_{u,v}}(u, u) = 0$. Thus from Lemma 11.34 (i) and the fact that $\beta(u, v) = 0$, the discriminant of $\chi_{\rho_{u,v}}$ is $\chi_{\rho_{u,v}}(u, v)^2\mathbb{F}^2$. Therefore $\theta(\rho_{u,v}) = \mathbb{F}^2$.

We have noted before that non-zero scalar multiples of Q define the same orthogonal group $O(V)$. On the other hand, the spinor norm depends on Q and we have

11.49 Lemma. *If θ' is the spinor norm corresponding to $Q' := aQ$, then for all $f \in O(V)$,*

$$\theta'(f) = a^{D(f)}\theta(f),$$

where $D(f)$ is the Dickson invariant of f.

Proof. This is immediate from the definition of the spinor norm. □

It follows from this lemma that the restriction of the spinor norm to $SO(V)$ depends only on $SO(V)$ and not on Q.

11.50 Theorem. *The spinor norm $\theta : O(V) \to \mathbb{F}^\times/\mathbb{F}^2$ is a homomorphism.*

Proof. Let t be the reflection in $\langle u \rangle^\perp$ and let f be an element of $O(V)$. If $u \in [V, f]$, then by Theorem 11.36 $f = tf'$, where $[V, f'] = \langle u \rangle^\triangleright$ and the Wall form of f' is the restriction of χ_f to $[V, f']$. It follows that $\theta(f) = \theta(t)\theta(f')$ and hence $\theta(tf) = \theta(t)\theta(f)$.

If $u \notin [V, f]$, then by Theorem 11.13 (*ii*), $[V, tf] = [V, f] \oplus \langle u \rangle$ and we may apply the previous argument to tf to conclude that $\theta(tf) = \theta(t)\theta(f)$.

If $t_1, t_2, \ldots, t_k$ are reflections, it follows by induction on k that $\theta(t_1 t_2 \cdots t_k) = \theta(t_1)\theta(t_2)\cdots\theta(t_k)$. If $O(V)$ is generated by reflections, this proves that θ is a homomorphism. The only orthogonal group not generated by reflections is $O^+(4, 2)$. But in this case $\mathbb{F}^\times = \mathbb{F}^2$ and so θ is the trivial homomorphism. □

It is an easy exercise to check that, except for $SO^+(4, 2)$, the group $SO(V)$ is generated by products $t_1 t_2$, where t_1 and t_2 are reflections. Similarly, for all V, $\Omega(V) = \langle f^2 \mid f \in O(V) \rangle$.

11.51 Theorem. *If $\dim V \geq 2$ and V contains singular vectors, then except for $\Omega^+(4, 2)$, we have*

$$\Omega(V) = \{ f \in SO(V) \mid \theta(f) = \mathbb{F}^2 \}$$

and $SO(V)/\Omega(V) \simeq \mathbb{F}^\times/\mathbb{F}^2$.

Proof. The group $\Omega(V)$ is generated by the elements g^2, where $g \in O(V)$. Thus $\theta(f) = \mathbb{F}^2$ for all $f \in \Omega(V)$. Let (u, v) be a hyperbolic pair in V and put $L := \langle u, v \rangle$. For $a \in \mathbb{F}^\times$, let t_1 and t_2 be the reflections in $\langle u - v \rangle$ and $\langle u - av \rangle$. Then $f := t_1 t_2 \in SO(V)$, $f(u) = au$, and $f(v) = a^{-1}v$. We have $\theta(f) = \theta(t_1)\theta(t_2) = a\mathbb{F}^2$ and therefore $\theta : SO(V) \to \mathbb{F}^\times/\mathbb{F}^2$ is surjective.

If $V = L$, then the isometry f just described is a typical element of $SO(L)$. If $\theta(f) = \mathbb{F}^2$, then $a = b^2$ for some b, and $f = g^2$, where $g(u) = bu$ and $g(v) = b^{-1}v$. It follows that $f \in \Omega(L)$ and hence that $\Omega(L) = SO(L) \cap \ker\theta$.

From now on suppose that $V \neq L$ and consider $f := t_1 t_2 \cdots t_k$, where t_i is the reflection in $\langle u_i \rangle^\perp$. For $i = 1, 2, \ldots, k$, choose $v_i \in L$ such that

$Q(v_i) = Q(u_i)$. In general, if $Q(x) = Q(y) \neq 0$, then by Witt's theorem there exists $h \in O(V)$ such that $h(x) = y$ and hence $ht_{\langle x \rangle}h^{-1} = t_{\langle y \rangle}$. By definition, $\Omega(V) = O(V)'$ and therefore $t_{\langle x \rangle} \equiv t_{\langle y \rangle} \pmod{\Omega(V)}$. In particular, $f \equiv g \pmod{\Omega(V)}$, where $g := t_{\langle v_1 \rangle} t_{\langle v_2 \rangle} \cdots t_{\langle v_k \rangle}$. If $f \in SO(V)$ and $\theta(f) = \mathbb{F}^2$, then k is even, $g \in SO(V)$, and $\theta(g) = \theta(f) = \mathbb{F}^2$. In order to show that $f \in \Omega(V)$ it suffices to show that $g \in \Omega(V)$, and as g acts trivially on $L^\perp$ this follows from the first part of the proof. □

Order Formulae, II

If V is a non-degenerate orthogonal geometry of dimension n over $\mathbb{F}_q$, then $|O(V) : SO(V)| = 2$, except when n is odd and q is even, in which case $O(V) = SO(V) = \Omega(V)$. If $\dim V = 2m$, let ε be $+1$ or -1 according to whether the Witt index is m or $m-1$. Then from our earlier formulae

$$|SO^\varepsilon(2m, q)| = q^{m(m-1)}(q^m - \varepsilon) \prod_{i=1}^{m-1} (q^{2i} - 1),$$

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

Theorem 11.51 shows that if q is odd, $|SO(V) : \Omega(V)| = 2$, and if q is even, $\Omega(V) = SO(V)$, except for $\Omega^+(4, 2)$. Moreover, if q is odd, $-\mathbf{1} \in SO(V)$ if and only if n is even. Thus $P\Omega(V) = \Omega(V)$, except possibly when n is even and q is odd. To settle this case we need the value of $\theta(-\mathbf{1})$.

Assume that q is odd and suppose at first that $V := \langle e, f \rangle$, where (e, f) is a hyperbolic pair. Then the matrix of the Wall form of $-\mathbf{1}$ with respect to the basis e, f is $\begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}$ and therefore $\theta(-\mathbf{1}) = (-1)\mathbb{F}^2$. If $\dim V = 2m$ and the Witt index of V is m, then V is the orthogonal sum of m hyperbolic lines. In this case $\theta(-\mathbf{1}) = (-1)^m\mathbb{F}^2$ and therefore $-\mathbf{1} \in \Omega^+(2m, q)$ if and only if $q^m \equiv 1 \pmod 4$.

If $\dim V = 2$ and V has no singular vectors, then case III of the section 'Finite Fields' shows that we may suppose that $V := \langle e, f \rangle$, where $Q(e) = \beta(e, f) = 1$, $Q(f) = a$, and the polynomial $X^2 + X + a$ is irreducible. The matrix of the Wall form of $-\mathbf{1}$ with respect to e, f is $\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & a \end{pmatrix}$ and thus $\theta(-\mathbf{1}) = (a - \frac{1}{4})\mathbb{F}^2 = (-s)\mathbb{F}^2$, where $s \notin \mathbb{F}^2$. If $\dim V = 2m$ and the Witt index of V is $m-1$, it follows that $\theta(-\mathbf{1}) = (-1)^m s\mathbb{F}^2$. Thus $-\mathbf{1} \in \Omega^-(2m, q)$ if and only if $q^m \equiv -1 \pmod 4$.

Thus, except for $P\Omega^+(4,2)$,

$$|P\Omega^\varepsilon(2m,q)| = \frac{1}{d} q^{m(m-1)}(q^m-\varepsilon)\prod_{i=1}^{m-1}(q^{2i}-1),$$

$$|P\Omega(2m+1,q)| = \begin{cases} q^{m^2}\displaystyle\prod_{i=1}^{m}(q^{2i}-1) & q \text{ even} \\ \dfrac{1}{2}q^{m^2}\displaystyle\prod_{i=1}^{m}(q^{2i}-1) & q \text{ odd} \end{cases}$$

where d is the greatest common divisor of 4 and $q^m - \varepsilon$.

From these formulae we see that the orders of the low dimensional orthogonal groups coincide with the orders of groups introduced in earlier chapters. Indeed,

$$\begin{aligned} |P\Omega(3,q)| &= |PSL(2,q)|, \\ |P\Omega^+(4,q)| &= |PSL(2,q)|^2, \\ |P\Omega^-(4,q)| &= |PSL(2,q^2)|, \\ |P\Omega(5,q)| &= |PSp(4,q)|, \\ |P\Omega^+(6,q)| &= |PSL(4,q)|, \quad \text{and} \\ |P\Omega^-(6,q)| &= |PSU(4,q)|. \end{aligned}$$

These coincidences of order all arise from isomorphisms of the corresponding groups. The first is familiar to us (Theorem 11.6) and the others will be dealt with in a uniform way in Chapter 12.

It is also true that $|P\Omega(2m+1,q)| = |PSp(2m,q)|$ for all q. In the next section we prove the remarkable result that when q is odd and $m > 2$, these groups are *not* isomorphic.

The Groups $PSp(2m,q)$ and $\Omega(2m+1,q)$, q odd

In the section 'Degenerate Polar Forms and the Groups $O(2m+1,2^k)$', we showed that, when q is even, the groups $PSp(2m,q)$ and $\Omega(2m+1,q)$ are isomorphic. However, in this section we shall show that when $m > 2$ and q is odd, these groups are not isomorphic. We do this by counting the conjugacy classes of elements of order 2.

When $m = 2$, it follows from Theorem 11.6 that $\Omega(3,q) \simeq PSp(2,q)$, and in the next chapter we complete the picture by showing that $\Omega(5,q) \simeq PSp(4,q)$.

11.52 Lemma. *If q is odd, $PSp(2m,q)$ has $\lfloor \frac{m}{2} \rfloor + 1$ conjugacy classes of elements of order 2.*

Proof. (cf. Exercises 8.10 and 10.16.) If $t \in Sp(2m,q)$ represents an element of order 2 in $PSp(2m,q)$, then either $t^2 = \mathbf{1}$ or $t^2 = -\mathbf{1}$.

Suppose at first that $t^2 = \mathbf{1}$ and let V be the underlying symplectic geometry with alternating form β. If $v \in V$, then $v = \frac{1}{2}(v + t(v)) + \frac{1}{2}(v - t(v))$ and therefore $V = V_+ \oplus V_-$, where

$$V_\varepsilon := \{\, v \in V \mid t(v) = \varepsilon v \,\}, \qquad \varepsilon = \pm 1.$$

If $u \in V_+$ and $v \in V_-$, then $\beta(u,v) = \beta(t(u),t(v)) = -\beta(u,v)$ and so $\beta(u,v) = 0$. Thus $V = V_+ \perp V_-$ and consequently V_+ and V_- are non-degenerate. In particular, $\dim V_+$ and $\dim V_-$ are even. As t and $-t$ represent the same element of $PSp(2m,q)$, we may suppose that $2 \le \dim V_- \le m$.

Suppose that t' is another element of order 2 in $Sp(2m,q)$ and that $V = V'_+ \perp V'_-$ is the associated decomposition of V, where we have replaced t' by $-t'$, if necessary, to ensure $\dim V'_- \le m$. If $\dim V_- = \dim V'_-$, there is an element $g \in Sp(2m,q)$ such that $g(V_+) = V'_+$ and $g(V_-) = V'_-$. In this case, $gtg^{-1} = t'$. Conversely, if t and t' are conjugate, then $\dim V_- = \dim V'_-$. There are $\lfloor \frac{m}{2} \rfloor$ choices for $\dim V_-$ and therefore $\lfloor \frac{m}{2} \rfloor$ conjugacy classes of the type being considered.

Next we deal with $t \in Sp(2m,q)$ such that $t^2 = -\mathbf{1}$. If 4 divides $q-1$, then there exists $\omega \in \mathbb{F}_q$ such that $\omega^2 = -1$. This time, writing $v = \frac{1}{2}(v - \omega t(v)) + \frac{1}{2}(v + \omega t(v))$ we see that $V = M_+ \oplus M_-$, where

$$M_\varepsilon := \{\, v \in V \mid t(v) = \varepsilon\omega v \,\}, \qquad \varepsilon = \pm 1.$$

For u, $v \in M_+$ we have $\beta(u,v) = \beta(t(u),t(v)) = -\beta(u,v)$ and so $\beta(u,v) = 0$. Thus in this case, M_+ and M_- are maximal totally isotropic subspaces of V. If $t'^2 = -\mathbf{1}$ and $V = M'_+ \oplus M'_-$ is the corresponding decomposition of V, then (by Lemma 7.5) there exists $g \in Sp(2m,q)$ such that $g(M_+) = M'_+$ and $g(M_-) = M'_-$; hence $gtg^{-1} = t'$.

Finally, suppose that 4 does not divide $q-1$. If for some $u \ne 0$, $t(u) = \lambda u$, then, on applying t again, we see that $\lambda^2 = -1$; a contradiction. Thus for all $u \ne 0$, $\dim\langle\, u, t(u)\,\rangle = 2$. If $\langle\, u, t(u)\,\rangle$ is totally isotropic, then (by Lemma 7.5) there exists v such that $\beta(u,v) = 1$ and $\beta(t(u),v) = 0$. Put $w := u + t(v)$ and observe that $\beta(w,t(w)) = -2$. Thus it is always the case that for some $w \in V$, $L_1 := \langle\, w, t(w)\,\rangle$ is a hyperbolic line. If $d := \beta(w,t(w))$ and $e_1 := aw + bt(w)$, then $\beta(e_1,t(e_1)) = a^2d + b^2d$. By Lemma 11.1 we may choose e_1 so that $\beta(e_1,t(e_1)) = 1$. By induction we may write $L_1^\perp = L_2 \perp \ldots \perp L_m$, where $L_i := \langle\, e_i, t(e_i)\,\rangle$ and $\beta(e_i,t(e_i)) = 1$. Once again we see that there is just one conjugacy class of elements t such that $t^2 = -\mathbf{1}$.

In all cases we have shown that $PSp(2m,q)$ has $\lfloor \frac{m}{2} \rfloor + 1$ conjugacy classes of elements of order 2. □

11.53 Lemma. *If q is odd, $\Omega(2m+1,q)$ has m conjugacy classes of elements of order 2.*

Proof. Suppose that $t \in \Omega(2m+1,q)$ and that $t^2 = \mathbf{1}$. Let V be the underlying orthogonal geometry. As in the previous lemma, $V = V_+ \perp V_-$, where $t(v) = v$ for all $v \in V_+$ and $t(v) = -v$ for all $v \in V_-$. The subspaces V_+ and V_- are non-degenerate and as $t \in SO(V)$ and $V_- = [V,t]$, it follows that $\dim V_- = 2k$ is even.

The restriction of t to V_- is $-\mathbf{1}$ and therefore $-\mathbf{1} \in \Omega(V_-)$. The calculation of the spinor norm of $-\mathbf{1}$ in the previous section shows that $\Omega(V_-) = \Omega^\varepsilon(2k,q)$, where ε is determined by the congruence $q^k \equiv \varepsilon \pmod 4$. Thus the Witt index of V_- is uniquely determined. The argument of the previous lemma shows that if $t'^2 = \mathbf{1}$ and $V := V'_+ \perp V'_-$ is the corresponding decomposition of V, then t is conjugate to t' if and only if $\dim V_- = \dim V'_-$. Conversely, for each integer k with $1 \le k \le m$ there exists $t \in \Omega(V)$ such that $\dim[V,t] = 2k$. Thus $\Omega(2m+1,q)$ has m conjugacy classes of elements of order 2. □

11.54 Theorem. *If q is odd and $m > 2$, then $PSp(2m,q)$ and $\Omega(2m+1,q)$ are non-isomorphic simple groups of the same order.* □

Orthogonal BN-pairs

Suppose that V is an orthogonal geometry of Witt index $m > 0$ defined by a quadratic form Q whose polar form is non-degenerate. Theorems 9.1 and 9.8 show that, except when $\dim V = 2m$, the group $O(V)$ has a BN-pair. We also have

11.55 Lemma. *If $\dim V > 2m$, then $\Omega(V)$ is strongly transitive.*

Proof. Suppose that $\mathcal{F}$ and $\mathcal{F}'$ are polar frames and that M and M' are chambers of $\Sigma(\mathcal{F})$ and $\Sigma(\mathcal{F}')$, respectively. By Lemmas 11.26 and 11.27, $\Omega(V)$ is transitive on the ordered pairs (P,Q) such that P and Q are singular and $Q \notin P^\perp$. Thus we may suppose that M and M' have a point P in common and that $\mathcal{F}$ and $\mathcal{F}'$ contain P and Q, where $Q \notin P^\perp$. If $m = 1$, the result is proved, otherwise we apply induction to $(P+Q)^\perp$. □

It is a consequence of this lemma and Theorem 9.8 that, for $\dim V > 2m$, the BN-pair for $O(V)$ restricts to a BN-pair for $\Omega(V)$. As in the case of the symplectic and unitary geometries, the group B is the stabilizer of a maximal flag of totally singular subspaces, the group N is the stabilizer of a polar frame, and the Weyl group $N/B\cap N$ is $\mathbb{Z}_2 \wr S_m$. The group $SO(V)$ contains $\Omega(V)$ and therefore it is strongly transitive, provided $\dim V > 2m$. In this case $SO(V)$ also inherits a BN-pair from $O(V)$.

If V is a hyperbolic line, then $\mathcal{P}(V)$ has just one polar frame. Thus the argument of the lemma just proved shows that the subgroup of $O(V)$ generated by the Siegel transformations is transitive on the polar frames of $\mathcal{P}(V)$, even when $\dim V = 2m$. In the case of $\Omega^+(4,2)$ there are 9 polar frames and a direct calculation shows that $\Omega^+(4,2)$ is transitive on them. Thus in all cases $\Omega(V)$ is transitive on the polar frames. However, in the next section, we show that when $\dim V = 2m$, the group $\Omega(V)$ has two orbits on the chambers of the polar building.

As in the case of the symplectic and unitary groups, it is possible to give explicit transformations $n_1, n_2, \ldots, n_m$ such that the cosets n_1H, n_2H, …, n_mH, generate the Weyl group N/H, where $H := B\cap N$. Suppose the polar frame is

$$\mathcal{F} := \{\, \langle e_i \rangle, \langle f_i \rangle \mid 1 \le i \le m \,\}, \tag{11.56}$$

where (e_1, f_1), …, (e_m, f_m) are mutually orthogonal hyperbolic pairs. We define n_i (for $1 \le i < m$) to be the linear transformation such that

$$\begin{aligned}
&n_i(e_i) := e_{i+1}, \qquad n_i(e_{i+1}) := -e_i,\\
&n_i(f_i) := f_{i+1}, \qquad n_i(f_{i+1}) := -f_i, \quad \text{and}\\
&n_i(v) := v \qquad \text{for all } v \in \langle\, e_i, e_{i+1}, f_i, f_{i+1} \,\rangle^{\perp}.
\end{aligned}$$

Then $\dim[V, n_i] = 4$ (or 2, if the characteristic of the field is 2), and so $n_i \in SO(V)$. A straightforward calculation using (11.31) shows that the spinor norm of n_i is trivial; hence $n_i \in \Omega(V)$. It remains to define n_m.

If $\dim V > 2m$, we may choose $w \in \langle\, e_i, f_i \mid 1 \le i \le m \,\rangle^{\perp}$ such that $Q(w) \ne 0$, and then we may replace Q by a scalar multiple to ensure that $Q(w) = -1$. (This does not change the group.) Put $n_m := t_{\langle w \rangle} t_{\langle e_m - f_m \rangle}$. Then $n_m \in SO(V)$ and the spinor norm of n_m is trivial; i.e., $n_m \in \Omega(V)$.

Note that if Q is non-degenerate but the polar form is degenerate and the field is perfect, then $\dim V = 2m+1$ and it is still the case that the elements n_1H, n_2H,…, n_mH defined above generate the Weyl group. But in this case we have $t_{\langle w \rangle} = \mathbf{1}$.

Suppose that $\dim V = 2m$. In this case the polar building is not thick: each panel of type $\{1, 2, \ldots, m-1\}$ is contained in exactly two chambers. On the other hand, every panel not of type $\{1, 2, \ldots, m-1\}$ is contained in at least three chambers. As described on p. 85, the Weyl group of $O(V)$ is generated by elements n_1H, n_2H, …, n_mH, and the proof of Theorem 9.8 shows that $n_iBn_i \ne B$ except for $i = m$, in which case $n_mBn_m = B$.

If $\Omega(V)$ were strongly transitive, then the stabilizer of the frame $\mathcal{F}$ would contain an element n_m inducing the transposition $(\langle e_m \rangle, \langle f_m \rangle)$. But for any such element, $\dim[V, n_m] = 1$ and so $n_m \notin SO(V)$. Thus neither $\Omega^+(2m, \mathbb{F})$ nor $SO^+(2m, \mathbb{F})$ can be strongly transitive.

On the other hand, it is clear that $n'_m := n_m n_{m-1} n_m^{-1} \in \Omega(V)$ and the cosets $n_1H, \ldots, n_{m-1}H, n'_mH$ generate the subgroup $(\mathbb{Z}_2 \wr S_m)^+$ of even permutations in $\mathbb{Z}_2 \wr S_m$. We know from Lemma 9.1 that $\mathbb{Z}_2 \wr S_m$ acts regularly on the $2^m\, m!$ chambers of $\Sigma(\mathcal{F})$. But $(\mathbb{Z}_2 \wr S_m)^+$ has index 2 in $\mathbb{Z}_2 \wr S_m$ and therefore $(\mathbb{Z}_2 \wr S_m)^+$ has two orbits on these chambers. As $\Omega(V)$ is transitive on polar frames, it follows that for $k < m$, $\Omega(V)$ is transitive on the totally singular subspaces of dimension k.

Even though the groups $SO^+(2m, \mathbb{F})$ and $\Omega^+(2m, \mathbb{F})$ do not act strongly transitively on the polar building, it is still the case that the stabilizer of a chamber and of an apartment (containing the chamber) form a BN-pair for these groups. The Weyl group is $(\mathbb{Z}_2 \wr S_m)^+$. Before proving this we look more closely at the action of orthogonal groups on the maximal totally singular subspaces.

Maximal Totally Singular Subspaces

As in the previous section, suppose that V is an orthogonal geometry of dimension n and Witt index $m > 0$ defined by a quadratic form Q whose polar form β is non-degenerate. Let Φ be the set of all maximal totally singular subspaces of V. Make Φ into a graph by joining E to F by an edge whenever $\dim(E \cap F) = m - 1$. The *distance* $d(E, F)$ from E to F is the length of a shortest path from E to F.

11.57 Lemma. *(i) For all E, $F \in \Phi$, if E is adjacent to F, then $t(E) = F$ for some reflection $t := t_{\langle u \rangle}$, where $Q(u) = 1$.*

(ii) If $E \in \Phi$ and if t is a reflection, then $t(E) = E$, or E is adjacent to $t(E)$.

Proof. (i) Choose $e \in E \setminus E \cap F$ and $f \in F \setminus E \cap F$ such that $\beta(e, f) = 1$. Then $Q(e + f) = \beta(e, f) = 1$ and $t_{\langle e+f \rangle}(E) = F$.

(ii) The reflection t fixes every vector in a hyperplane of V and therefore $\dim(E \cap t(E)) \geq m - 1$. □

11.58 Lemma. *For all E, $F \in \Phi$, $d(E, F) = m - \dim(E \cap F)$.*

Proof. Suppose that $E \neq F$ and choose $e \in F \setminus E \cap F$. Then $E' := \langle e \rangle + E \cap \langle e \rangle^\perp \in \Phi$ and E' is adjacent to E. Since $\dim(E' \cap F) = \dim(E \cap F) + 1$, we can continue in this fashion and construct a path of length $m - \dim(E \cap F)$ from E to F. Hence $d(E, F) \leq m - \dim(E \cap F)$, and in particular, Φ is connected.

Now suppose that E, F and $E' \in \Phi$, where E' is adjacent to E. Then $E \cap E'$ is a hyperplane of E', and therefore

$$\dim(E \cap F) \geq \dim(E \cap E' \cap F) \geq \dim(E' \cap F) - 1.$$

It follows by induction that $\dim(E \cap F) \geq m - d(E,F)$, and by the previous paragraph, $d(E,F) = m - \dim(E \cap F)$. □

11.59 Theorem. *The group $O(V)$ is transitive on Φ and its orbits on $\Phi \times \Phi$ are the sets*

$$\Delta_k := \{\, (E,F) \in \Phi \times \Phi \mid d(E,F) = k \,\} \qquad (0 \leq k \leq m).$$

Proof. Suppose that (E_1, F_1) and (E_2, F_2) belong to Δ_k. By Witt's theorem we may assume that $E_1 \cap F_1 = E_2 \cap F_2 = W$, say. Now write $E_1 = E_1' \oplus W$, $E_2 = E_2' \oplus W$, $F_1 = F_1' \oplus W$, and $F_2 = F_2' \oplus W$. By Lemma 7.5 there is a basis $e_1, e_2, \ldots, e_k$ for E_1' and a basis $f_1, f_2, \ldots, f_k$ for F_1' such that (e_1, f_1), (e_2, f_2), $\ldots$, (e_k, f_k) are mutually orthogonal hyperbolic pairs. The same is true of E_2' and F_2', hence by Witt's theorem there exists $g \in O(V)$ that fixes every vector of W, takes E_1' to E_1, and takes F_1' to F_2. Then $g(E_1, F_1) = (E_2, F_2)$. □

A graph is said to be *bipartite* if it can be written as the disjoint union of two non-empty subsets (called the *parts* of the *bipartition*) such that the edges of the graph only join vertices in distinct subsets; equivalently, the graph has no circuits of odd length.

11.60 Theorem. *Suppose that V is an orthogonal geometry of dimension n and Witt index $m > 0$. Let Φ be the graph of maximal totally singular subspaces.*

(*i*) *If $n \neq 2m$, then the action of $O(V)$ on Φ is primitive.*

(*ii*) *If $n = 2m$, then Φ is bipartite; hence if $m > 1$, the action of $O(V)$ on Φ is imprimitive.*

Proof. Suppose that $\Gamma \subseteq \Phi \times \Phi$ is an $O(V)$-invariant equivalence relation such that $\Gamma \neq \Delta_0$. If $\Delta_k \cap \Gamma \neq \emptyset$, then $\Delta_k \subseteq \Gamma$, and if $\Delta_1 \subseteq \Gamma$, then $\Gamma = \Phi \times \Phi$. So suppose that, for some $k > 1$, $\Delta_k \subseteq \Gamma$, but $\Delta_{k-1} \not\subseteq \Gamma$. Choose $(E,F) \in \Delta_{k-1}$ and let H be a hyperplane of $E \cap F$. Let H_1 be a hyperplane of F containing H and choose $F_1 \in \Phi$ so that $F_1 \cap F = H_1$. If $e \in E \cap F_1^\perp$, then $e \in F^\perp$ and as F is a maximal totally singular subspace, $e \in F$. Thus $E \cap F_1 = H$ and so $(E, F_1) \in \Delta_k$. If $H_2 \neq H_1$ is another hyperplane of F such that $E \cap H_2 = H$ and if $F_2 \neq F$ is an element of Φ that contains H_2, then $(E, F_2) \in \Delta_k$ and $(F_1, F_2) \in \Delta_2$. It follows that $\Delta_2 \subseteq \Gamma$ and therefore

$$\bigcup_{i \geq 0} \Delta_{2i} \subseteq \Gamma.$$

If $\Delta_\ell \subseteq \Gamma$ for some odd ℓ, then $\Delta_1 \subset \Gamma$ and hence $\Gamma = \Phi \times \Phi$. If $n \neq 2m$ and H is a totally isotropic subspace of dimension $m-1$, then $H^\perp/H$ contains at least three singular points and hence the graph Φ contains circuits of length three. Thus in this case we have $\Gamma = \Phi \times \Phi$ and therefore $O(V)$ is primitive.

Now suppose that $n = 2m$. We shall show that Φ does not have circuits of odd length. Suppose, on the contrary, that E, F_1, $F_2 \in \Phi$ and that (E, F_1), $(E, F_2) \in \Delta_k$ and $(F_1, F_2) \in \Delta_1$. Then $E \cap F_1 = E \cap F_2$, and if $H := F_1 \cap F_2$, the subspaces $E \cap H^\perp$, F_1 and F_2 correspond to three distinct singular points of $H^\perp/H$. This is a contradiction as $H^\perp/H$ is a hyperbolic line and has only two singular points. Thus Φ has no circuits of odd length and hence it is bipartite. □

11.61 Theorem. *If V is an orthogonal geometry of Witt index m and dimension $2m$, then both $SO(V)$ and $\Omega(V)$ have two orbits on the set Φ of maximal totally singular subspaces: two subspaces E and F are in the same orbit if and only if $d(E, F)$ is even. Furthermore, both $SO(V)$ and $\Omega(V)$ have two orbits on the chambers of the polar building.*

Proof. If t is a reflection and $E \in \Phi$, then $t(E) \neq E$. By Lemma 11.57 (*ii*), E is adjacent to $t(E)$. Thus t interchanges the two parts of the bipartite graph Φ. Except for $SO^+(4,2)$, every element of $SO(V)$ is a product of an even number of reflections; hence $SO(V)$ fixes the two parts of the bipartition of Φ. It follows from Lemma 11.57 (*i*) that $\Omega(V)$ (and hence $SO(V)$) has two orbits on Φ. The corresponding statements for $\Omega^+(4,2)$ and $SO^+(4,2)$ are easy exercises.

It is now clear that $SO(V)$ and $\Omega(V)$ have at least two orbits on the chambers of the polar building. On the other hand, any two chambers belong to a common apartment and we saw in the previous section that the stabilizer of an apartment has two orbits on its chambers. This completes the proof.

□

A detailed description of the stabilizer of a maximal totally singular subspace is given in Exercise 11.19.

The Oriflamme Geometry

In this section V is an orthogonal geometry of Witt index m and dimension $2m$. For each chamber M of the polar building of V, there is a unique chamber M' such that $M \cap M'$ is the panel obtained from M by omitting the subspace of dimension m. The subspaces of dimension m in M and M' are adjacent in the graph of maximal totally singular subspaces and the theorems of the previous section show that M and M' are in different orbits of $\Omega(V)$, but $\Omega(V)$ is transitive on the set of pairs $\{M, M'\}$.

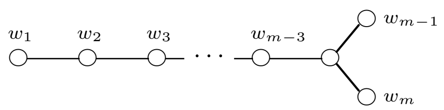
Let $\mathcal{F}$ be the polar frame (11.56) and let M be the chamber

$$\{\,\langle e_1, e_2, \ldots, e_i\rangle \mid 1 \le i \le m\,\}$$

in the apartment Σ of $\mathcal{F}$. Then M' is also in Σ and the stabilizer B of M fixes M'. In the section 'Orthogonal BN-pairs' we showed that the stabilizer N of $\mathcal{F}$ has two orbits on the chambers of Σ and it follows that B is transitive on the apartments which contain M. This is all that is required for the proofs of Theorems 9.3 and 9.6 to hold, where for the generators of the Weyl group we use the elements $w_i := n_i B \cap N$, for $1 \le i \le m-1$, and $w_m := n'_m B \cap N$, described on pages 169 and 170.

The proof of Theorem 9.8 also depends on knowing that each panel of the form $M \cap n_i(M)$ is in at least three chambers. This has been proved in Lemma 9.4 for n_1, n_2, …, n_{m-1}. The panel $M' \cap n'_m(M')$ is the same as $M \cap n_{m-1}(M)$ and therefore the proof of Theorem 9.8 goes through for n'_m with M replaced by M'. Thus the groups B and N form a BN-pair for $\Omega(V)$ and the Weyl group is $(\mathbb{Z}_2 \wr S_m)^+$. Its Coxeter-Dynkin diagram (said to be of type D_m) is

D_m



where the vertex at the fork corresponds to w_{m-2}. The same construction provides a BN-pair for $SO(V)$.

In Chapter 9, in the section 'Diagram Geometries', we indicated that the Coxeter-Dynkin diagrams of type A_n and C_n describe the incidence relations between the varieties of projective geometry and polar geometry, respectively.

The geometry corresponding to the Coxeter-Dynkin diagram of type D_n is known as the *oriflamme geometry.* The varieties of type i (for $1 \le i \le m-2$) are the totally isotropic subspaces of dimension i. In the bipartite graph Φ, introduced in the previous section, let Φ_1 and Φ_2 be the two parts of the bipartition. The elements of Φ_1 are the varieties of type $m-1$ and the elements of Φ_2 are the varieties of type m. Two varieties are adjacent if one is contained in the other or if they are both of dimension m and adjacent in the graph Φ (i.e., their intersection has dimension $m-1$). The *oriflamme building* is the set of oriflammes of this geometry, where an *oriflamme* is a set of mutually adjacent varieties. A *chamber* is a maximal oriflamme.

If M and M' are chambers of the polar building that differ only in their subspaces E and E' of dimension m, then their subspaces of dimensions $1, 2, \dots, m-2$ together with E and E' form a chamber of the oriflamme building. Conversely, every chamber of the oriflamme building arises from a unique pair of polar chambers of this form.

For each polar frame $\mathcal{F}$, there is an apartment in the oriflamme building. It consists of the oriflammes whose subspaces are spanned by subsets of $\mathcal{F}$.

EXERCISES

11.1 Let V be a vector space over a field $\mathbb{F}$ of characteristic 2 and suppose that β is a non-degenerate symmetric bilinear form on V. Let G be the group of linear transformations of V that preserve β.

  (*i*) If $W := \{ v \in V \mid \beta(v,v) = 0 \}$, show that G stabilizes the flag $\operatorname{rad} W \subset W$, and that G acts as the identity on $\operatorname{rad} W$ and V/W.

  (*ii*) Show that the map $G \to Sp(W/\operatorname{rad} W)$, which assigns $f \in G$ to the transformation of $W/\operatorname{rad} W$ induced by f, is a homomorphism onto $Sp(W/\operatorname{rad} W)$ whose kernel is a nilpotent group of class at most 2.

  (*iii*) If $\mathbb{F}$ is perfect, show that W is a hyperplane of V, and $\dim \operatorname{rad} W$ is 0 or 1 according to whether $\dim V$ is odd or even.

11.2 Following Theorem 11.4 describe $O(L)$, where L is an orthogonal hyperbolic line over an arbitrary field $\mathbb{F}$. Show that $SO(L)$ is isomorphic to the multiplicative group of the field.

11.3 Let V be an orthogonal geometry of dimension n over $\mathbb{F}_q$ and let $m > 0$ be the Witt index of V. Show that the number of totally singular subspaces of dimension k in V is

$$\begin{bmatrix} m \\ k \end{bmatrix}_q \prod_{i=0}^{k-1} (q^{m-\varepsilon-i} + 1)$$

where $\varepsilon = 2m - n + 1$ and

$$\begin{bmatrix} m \\ k \end{bmatrix}_q = \prod_{i=0}^{k-1} (q^m - q^i)/(q^k - q^i).$$

Observe that the same expression gives the number of totally isotropic subspaces of dimension k for the geometries associated with the groups $Sp(2m, q)$, $U(2m+1, q^{1/2})$ and $U(2m, q^{1/2})$ provided we take ε to be 0, $-\frac{1}{2}$ and $\frac{1}{2}$, respectively.

11.4 Let E be a maximal totally singular subspace of an orthogonal geometry of Witt index $m > 0$. Defining ε as in the previous exercise, show that the number of maximal totally singular subspaces F such that $E \cap F = \{0\}$ is $q^{\frac{1}{2}m(m+1)-m\varepsilon}$.

11.5 In the proof Theorem 11.6 show that every element of $SO(V)_{\langle e \rangle}$ can be written as a product $H(a)S(b)$ for suitable a and b.

11.6 Let V be an orthogonal geometry over a perfect field $\mathbb{F}$ defined by a non-degenerate quadratic form Q whose polar form is *degenerate*. Show that the orthogonal building of $\mathcal{P}(V)$ is isomorphic to the symplectic building of $\mathcal{P}(V/V^\perp)$.

11.7 In case II in the section on root groups check that $\widehat{X}_{u,v}$ depends only on the totally singular subspace $\langle u, v \rangle$ and not on the particular choice of basis u, v.

11.8 Write out the elements of $\Omega^+(4,2)$ as permutations on the six non-singular vectors of the geometry and identify the Siegel transformations.

11.9 Find two regular elements in $O^+(4,3)$ which commute but whose product is not regular.

11.10 Let Ω be a set of size $2m+2$ and let V be the vector space of partitions $\{\Gamma, \Delta\}$ of Ω into pairs of even subsets; addition for V is symmetric difference. This is the vector space introduced in Theorem 8.9. Suppose that m is odd, and define a quadratic form $Q : V \to \mathbb{F}_2$ by

$$Q(\{\Gamma, \Delta\}) := \frac{1}{2}|\Gamma| \pmod 2.$$

- (i) Show that the polar form of Q is the alternating form β defined in the proof of Theorem 8.9.
- (ii) If $m \equiv 1 \pmod 4$, show that $S_{2m+2} \subseteq O^-(2m,2)$.
- (iii) If $m \equiv 3 \pmod 4$, show that $S_{2m+2} \subseteq O^+(2m,2)$.
- (iv) Deduce that $O^+(6,2) \simeq S_8$ and that $\Omega^+(6,2) \simeq PSL(4,2)$.
- (v) Use a similar argument to show that $O^-(4,2) \simeq S_5$

11.11 Show that under the isomorphism $O^+(6,2) \simeq S_8$, the reflections of $O^+(6,2)$ correspond to the transpositions of S_8 and the elements f such that $[V,f]$ is totally singular correspond to the products of 4 commuting transpositions. Deduce that if f corresponds to a conjugate of $(1\ 2\ 3\ 4\ 5\ 6\ 7\ 8)$, $(1\ 2\ 3\ 4\ 5\ 6)(7\ 8)$, $(1\ 2\ 3\ 4)(5\ 6\ 7\ 8)$ or $(1\ 2\ 3\ 4)(5\ 6)(7\ 8)$, then $[V,f]$ is not totally singular and f is a product of $\dim[V,f]+2$ reflections but no fewer.

11.12 Let (W,R) be a Coxeter system and let V be the orthogonal geometry defined in Exercise 9.8. For $w \in W$, put

$$A(w) := \{\, d \in W \mid [V,d] \subseteq [V,w] \,\}.$$

Show that

(i) $A(w)$ is a subgroup of W.

(ii) If $w' \in A(w)$, then $A(w') \subseteq A(w)$.

(iii) If $x \in W$, then $xA(w)x^{-1} = A(xwx^{-1})$.

(iv) For all $w \in W$, $\dim[V,w]$ is the shortest length of any expression for w as a product of reflections. (See Steinberg (1967) or Steinberg (1968).)

(v) If $J \subseteq R$ and c_J is obtained by taking the product of the elements of J in some order, then $[V,c_J] = \langle\, e_r \mid r \in J \,\rangle$ and $A(c_J) = W_J$.

11.13 (i) Show that except for $SO^+(4,2)$, the group $SO(V)$ is generated by the products t_1t_2, where t_1 and t_2 are reflections.

(ii) Show that for all V, $\Omega(V)$ is generated by the elements f^2, where $f \in O(V)$.

11.14 Let P be a non-singular point of an orthogonal geometry V of dimension at least 3 over the field $\mathbb{F}_q$, defined by a quadratic form whose polar form is non-degenerate. If t is the reflection in $P^\perp$, show that, when q is even, $C_{O(V)}(t)$ is transitive on the non-singular points of $P^\perp \setminus \{P\}$ and that $C_{O(V)}(t)/\langle\, t \,\rangle$ is isomorphic to $Sp(P^\perp/P)$. When q is odd, show that $C_{O(V)}(t) = \langle\, t \,\rangle \times O(P^\perp)$. If $\dim V$ is odd, show that the Witt index of $P^\perp$ depends on the choice of P.

11.15 Let V be a vector space of dimension 4 and Witt index 2 over a field $\mathbb{F}$. Show that the action of $O^+(4,\mathbb{F})$ on the singular points of V is primitive.

11.16 Show that $SU(2m,q)$ is contained in $\Omega^+(4m,q)$ and that $SU(2m{+}1,q)$ is contained in $\Omega^-(4m+2,q)$.

11.17 Let β be a non-degenerate alternating form on the vector space V of dimension $2m$ over the field $\mathbb{F}_2$. Let $\mathcal{Q}$ be the set of all quadratic forms that polarize to β.

(i) Define an action of V on $\mathcal{Q}$ as follows. For $v \in V$ and $Q \in \mathcal{Q}$, put

$$v \cdot Q := Q + \beta(v, -).$$

Show that V acts regularly on $\mathcal{Q}$.

(ii) For $f \in Sp(V)$ and $Q \in \mathcal{Q}$, define $f \cdot Q$ by

$$(f \cdot Q)(v) := Q(f^{-1}v).$$

Show that the actions of V and $Sp(V)$ on $\mathcal{Q}$ extend to an action of the semidirect product $V\,Sp(V)$ on $\mathcal{Q}$ and that $V\,Sp(V)$ acts doubly transitively on V and $\mathcal{Q}$.

(iii) Show that $Sp(V)$ has two orbits on $\mathcal{Q}$: the forms of Witt index m and the forms of Witt index $m-1$.

(iii) Show that the orbits of $Sp(V)$ on $\mathcal{Q}$ have lengths $2^{2m-1} + 2^{m-1}$ and $2^{2m-1} - 2^{m-1}$, and that $Sp(V)$ acts doubly transitively on both of them. (Hint. Show that $Sp(V)$ has 6 orbits on $\mathcal{Q} \times \mathcal{Q}$: use the fact that a group acting on a block design with the same number of points as blocks has the same number of orbits on points as on blocks. See also Jordan (1870) and Taylor (1977).)

11.18 Let V be an orthogonal geometry of Witt index $m > 0$ and dimension at least $2m+1$. Show that $\Omega(V)$ acts primitively on the set of maximal totally singular subspaces.

11.19 Let E be a vector space of dimension m over $\mathbb{F}$.

(i) Set $V := E^* \oplus E$, where E^* is the dual space of E, and define a quadratic form Q on E by

$$Q((\varphi, v)) := \varphi(v).$$

Show that Q is non-degenerate and that its polar form β is given by

$$\beta((\varphi, v), (\psi, w)) := \varphi(w) + \psi(v).$$

(ii) Show that every orthogonal geometry of Witt index m and dimension $2m$ over $\mathbb{F}$ is isometrically isomorphic to V.

(*iii*) Let $\Lambda^2(E)$ be the set of all (possibly degenerate) alternating forms on E. For $\gamma_1, \gamma_2 \in \Lambda^2 E$ and $a_1, a_2 \in \mathbb{F}$, define

$$(a_1\gamma_1 + a_2\gamma_2)(u, v) := a_1\gamma_1(u, v) + a_2\gamma_2(u, v)$$

and show that $\Lambda^2 E$ is a vector space of dimension $\frac{1}{2}m(m-1)$ over $\mathbb{F}$.

(*iv*) Identify E^* with the subspace $\{ (\varphi, 0) \mid \varphi \in E^* \}$ and identify E with the subspace $\{ (0, v) \mid v \in E \}$ of V. If $f \in O(V)$ fixes every element of E^*, show that for all $v \in E$, $(1-f)(v) \in E^*$ and that

$$(1-f)(v) = \gamma(-, v),$$

for some $\gamma \in \Lambda^2 E$. Conversely, for $\gamma \in \Lambda^2 E$ show that the linear transformation $\tilde{\gamma}$ defined by

$$\tilde{\gamma}(\varphi, v) := (\varphi + \gamma(v, -), v)$$

preserves Q and fixes every vector of E^*. Deduce that the subgroup $O(V)(E^*)$ of $O(V)$ which fixes every element of E^* is isomorphic to $\Lambda^2 E$. (cf. Exercise 8.2.)

(*v*) For $\gamma \in \Lambda^2 E$, show that $[V, \tilde{\gamma}]$ is isomorphic to $E/\operatorname{rad} E$, where $\operatorname{rad} E$ denotes the radical of E with respect to γ. Using the fact that the dimension of a symplectic geometry is even, show that the Dickson invariant of $\tilde{\gamma}$ is 0 and hence $\tilde{\gamma} \in SO(V)$.

(*vi*) Let χ be the Wall form of $\tilde{\gamma}$ and show that

$$\chi(\gamma(-, u), \gamma(-, v)) = \gamma(u, v).$$

By calculating determinants with respect to a symplectic basis, show that the discriminant of χ is a square and hence $\tilde{\gamma} \in \Omega(V)$.

(*vii*) Show that $O(V)(E^*)$ acts regularly on the set of maximal totally singular subspaces F of V such that $E^* \cap F = \{0\}$.

(*viii*) For $f \in GL(E)$ and $\gamma \in \Lambda^2 E$, define $f\gamma$ by

$$(f\gamma)(u, v) := \gamma(f^{-1}(u), f^{-1}(v)).$$

Show that in this way $GL(E)$ may be regarded as a group of linear transformations of $\Lambda^2 E$.

(*ix*) For $f \in GL(E)$, let $\bar{f}$ be the element of $GL(E^*)$ that takes $\varphi \in E^*$ to φf^{-1}. Show that every element of the subgroup $O(V)_{E^*,E}$ fixing both E^* and E can be written in the form

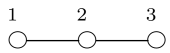
$$(\varphi, v) \mapsto (\bar{f}(\varphi), f(v)) \qquad \text{for some } f \in GL(E),$$

and hence $O(V)_{E^*,E} \subseteq SO(V)$. Deduce that the stabilizer $O(V)_{E^*}$ of E^* in $O(V)$ is a subgroup of $SO(V)$ and that it is isomorphic to the semidirect product $(\Lambda^2 E)\, GL(E)$.

12

The Klein Correspondence

An inspection of the Coxeter-Dynkin diagrams associated with the orthogonal groups and geometries defined on vector spaces of dimension at most 6 shows that in every case the diagram has occurred before in connection with a projective, symplectic, or unitary geometry. For example, at the end of the previous chapter it was shown that when V is an orthogonal geometry of dimension 6 and Witt index 3 over a field $\mathbb{F}$, the Coxeter-Dynkin diagram for the oriflamme geometry of V is



The end nodes represent the two classes of totally singular planes of $\mathcal{P}(V)$ and the middle node represents the singular points. This is also the diagram of type A_3 and suggests that the oriflamme geometry is none other than a projective geometry of (projective) dimension 3. Indeed this is the case. The varieties of types 1, 2 and 3 correspond to points, lines and planes, and it is easy to check directly that the axioms for projective geometry (given in Chapter 3) are satisfied. For example, given a variety M of type 3 (i.e., a totally singular plane), the varieties of type 1 adjacent to M in the oriflamme geometry are the totally singular planes M' such that $M \cap M'$ is a line, and the varieties of type 2 adjacent to M are the points of M. Thus the residue of M is the dual of the projective plane $\mathcal{P}(M)$.

This correspondence between the orthogonal geometry of a 5-dimensional projective space and the geometry of a 3-dimensional projective space was first studied (over the complex numbers) by Felix Klein in his dissertation of 1868 (reprinted in his collected works: Klein (1921)). In its general form it is the basis for various isomorphisms (hinted at on p. 166) between orthogonal groups in dimensions 4, 5 and 6, and other linear groups. (See van der Waerden (1935), I, §7 and Dieudonné (1971), Chap. III, §8.)

The isomorphism between the geometries of type A_3 and D_3 can be derived from a more general correspondence between the geometry of a vector space and its exterior square. Essentially, this is the approach taken by Klein in his dissertation and subsequent papers on line geometry. It is also the approach that we take in this chapter, and therefore we devote the next few sections to a review of the fundamentals of exterior algebra. A more detailed account can be found in Bourbaki (1970), Chap. III, §§7 and 11.

The Exterior Algebra of a Vector Space

Given a vector space V of dimension n over the field $\mathbb{F}$, the *exterior square* of V is a vector space $\Lambda_2 V$ (also over $\mathbb{F}$) together with a bilinear map

$$\alpha_2 : V \times V \to \Lambda_2 V$$

such that $\alpha_2(v, v) = 0$ for all $v \in V$, and such that for every alternating form $\gamma : V \times V \to \mathbb{F}$, there is a unique linear functional $\widehat{\gamma} : \Lambda_2 V \to \mathbb{F}$ satisfying $\widehat{\gamma}\alpha_2 = \gamma$; or, as is often said, such that the following diagram commutes:

$$\begin{tikzcd} V \times V \arrow[r, "\alpha_2"] \arrow[dr, "\gamma"'] & \Lambda_2 V \arrow[d, "\widehat{\gamma}"] \\ & \mathbb{F} \end{tikzcd}$$

The correspondence between γ and $\widehat{\gamma}$ is one-to-one and thus the dual of $\Lambda_2 V$ may be identified with the space $\Lambda^2 V$ of all alternating forms on V (introduced in Exercise 11.19). However, it is easy to construct $\Lambda_2 V$ directly.

If $e_1, e_2, \ldots, e_n$ is a basis for V, then $\Lambda_2 V$ can be defined to be the vector space over $\mathbb{F}$ with the $\binom{n}{2}$ symbols $e_i \wedge e_j$ $(1 \le i < j \le n)$ as a basis. For $1 \le i < j \le n$, we define

$$\begin{aligned} e_i \wedge e_i &:= 0, \qquad \text{and} \\ e_j \wedge e_i &:= -e_i \wedge e_j. \end{aligned}$$

Then for $u := \sum_{i=1}^{n} a_i e_i$ and $v := \sum_{i=1}^{n} b_i e_i$, we put

$$u \wedge v := \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j e_i \wedge e_j.$$

The transformation $\alpha_2 : V \times V \to \Lambda_2 V$ defined by $\alpha_2(u, v) := u \wedge v$ is bilinear and satisfies $\alpha_2(v, v) = 0$ for all $v \in V$.

If γ is an alternating form on V, the linear functional $\widehat{\gamma} : \Lambda_2 V \to \mathbb{F}$ is defined on the basis element $e_i \wedge e_j$ by

$$\widehat{\gamma}(e_i \wedge e_j) := \gamma(e_i, e_j).$$

It then follows that $\widehat{\gamma}(u \wedge v) = \gamma(u, v)$ for all $u, v \in V$. By construction, $\dim \Lambda_2 V = \frac{1}{2} n(n-1)$, and in particular, $\dim \Lambda_2 V = 6$, when $\dim V = 4$.

The notation $u \wedge v$ is meant to suggest that V and $\Lambda_2 V$ are subspaces of a larger algebra in which $\wedge$ is the multiplication: this is the *exterior algebra* $\Lambda_\bullet V$ of V. The construction of $\Lambda_\bullet V$ found in Bourbaki (1970), Chap. III, §7 is typical of modern treatments and makes use of the tensor algebra of V. However, just as for $\Lambda_2 V$, a direct construction is possible in terms of an explicit basis. This approach is a special case of the construction of a Clifford algebra given by Artin (1957), p. 186.

For each subset S of $I := \{1, 2, \dots, n\}$ we introduce a symbol e_S and we let $\Lambda_\bullet V$ be the vector space of dimension 2^n over $\mathbb{F}$ with basis $\{\, e_S \mid S \subseteq I \,\}$. If S and T are subsets of I, we define

$$\varepsilon(S,T) := \begin{cases} (-1)^{k(S,T)} & S \cap T = \emptyset, \\ 0 & S \cap T \neq \emptyset, \end{cases}$$

where $k(S,T) := |\{\, (i,j) \mid i \in S, j \in T \text{ and } i > j \,\}|$. Then we put

$$e_S \wedge e_T := \varepsilon(S,T) e_{S \cup T} \tag{12.1}$$

and extend this product to all of $\Lambda_\bullet V$ by requiring it to be distributive. That is, if $\xi := \sum_{S \subseteq I} a_S e_S$ and $\eta := \sum_{S \subseteq I} b_S e_S$, the *exterior product* of ξ and η is

$$\xi \wedge \eta := \sum_{S \subseteq I} \sum_{T \subseteq I} a_S b_T e_S \wedge e_T. \tag{12.2}$$

There is a linear transformation $\alpha : V \to \Lambda_\bullet V$ such that $\alpha(e_i) := e_{\{i\}}$ for all $i \in I$, and we have $\alpha(v) \wedge \alpha(v) = 0$ for all $v \in V$. This allows us to identify V with its image in $\Lambda_\bullet V$ and from now on we write e_i instead of $e_{\{i\}}$. The basis element $e_\emptyset$ is the identity element of $\Lambda_\bullet V$ and we denote it by 1.

12.3 Theorem. *The vector space $\Lambda_\bullet V$ is an associative algebra over $\mathbb{F}$ and for every associative algebra A and every linear transformation $h : V \to A$ such that $h(v)^2 = 0$ for all $v \in V$, there exists a unique algebra homomorphism $\widehat{h} : \Lambda_\bullet V \to A$ such that the following diagram commutes.*

$$\begin{tikzcd} V \arrow[r, "\alpha"] \arrow[dr, "h"'] & \Lambda_\bullet V \arrow[d, "\widehat{h}"] \\ & A \end{tikzcd}$$

Moreover (up to isomorphism) $\Lambda_\bullet V$ is the unique associative algebra with this property.

Proof. The distributive law for $\Lambda_\bullet V$ follows immediately from (12.2) and so, in order to show that $\Lambda_\bullet V$ is an associative algebra, all that really needs proving is the associative law. For the basis elements we have

$$\begin{aligned}(e_R \wedge e_S) \wedge e_T &= \varepsilon(R,S)\varepsilon(R \cup S, T)e_{R\cup S\cup T}\\ &= \varepsilon(R,S)\varepsilon(R,T)\varepsilon(S,T)e_{R\cup S\cup T}\\ &= e_R \wedge (e_S \wedge e_T).\end{aligned}$$

The associative law for $\Lambda_\bullet V$ now follows from (12.2).

If $h : V \to A$ is a linear transformation such that $h(v)^2 = 0$ for all $v \in V$, then replacing v by $u + v$ shows that

$$h(u)h(v) = -h(v)h(u).$$

For $S := \{i_1, i_2, \dots, i_k\}$, where $1 \le i_1 < i_2 < \cdots < i_k \le n$, define

$$\widehat{h}(e_S) := h(e_{i_1})h(e_{i_2})\cdots h(e_{i_k})$$

and extend $\widehat{h}$ to $\Lambda_\bullet V$ by linearity. It follows from (12.1) that $\widehat{h}(e_S e_T) = \widehat{h}(e_S)\widehat{h}(e_T)$ and therefore $\widehat{h}$ is a homomorphism.

The uniqueness of $\Lambda_\bullet V$ is Exercise 12.2. □

If we put $S := \{i_1, i_2, \dots, i_k\}$, where $1 \le i_1 < i_2 < \cdots < i_k \le n$, then $e_S = e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}$. Thus we may identify $\Lambda_2 V$ with the subspace spanned by the basis vectors e_S, where $|S| = 2$. More generally, we define the k-th *exterior power* of V to be the subspace $\Lambda_k V$ spanned by the vectors e_S, where $|S| = k$. The elements of $\Lambda_k V$ are called *k-vectors* and, in particular, the elements of $\Lambda_2 V$ are called *bivectors*.

It follows directly from the definitions that $\dim \Lambda_k V = \binom{n}{k}$. Thus both $\Lambda_0 V$ and $\Lambda_n V$ are 1-dimensional. We identify $\Lambda_0 V$ with the field $\mathbb{F}$.

The multiplication defined by (12.1) and (12.2) is anticommutative in the sense that for all $\xi \in \Lambda_k V$ and all $\eta \in \Lambda_\ell V$ we have

$$\xi \wedge \eta = (-1)^{k\ell} \eta \wedge \xi.$$

If V and W are vector spaces, and if $f : V \to W$ is a linear transformation, then by Theorem 12.3 there is a unique algebra homomorphism $\Lambda_\bullet f : \Lambda_\bullet V \to \Lambda_\bullet W$ such that $(\Lambda_\bullet f)(v) = f(v)$ for all $v \in V$ (where we have identified v with its image in $\Lambda_\bullet V$ and $f(v)$ with its image in $\Lambda_\bullet W$). Furthermore, if $g : W \to X$ is another linear transformation, then $\Lambda_\bullet(gf) = \Lambda_\bullet(g)\Lambda_\bullet(f)$. In particular, taking $V = X$ and taking f to be invertible, it follows that $\Lambda_\bullet f$ is invertible and its inverse is $\Lambda_\bullet f^{-1}$.

For $0 \le k \le n$, the algebra homomorphism $\Lambda_\bullet f$ restricts to a linear transformation $\Lambda_k f : \Lambda_k V \to \Lambda_k W$. The map $\Lambda_0 f$ is the identity and $\Lambda_1 f$

coincides with f. The space $\Lambda_n V$ is spanned by $e_I = e_1 \wedge e_2 \wedge \cdots \wedge e_n$ and if $f : V \to V$, it follows readily from the definitions that

$$(\Lambda_n f)e_I = \det(f)\, e_I. \tag{12.4}$$

More generally, if $f : V \to W$ is σ-semilinear, then for each k, f induces a σ-semilinear transformation $\Lambda_k f : \Lambda_k V \to \Lambda_k W$ such that for all vectors v_1, v_2, …, $v_k \in V$,

$$(\Lambda_k f)(v_1 \wedge v_2 \wedge \cdots \wedge v_k) = f(v_1) \wedge f(v_2) \wedge \cdots \wedge f(v_k).$$

The map $\Lambda_0 f$ is the field automorphism σ and when $W = V$, we have $(\Lambda_n f)e_I = d\, e_I$, where d is the determinant of the matrix of f with respect to the basis e_1, e_2, …, e_n of V. But if $\sigma \neq 1$, then $\Lambda_n f$ does *not* act on $\Lambda_n V$ as multiplication by d. In fact, if $a \neq 0$ and $e'_I := a e_I$, then $(\Lambda_n f)e'_I = a^{-1}\sigma(a) d\, e'_I$.

The Dual Space

Let ω_1, ω_2, …, ω_n be the basis of V^* dual to the basis e_1, e_2, …, e_n of V considered in the previous section. We shall show that for all k, $\Lambda_k V^*$ may be identified with the dual space $\Lambda^k V$ of $\Lambda_k V$.

First of all, for $S := \{i_1, i_2, \ldots, i_k\}$, where $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, we may regard $\omega_S := \omega_{i_1} \wedge \omega_{i_2} \wedge \cdots \wedge \omega_{i_k}$ as a linear functional on $\Lambda_k V$ by putting

$$\omega_S(e_{j_1} \wedge \cdots \wedge e_{j_k}) := \det\bigl(\omega_{i_s}(e_{j_t})\bigr)$$

and extending to $\Lambda_k V$ by linearity. Then for $\Phi := \sum a_S \omega_S$, and $\xi \in \Lambda_k V$, we define $\Phi(\xi) := \sum_S a_S \omega_S(\xi)$. It follows that for all φ_1, φ_2, …, $\varphi_k \in V^*$ and for all v_1, v_2, …, $v_k \in V$, we have

$$(\varphi_1 \wedge \cdots \wedge \varphi_k)(v_1 \wedge \cdots \wedge v_k) = \det\bigl(\varphi_i(v_j)\bigr). \tag{12.5}$$

Therefore the definition of the action of $\Phi \in \Lambda_k V^*$ on $\Lambda_k V$ does not depend on the choice of basis for V.

This identification of $\Lambda_k V^*$ with $(\Lambda_k V)^*$ extends to an identification of $\Lambda_\bullet V^*$ with the dual space $\Lambda^\bullet V$ of $\Lambda_\bullet V$. The basis $\{\, \omega_S \mid S \subseteq I \,\}$ is dual to the basis $\{\, e_S \mid S \subseteq I \,\}$.

Decomposable k-vectors

If V is a vector space with basis e_1, e_2, …, e_n, then from (12.1) the exterior product $e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}$ vanishes whenever any two of its terms coincide.

From this it follows that the vectors $v_1, v_2, \ldots, v_k$ of V are linearly dependent if and only if $v_1 \wedge v_2 \wedge \cdots \wedge v_k = 0$.

A non-zero k-vector is said to be *decomposable* (or *pure*) if it can be written in the form $v_1 \wedge v_2 \wedge \cdots \wedge v_k$ for some $v_1, v_2, \ldots, v_k \in V$.

If W is a subspace of V, the inclusion of W in V extends to an inclusion of $\Lambda_\bullet W$ in $\Lambda_\bullet V$ and we may identify $\Lambda_\bullet W$ with its image in $\Lambda_\bullet V$. In particular, if $\dim W = k$, then $\Lambda_k W$ is a 1-dimensional subspace of $\Lambda_k V$ spanned by the decomposable k-vector $w_1 \wedge w_2 \wedge \cdots \wedge w_k$, where $w_1, w_2, \ldots, w_k$ is a basis for W. Thus there is a mapping $W \mapsto \Lambda_k W$ from the set $G_k(V)$ of k-dimensional subspaces of V to the set $G_1(\Lambda_k V)$ of points of $\mathcal{P}(\Lambda_k V)$. The next lemma shows that this mapping is one-to-one.

12.6 Lemma. *If $w_1, w_2, \ldots, w_k$ and $w_1', w_2', \ldots, w_k'$ are two sets of linearly independent vectors of V, then*

$$\langle\, w_1, w_2, \ldots, w_k \,\rangle = \langle\, w_1', w_2', \ldots, w_k' \,\rangle$$

if and only if $w_1 \wedge w_2 \wedge \cdots \wedge w_k = a w_1' \wedge w_2' \wedge \cdots \wedge w_k'$ for some $a \in \mathbb{F}$.

Proof. If $\langle\, w_1, w_2, \ldots, w_k \,\rangle = \langle\, w_1', w_2', \ldots, w_k' \,\rangle$, and if W denotes this subspace, then both $w_1 \wedge w_2 \wedge \cdots \wedge w_k$ and $w_1' \wedge w_2' \wedge \cdots \wedge w_k'$ span $\Lambda_k W$; hence each is a scalar multiple of the other.

Conversely, if $w_1 \wedge w_2 \wedge \cdots \wedge w_k = a w_1' \wedge w_2' \wedge \cdots \wedge w_k'$, then for all i, $w_i' \wedge w_1 \wedge \cdots \wedge w_k = 0$, and therefore w_i' is a linear combination of $w_1, w_2, \ldots, w_k$. It follows that $w_1', w_2', \ldots, w_k' \in \langle\, w_1, w_2, \ldots, w_k \,\rangle$ and hence $\langle\, w_1, w_2, \ldots, w_k \,\rangle = \langle\, w_1', w_2', \ldots, w_k' \,\rangle$. □

It turns out that not every element of $\Lambda_k V$ is decomposable, and therefore the map $G_k(V) \to G_1(\Lambda_k V)$ is not onto. In general, the decomposable k-vectors can be described succinctly by means of the interior product between $\Lambda_\bullet V$ and its dual $\Lambda^\bullet V$ (see Bourbaki (1970), Chap. III, §11, Proposition 16). However, for the elements of $\Lambda_2 V$ the situation is somewhat simpler and the equations characterizing the decomposable bivectors can be expressed in terms of annihilation operators—a special case of the interior product.

Creation and Annihilation Operators

For $v \in V$, the linear transformation

$$\alpha_v^+ : \Lambda_\bullet V \to \Lambda_\bullet V$$

defined by $\alpha_v^+(\xi) := v \wedge \xi$ is called a *creation operator*. It is clear from this definition that $\alpha_v^+(\Lambda_k V) \subseteq \Lambda_{k+1} V$ for all k.

12.7 Lemma. *For all $\varphi \in V^*$ there is a unique linear transformation $\alpha_\varphi^- : \Lambda_\bullet V \to \Lambda_\bullet V$ such that $\alpha_\varphi^-(1) = 0$ and, for all $v \in V$,*

$$\alpha_\varphi^- \alpha_v^+ + \alpha_v^+ \alpha_\varphi^- = \varphi(v)\mathbf{1}. \tag{12.8}$$

Furthermore,

(*i*) $\alpha_{a\varphi+b\psi}^- = a\alpha_\varphi^- + b\alpha_\psi^-$ *for all a, $b \in \mathbb{F}$ and all φ, $\psi \in V^*$,*

(*ii*) $(\alpha_\varphi^-)^2 = 0$,

(*iii*) $\alpha_\varphi^-\alpha_\psi^- + \alpha_\psi^-\alpha_\varphi^- = 0$, *for all φ, $\psi \in V^*$, and*

(*iv*) $\alpha_\varphi^-(\xi\wedge\eta) = \alpha_\varphi^-(\xi)\wedge\eta + (-1)^k\xi\wedge\alpha_\varphi^-(\eta)$, *for all $\xi \in \Lambda_k V$ and all $\eta \in \Lambda_\bullet V$.*

(*v*) *If $\alpha_\varphi^-(\xi) = 0$ for all $\varphi \in V^*$, then $\xi = 0$.*

Proof. For $\xi \in \Lambda_k V$, equation (12.8) can be written in the form

$$\alpha_\varphi^-(v \wedge \xi) = \varphi(v)\xi - v \wedge \alpha_\varphi^-(\xi).$$

In particular, if e_1, e_2, ..., e_n is a basis for V, it is equivalent to

$$\alpha_\varphi^-(e_{i_1} \wedge \cdots \wedge e_{i_k}) = \varphi(e_{i_1})e_{i_2} \wedge \cdots \wedge e_{i_k} - e_{i_1} \wedge \alpha_\varphi^-(e_{i_2} \wedge \cdots \wedge e_{i_k}) \tag{12.9}$$

and the existence and uniqueness of α_φ^- follow by induction on k.

Equation (12.8) is linear in φ and therefore (*i*) holds by the uniqueness result just proved. On applying α_φ^- to (12.9) we see that $(\alpha_\varphi^-)^2$ vanishes on the basis elements of $\Lambda_\bullet V$ and therefore $(\alpha_\varphi^-)^2 = 0$, proving (*ii*). We obtain (*iii*) by replacing φ by $\varphi + \psi$ in (*ii*).

Suppose that (*iv*) holds for $\xi \in \Lambda_k V$. (This is certainly true when $k = 0$ or $k = 1$.) Then for $v \in V$, $\xi \in \Lambda_k V$ and $\eta \in \Lambda_\bullet V$,

$$\begin{aligned}\alpha_\varphi^-(v \wedge \xi \wedge \eta) &= \varphi(v)\xi \wedge \eta - v \wedge \alpha_\varphi^-(\xi \wedge \eta)\\ &= \varphi(v)\xi \wedge \eta - v \wedge \alpha_\varphi^-(\xi) \wedge \eta - (-1)^k v \wedge \xi \wedge \alpha_\varphi^-(\eta)\\ &= \alpha_\varphi^-(v \wedge \xi) \wedge \eta + (-1)^{k+1}(v \wedge \xi) \wedge \alpha_\varphi^-(\eta).\end{aligned}$$

The elements $v\wedge\xi$ span $\Lambda_{k+1}V$ and therefore (*iv*) holds for all k by induction.

Finally, suppose that $\alpha_\varphi^-(\xi) = 0$ for all $\varphi \in V^*$. If $\xi = \xi_0+\xi_1+\cdots+\xi_n$, where $\xi_k \in \Lambda_k V$, then it follows from (12.9), by induction, that $\alpha_\varphi^-(\xi_k) \in \Lambda_{k-1}V$. Therefore $\alpha_\varphi^-(\xi_k) = 0$ for all k. Choose a basis e_1, e_2, ..., e_n for V and let ω_1, ω_2, ..., ω_n be the dual basis. If $H := \langle e_2, \ldots, e_n \rangle$, we may choose η_1 and η in $\Lambda_\bullet H$ such that $\xi_k = e_1 \wedge \eta_1 + \eta$. From (12.9), by induction, we have $\alpha_{\omega_1}^-(\eta_1) = \alpha_{\omega_1}^-(\eta) = 0$, and therefore $\eta_1 = \alpha_{\omega_1}^-(\xi_k) = 0$. Consequently, $\xi_k \in \Lambda_\bullet H$ and by induction on $\dim V$ we have $\xi_k = 0$ for all k. □

The linear transformation α_φ^- is called an *annihilation operator* and for all k we have $\alpha_\varphi^-(\Lambda_k V) \subseteq \Lambda_{k-1}V$.

12.10 Theorem. *A non-zero element ξ of $\Lambda_2 V$ is decomposable if and only if $\alpha^-_\varphi(\xi) \wedge \xi = 0$ for all $\varphi \in V^*$.*

Proof. If $\xi = u \wedge v$, then $\alpha^-_\varphi(\xi) = \varphi(u)v - \varphi(v)u$ and therefore $\alpha^-_\varphi(\xi) \wedge \xi = 0$. Conversely, suppose that $\alpha^-_\varphi(\xi) \wedge \xi = 0$ for all $\varphi \in V^*$. By Lemma 12.7 (*v*) we may choose $\varphi \in V^*$ so that $e_1 := \alpha^-_\varphi(\xi) \neq 0$. Extend e_1 to a basis e_1, e_2, $\dots$, e_n for V and write $\xi = \sum_{i<j} p_{ij} e_i \wedge e_j$. Then the condition $e_1 \wedge \xi = 0$ implies $\xi = e_1 \wedge \big(\sum_{1<j} p_{1j} e_j\big)$. □

There is a duality between the annihilation operators of $\Lambda_\bullet V$ and the creation operators of $\Lambda^\bullet V$. The precise connection is given in the next theorem. Recall (from Chapter 7) that the *transpose* of a σ-semilinear transformation $f : V_1 \to V_2$ is the map $f^* : V_2^* \to V_1^*$ defined by $f^*(\varphi) := \sigma^{-1}\varphi f$.

12.11 Theorem. *For all $\varphi \in V^*$ we have $\alpha^+_\varphi = (\alpha^-_\varphi)^*$.*

Proof. We must show that for all $\xi \in \Lambda_k V$ and all $\Phi \in \Lambda^{k-1} V$, we have

$$(\varphi \wedge \Phi)\xi = \Phi(\alpha^-_\varphi(\xi)).$$

To this end it suffices to take $\xi := v_1 \wedge \cdots \wedge v_k$, $\Phi := \varphi_2 \wedge \cdots \wedge \varphi_k$, and to put $\varphi_1 := \varphi$. Then from (12.5) we have $(\varphi \wedge \Phi)\xi = \det\big(\varphi_i(v_j)\big)$, and from the formula given in Exercise 12.4 (an immediate consequence of (12.8) and induction), it follows that $\Phi(\alpha^-_\varphi(\xi))$ is simply the expansion of the determinant of $\big(\varphi_i(v_j)\big)$ according to its first row. Thus $(\varphi \wedge \Phi)\xi = \Phi(\alpha^-_\varphi(\xi))$ and the theorem is proved. □

We conclude this section by describing the interaction between semilinear transformations and the annihilation operators.

12.12 Theorem. *If $f : V \to W$ is a σ-semilinear transformation, then for all $\varphi \in W^*$, and for all $\xi \in \Lambda_k V$ we have*

$$(\Lambda_{k-1} f)\alpha^-_{f^*(\varphi)}(\xi) = \alpha^-_\varphi\big((\Lambda_k f)\xi\big).$$

Proof. We argue by induction on k. If $k = 1$, then $\xi \in V$ and it follows that $(\Lambda_0 f)\alpha^-_{f^*(\varphi)}(\xi) = \varphi f(\xi)$, as required. If $k > 1$ and $(\Lambda_{k-1} f)\alpha^-_{f^*(\varphi)}(\xi) = \alpha^-_\varphi\big((\Lambda_k f)\xi\big)$, then for all $v \in V$,

$$\begin{aligned}(\Lambda_k f)\alpha^-_{f^*(\varphi)}(v \wedge \xi) &= \varphi f(v)(\Lambda_k f)\xi - f(v) \wedge (\Lambda_{k-1} f)\alpha^-_{f^*(\varphi)}(\xi)\\ &= \varphi f(v)(\Lambda_k f)\xi - f(v) \wedge \alpha^-_\varphi\big((\Lambda_k f)\xi\big)\\ &= \alpha^-_\varphi\big((\Lambda_{k+1} f)(v \wedge \xi)\big).\end{aligned}$$

The elements $v \wedge \xi$ span $\Lambda_{k+1}V$ and so the result holds for $k+1$. This completes the induction. □

The Klein Quadric

From now on suppose that $\dim V = 4$. Let e_1, e_2, e_3, e_4 be a basis for V and put

$$\tilde{e} := e_1 \wedge e_2 \wedge e_3 \wedge e_4.$$

We shall show that the decomposable elements of $\Lambda_2 V$ are the singular vectors for a quadratic form Q of Witt index 3. To define Q, first write $\xi := \sum_{i<j} p_{ij} e_i \wedge e_j$ and then put

$$Q(\xi) := p_{12}p_{34} - p_{13}p_{24} + p_{14}p_{23}. \tag{12.13}$$

It is immediate from this description that $(e_1 \wedge e_2, e_3 \wedge e_4)$, $(e_2 \wedge e_4, e_1 \wedge e_3)$ and $(e_1 \wedge e_4, e_2 \wedge e_3)$ are mutually orthogonal hyperbolic pairs for Q; hence Q is non-degenerate and its Witt index is 3. We let β denote the polar form of Q.

At first sight it might appear that Q depends on the basis chosen for V, but the next theorem shows that it depends only on $\tilde{e}$; i.e., it is uniquely determined up to a non-zero scalar multiple.

12.14 Theorem. *For all ξ, $\eta \in \Lambda_2 V$ and all $\varphi \in V^*$ we have*

(i) $\alpha_\varphi^-(\xi) \wedge \xi = Q(\xi)\alpha_\varphi^-(\tilde{e})$, *and*

(ii) $\xi \wedge \eta = \beta(\xi, \eta)\tilde{e}$.

Proof. Equation (i) is linear in φ and therefore it suffices to check it for the linear functionals ω_1, ω_2, ω_3, ω_4 which form the basis dual to e_1, e_2, e_3, e_4. In this case (i) is a short and easy calculation.

Replacing ξ by $\xi + \eta$ in (i), and simplifying, we obtain

$$\alpha_\varphi^-(\xi) \wedge \eta + \alpha_\varphi^-(\eta) \wedge \xi = \beta(\xi, \eta)\alpha_\varphi^-(\tilde{e}).$$

By Lemma 12.7 (iv), the left hand side is $\alpha_\varphi^-(\xi \wedge \eta)$ and then (ii) follows from Lemma 12.7 (v). □

It follows from (i) that the conditions given in Theorem 12.10 for the non-zero bivector ξ to be decomposable reduce to the single equation $Q(\xi) = 0$. This means that the map $L \mapsto \Lambda_2 L$ is a bijection between the lines of $\mathcal{P}(V)$ and the singular points of $\mathcal{P}(\Lambda_2 V)$. The set of singular points of $\mathcal{P}(\Lambda_2 V)$ is called the *Klein quadric*.

12.15 Lemma. *The lines L_1 and L_2 of $\mathcal{P}(V)$ have a common point if and only if $\Lambda_2 L_1$ and $\Lambda_2 L_2$ are orthogonal (with respect to β).*

Proof. Suppose that $L_1 := \langle u_1, v_1 \rangle$ and $L_2 := \langle u_2, v_2 \rangle$. Then L_1 and L_2 have a common point if and only if u_1, v_1, u_2 and v_2 are linearly dependent. This is the case if and only if $u_1 \wedge v_1 \wedge u_2 \wedge v_2 = 0$ and so the result now follows from Theorem 12.14 (*ii*). □

We know from the last two sections of Chapter 11 that the maximal totally singular subspaces of $\Lambda_2 V$ (which have dimension 3) form two classes: two maximal totally singular subspaces are *adjacent* if the dimension of their intersection is 2; and this adjacency relation defines a bipartite graph such that distinct subspaces are in the same class if and only if the dimension of their intersection is 1.

We can see all this directly, using the geometry of $\mathcal{P}(V)$. If $P := \langle u \rangle$ is a point of $\mathcal{P}(V)$, then the lines through P have the form $\langle u, v \rangle$, $v \in V$, and it is clear from Lemma 12.15 that

$$\kappa(P) := \{ \Lambda_2 L \mid P \in L, \dim L = 2 \}$$

is a totally singular subspace of dimension 3. Similarly, if H is a plane of $\mathcal{P}(V)$, then every pair of lines of H have a common point and, again from Lemma 12.15,

$$\kappa(H) := \{ \Lambda_2 L \mid L \subseteq H, \dim L = 2 \}$$

is a maximal totally singular subspace.

12.16 Theorem. *The two classes of maximal totally singular subspaces of $\Lambda_2 V$ are*

$$\{ \kappa(P) \mid P \text{ is a point of } \mathcal{P}(V) \}, \quad \text{and}$$
$$\{ \kappa(H) \mid H \text{ is a plane of } \mathcal{P}(V) \}.$$

Moreover, if P is a point and H is a plane of $\mathcal{P}(V)$, then $\kappa(P)$ is adjacent to $\kappa(H)$ if and only if $P \in H$.

Proof. Suppose that E is a maximal totally singular subspace of $\Lambda_2 V$ and let ξ_1, ξ_2, ξ_3 be a basis for E. If L_1, L_2 and L_3 are the lines of $\mathcal{P}(V)$ corresponding to ξ_1, ξ_2, and ξ_3 respectively, then either L_1, L_2 and L_3 have a point P in common, or else $H := \langle L_1, L_2, L_3 \rangle$ is a plane. In the first case E is $\kappa(P)$; in the second case it is $\kappa(H)$.

If $\kappa(P)$ is adjacent to $\kappa(H)$, then $\kappa(P) \cap \kappa(H) = \langle \xi, \eta \rangle$, where $\langle \xi \rangle = \Lambda_2 L$ and $\langle \eta \rangle = \Lambda_2 M$. Then $P \in L \cap M \subseteq H$. Conversely, if $P \in H$, the lines of H through P correspond to a 2-dimensional subspace of $\Lambda_2 V$ and hence $\kappa(P)$ is adjacent to $\kappa(H)$. □

For each line L of $\mathcal{P}(V)$, put $\kappa(L) := \Lambda_2 L$. Then κ defines an incidence preserving bijection between the projective geometry $\mathcal{P}(V)$ and the oriflamme geometry of $\mathcal{P}(\Lambda_2 V)$. This is the *Klein correspondence*.

Each totally singular line of $\mathcal{P}(\Lambda_2 V)$ has the form $\kappa(P) \cap \kappa(H)$ for a unique point P and a unique plane H which contains P. Thus, under the Klein correspondence, the totally singular lines of $\mathcal{P}(\Lambda_2 V)$ correspond to the incident point-plane pairs of $\mathcal{P}(V)$. An interpretation of the non-singular points of $\mathcal{P}(\Lambda_2 V)$ will be given later in the section 'Alternating Forms and Reflections'.

The collineations of $\mathcal{P}(V)$ induce automorphisms of the oriflamme geometry whereas the correlations of $\mathcal{P}(V)$ interchange the two classes of maximal totally singular subspaces. As a first step towards understanding this correspondence we have the following result.

12.17 Theorem. *For all $f \in GL(V)$ and for all $\xi \in \Lambda_2 V$ we have*

$$Q\big((\Lambda_2 f)\xi\big) = \det(f)\, Q(\xi).$$

Proof. Using Theorems 12.12 and 12.14, and equation (12.4) we have, for all $\varphi \in V^*$,

$$\begin{aligned} Q\big((\Lambda_2 f)\xi\big)\, \alpha_\varphi^-(\tilde{e}) &= \alpha_\varphi^-\big((\Lambda_2 f)\xi\big) \wedge (\Lambda_2 f)\xi \\ &= (\Lambda_3 f)(\alpha_{\varphi f}^-(\xi) \wedge \xi) \\ &= Q(\xi)\, (\Lambda_3 f)\alpha_{\varphi f}^-(\tilde{e}) \\ &= Q(\xi)\, \alpha_\varphi^-\big((\Lambda_4 f)\tilde{e}\big) \\ &= \det(f)\, Q(\xi)\, \alpha_\varphi^-(\tilde{e}). \end{aligned}$$

Thus $Q\big((\Lambda_2 f)\xi\big) = \det(f)\, Q(\xi)$, as required. □

For $\xi \in \Lambda_2 V$, we have $\xi \wedge \xi = 2Q(\xi)\,\tilde{e}$ and therefore, if the characteristic of $\mathbb{F}$ is not 2, the above calculations can be simplified considerably.

If σ is an automorphism of $\mathbb{F}$, and if f is the σ-semilinear transformation of V such that $f(e_i) = e_i$, for $1 \le i \le 4$, then for all bivectors ξ, we have $Q\big((\Lambda_2 f)\xi\big) = \sigma Q(\xi)$. Combining this observation with the theorem just proved, we see that the map $f \mapsto \Lambda_2 f$ is a homomorphism from $\Gamma L(V)$ to the full orthogonal group $\Gamma O(\Lambda_2 V)$. Every element of $\Gamma O(\Lambda_2 V)$ either fixes or interchanges the two classes of maximal totally isotropic subspaces of $\Lambda_2 V$ and thus induces either a collineation or a correlation of $\mathcal{P}(V)$. From what we have just seen, every collineation and correlation of $\mathcal{P}(V)$ must arise in this way. Furthermore, as $\Lambda_2 V$ is spanned by its singular points, an element of $\Gamma O(\Lambda_2 V)$ that fixes every singular point is a scalar multiple of $\mathbf{1}$. The group of all collineations and correlations of $\mathcal{P}(V)$ is the group $P\Gamma L^*(V)$, introduced in Chapter 7, and we have established that

12.18 Theorem. $P\Gamma L^*(V) \simeq P\Gamma O(\Lambda_2 V)$. □

This theorem is another expression of the Klein correspondence and from it we can derive further isomorphisms between linear groups defined on V and $\Lambda_2 V$. For example, if $h \in SO(\Lambda_2 V)$, then h induces a collineation of $\mathcal{P}(V)$ and therefore $h = a\Lambda_2 f$, for some $a \in \mathbb{F}^\times$ and some $f \in \Gamma L(V)$. In fact $f \in GL(V)$ and from Theorem 12.17 we have $a^2 \det(f) = 1$. If we put

$$K(V) := \{\, (a, f) \in \mathbb{F}^\times \times GL(V) \mid \det(f) = a^{-2} \,\},$$

then the map

$$\kappa : K(V) \to SO(\Lambda_2 V), \tag{12.19}$$

defined by $\kappa(a, f) := a\Lambda_2 f$, is a homomorphism onto $SO(\Lambda_2 V)$, and the kernel of κ is $\{\, (a^{-2}, a\,\mathbf{1}) \mid a \in \mathbb{F}^\times \,\} \simeq \mathbb{F}^\times$.

Note that $K(V)/\langle\, (-1, \mathbf{1})\,\rangle \simeq \{\, f \in GL(V) \mid \det(f) \in \mathbb{F}^2 \,\}$ and if $\mathbb{F}$ is a finite field, then $|K(V)| = |GL(V)|$ and $|SO(\Lambda_2 V)| = |SL(V)|$.

The Groups $SL(V)$ and $\Omega(\Lambda_2(V))$

By Theorem 4.3 *(ii)* we have $SL(V)' = SL(V)$ and therefore $\Lambda_2 f \in \Omega(\Lambda_2 V)$ for all $f \in SL(V)$. We identify the subgroup $\{\, (1, f) \mid f \in SL(V) \,\}$ of $K(V)$ with $SL(V)$ and consider the restriction of κ to this subgroup.

12.20 Theorem. *The map* $\kappa : SL(V) \to \Omega(\Lambda_2 V)$ *is a homomorphism onto* $\Omega(\Lambda_2 V)$ *whose kernel is* $\{\pm\mathbf{1}\}$.

Proof. We have $K(V)' = SL(V)$, and by Theorem 11.45, $SO(\Lambda_2 V)' = \Omega(\Lambda_2 V)$. The result now follows from (12.19). □

The centre of $SL(V)$ consists of the scalar transformations $a\mathbf{1}$ such that $a^4 = 1$ and it is mapped by κ to the centre of $\Omega(\Lambda_2 V)$. Thus $-\mathbf{1} \in \Omega(\Lambda_2 V)$ if and only if $-1 \in \mathbb{F}^2$, a result which was obtained in Chapter 11 by calculating the spinor norm of $-\mathbf{1}$. We also see that one of the order coincidences of p. 166 comes from an isomorphism of groups, namely

12.21 Corollary. $PSL(4, q) \simeq P\Omega^+(6, q)$, *for all prime powers* q. □

The group $SL(V)$ is generated by transvections and therefore it is of some interest to determine the image of a transvection under the map κ. Not surprisingly, the image is a Siegel transformation.

If $\varphi \in V^*$, $u \in V$ and $\varphi(u) = 0$, the transvection $t := t_{\varphi,u}$ is defined by

$$t(v) := v + \varphi(v)u.$$

Extend u to a basis $u_1 := u$, u_2, u_3, u_4 such that u_1, u_2, u_3 is a basis for $\ker\varphi$, $\varphi(u_4) = 1$, and $u_1 \wedge u_2 \wedge u_3 \wedge u_4 = \tilde{e}$. Then for $\xi \in \langle\, u_1 \wedge u_2 \,\rangle^\perp$ we have

$$(\Lambda_2 t)\xi = \xi + \beta(\xi, u_1 \wedge u_3)\, u_1 \wedge u_2 \tag{12.22}$$

and it follows from Theorem 11.18 (or directly from (11.17)) that $\kappa(t) = \Lambda_2 t$ is the Siegel transformation $\rho_{u_1\wedge u_2, u_1\wedge u_3}$. Note that $\operatorname{im}(\mathbf{1} - \Lambda_2 t)$ is the totally singular subspace $\langle\, u_1 \wedge u_2, u_1 \wedge u_3 \,\rangle$ corresponding to the pair $(\langle\, u \,\rangle, \ker\varphi)$, via the Klein correspondence.

Correlations

It is time to study the elements of $O(\Lambda_2 V)$ that interchange the two classes of maximal totally singular subspaces. We have already observed that these transformations induce correlations of $\mathcal{P}(V)$ and we know from Chapter 7 that every correlation is also induced by a semilinear map $g : v \mapsto \gamma(-, v)$, where γ is a non-degenerate sesquilinear form.

Therefore, we begin with a non-degenerate σ-sesquilinear form γ and note that $\Lambda_2 g : \Lambda_2 V \to \Lambda^2 V$ is a σ-semilinear isomorphism which sends $u \wedge v$ to $\gamma(-, u) \wedge \gamma(-, v)$.

The symmetric form β is non-degenerate and it too defines an isomorphism $\tilde{\beta} : \Lambda_2 V \to \Lambda^2 V : \xi \mapsto \beta(-, \xi)$. Combining these isomorphisms we obtain a σ-semilinear transformation $\kappa(g) := \tilde{\beta}^{-1}\Lambda_2 g$ of $\Lambda_2 V$. In order to study the properties of $\kappa(g)$ we need an explicit formula for $\tilde{\beta}^{-1}$, namely

12.23 Lemma. *For all φ, $\psi \in V^*$ we have $\tilde{\beta}^{-1}(\varphi \wedge \psi) = \alpha_\psi^- \alpha_\varphi^-(\tilde{e})$.*

Proof. Given u, $v \in V$, choose u', $v' \in V$ such that $\tilde{e} = u \wedge v \wedge u' \wedge v'$. Then from Theorem 12.14 (*ii*) and equations (12.9) and (12.5),

$$\begin{aligned} \beta(u \wedge v, \alpha_\psi^- \alpha_\varphi^-(\tilde{e}))\tilde{e} &= u \wedge v \wedge \alpha_\psi^- \alpha_\varphi^-(\tilde{e}) \\ &= (\varphi(u)\psi(v) - \varphi(v)\psi(u))\tilde{e} \\ &= (\varphi \wedge \psi)(u \wedge v)\tilde{e}. \end{aligned}$$

Thus $\beta(-, \alpha_\psi^- \alpha_\varphi^-(\tilde{e})) = \varphi \wedge \psi$ and therefore $\tilde{\beta}^{-1}(\varphi \wedge \psi) = \alpha_\psi^- \alpha_\varphi^-(\tilde{e})$. □

(The quantity $\alpha_\psi^- \alpha_\varphi^-(\tilde{e})$ is the interior product $\tilde{e} \llcorner (\varphi \wedge \psi)$ defined in Chap. III, §11 of Bourbaki (1970). Thus $\tilde{\beta}^{-1}\Phi = \tilde{e} \llcorner \Phi$ for all $\Phi \in \Lambda^2 V$.)

Let $\tilde{\omega}$ be the element of $\Lambda^4 V$ dual to $\tilde{e}$, i.e., $\tilde{\omega}(\tilde{e}) = 1$. On applying Theorem 12.14 (*ii*) to $\Lambda^2 V$ we see that there is a symmetric bilinear form B

defined on $\Lambda^2 V$ such that for all Φ, $\Psi \in \Lambda^2 V$, we have $\Phi \wedge \Psi = B(\Phi, \Psi)\,\tilde{\omega}$. There is a corresponding isomorphism $\widetilde{B} : \Lambda^2 V \to \Lambda_2 V$ given by

$$\widetilde{B}(\Phi) := B(\Phi, -).$$

12.24 Lemma. $\widetilde{B} = \tilde{\beta}^{-1}$.

Proof. Suppose that φ, $\psi \in V^*$ and $\Psi \in \Lambda^2 V$. Then from the previous lemma and Theorem 12.11,

$$\begin{aligned} \Psi\tilde{\beta}^{-1}(\varphi \wedge \psi) &= \Psi\alpha_\varphi^- \alpha_\psi^-(\tilde{e}) \\ &= (\varphi \wedge \psi \wedge \Psi)(\tilde{e}) \\ &= B(\varphi \wedge \psi, \Psi)\tilde{\omega}(\tilde{e}). \end{aligned}$$

It follows that $\tilde{\beta}^{-1}(\varphi \wedge \psi) = B(\varphi \wedge \psi, -)$. The elements $\varphi \wedge \psi$ span $\Lambda^2 V$ and thus the lemma is proved. □

If W is a subspace of V, let $W^\perp$ denote the orthogonal complement of W with respect to γ (defined on p. 52).

12.25 Theorem. *For all lines L of $\mathcal{P}(V)$, we have $\kappa(g)(\Lambda_2 L) = \Lambda_2 L^\perp$.*

Proof. Suppose that $L := \langle\, u_1, u_2 \,\rangle$ and extend u_1, u_2 to a basis u_1, u_2, u_3, u_4 for V. For $1 \le i \le 4$, put $\varphi_i := \gamma(-, u_i)$. Then φ_1, φ_2, φ_3, φ_4 is a basis for V^* and we may choose u_3 and u_4 so that the dual basis v_1, v_2, v_3, v_4 for V satisfies $v_1 \wedge v_2 \wedge v_3 \wedge v_4 = \tilde{e}$. It follows from Theorem 12.11 that

$$\begin{aligned} (\varphi_i \wedge \varphi_j)(\alpha_{\varphi_2}^- \alpha_{\varphi_1}^-(\tilde{e})) &= (\varphi_1 \wedge \varphi_2 \wedge \varphi_i \wedge \varphi_j)(\tilde{e}) \\ &= \begin{cases} 1 & \text{if } (i,j) = (3,4) \\ 0 & \text{otherwise,} \end{cases} \end{aligned}$$

and therefore

$$\alpha_{\varphi_2}^- \alpha_{\varphi_1}^-(\tilde{e}) = v_3 \wedge v_4.$$

That is, from Lemma 12.23 and the definition of $\kappa(g)$, we have

$$\kappa(g)(\Lambda_2 L) = \langle\, \kappa(g)(u_1 \wedge u_2) \,\rangle = \langle\, v_3 \wedge v_4 \,\rangle = \Lambda_2 L^\perp. \qquad \square$$

It is a consequence of this theorem that the collineation of $\mathcal{P}(\Lambda_2 V)$ induced by $\kappa(g)$ coincides with the collineation obtained, via the Klein correspondence, from the correlation of $\mathcal{P}(V)$ induced by g. (We leave it as an exercise to check the details.)

If $f \in \mathit{\Gamma L}(V)$ is σ-semilinear, the proof of Theorem 12.17 shows that for all $\xi \in \Lambda_2 V$, $Q(\kappa(f)\xi) = d(f)\,\sigma Q(\xi)$, where $d(f) \in \mathbb{F}$ and $(\Lambda_4 f)\tilde{e} = d(f)\,\tilde{e}$. Our next task is to establish an analogue of this result for $\kappa(g)$.

12.26 Theorem. *For all σ-semilinear transformations $g: V \to V^*$ and all $\xi \in \Lambda_2 V$ we have*

$$Q(\kappa(g)\xi) = d(g)\,\sigma Q(\xi),$$

where $d(g) := (\Lambda_4 g)(\tilde{e})\tilde{e}$.

Proof. From what we have just seen, this is true when ξ is decomposable, for then both sides vanish. The decomposable bivectors span $\Lambda_2 V$ and so the theorem will be proved provided

$$\beta(\kappa(g)\xi, \kappa(g)\eta) = d(g)\,\sigma\beta(\xi, \eta)$$

for all decomposable ξ, $\eta \in \Lambda_2 V$.

Taking $\xi := u \wedge v$, it follows from Lemma 12.23, the definition of κ, and Theorem 12.11 that

$$\begin{aligned}
\beta(\kappa(g)\xi, \kappa(g)\eta) &= \big((\Lambda_2 g)\eta\big)(\alpha^-_{g(v)}\alpha^-_{g(u)}(\tilde{e})) \\
&= (g(u) \wedge g(v) \wedge (\Lambda_2 g)\eta)(\tilde{e}) \\
&= (\Lambda_4 g)(\xi \wedge \eta)(\tilde{e}) \\
&= d(g)\,\sigma\beta(\xi, \eta),
\end{aligned}$$

where in the last step we have used Theorem 12.14 (*ii*). □

(If γ is a bilinear form, then $d(g) = \det\big(\gamma(e_i, e_j)\big)$ is the discriminant of γ.)

For all $f \in \mathit{\Gamma L}^*(V)$ we have an element $\kappa(f)$ of $\mathit{\Gamma O}(\Lambda_2 V)$, but the map $\kappa: \mathit{\Gamma L}^*(V) \to \mathit{\Gamma O}(\Lambda_2 V)$ is not quite a homomorphism. The extent to which it differs from a homomorphism is governed by the next lemma. The scalar factor $d(f)$ appearing there is $\tilde{\omega}(\Lambda_4 f)\tilde{e}$ if f is an element of $GL(V)$, or $(\Lambda_4 f)(\tilde{e})\tilde{e}$ if $f: V \to V^*$. For $f \in \mathit{\Gamma L}^*(V)$, $\overline{f}$ denotes the inverse of the transpose f^* of f.

12.27 Lemma. *For all $f \in \mathit{\Gamma L}^*(V)$,*

$$(\Lambda_2 f)\tilde{\beta}^{-1} = \begin{cases} d(f)\,\tilde{\beta}^{-1}\Lambda_2 \overline{f} & \text{if } f: V \to V, \\ d(f)\,\tilde{\beta}\Lambda_2 \overline{f} & \text{if } f: V \to V^*. \end{cases}$$

Proof. If $f \in \mathit{\Gamma L}(V)$, then from Lemma 12.23 and Theorem 12.12, and for all φ, $\psi \in V^*$, we have

$$\begin{aligned}
(\Lambda_2 f)\tilde{\beta}^{-1}(\varphi \wedge \psi) &= (\Lambda_2 f)\alpha^-_{\psi}\alpha^-_{\varphi}(\tilde{e}) \\
&= \alpha^-_{\bar{f}(\psi)}\alpha^-_{\bar{f}(\varphi)}(\tilde{e}) \\
&= d(f)\,\tilde{\beta}^{-1}(\Lambda_2 \overline{f})(\varphi \wedge \psi),
\end{aligned}$$

and hence $(\Lambda_2 f)\tilde{\beta}^{-1} = d(f)\,\tilde{\beta}^{-1}(\Lambda_2\overline{f})$.

Similarly, if $f : V \to V^*$, then for all u, $v \in V$,

$$\begin{aligned}(\Lambda_2 f)\tilde{\beta}^{-1}(\Lambda_2 f^*)(u\wedge v) &= (\Lambda_2 f)\alpha^-_{f^*(v)}\alpha^-_{f^*(u)}(\tilde{e})\\ &= \alpha^-_v\alpha^-_u(\Lambda_4 f)(\tilde{e})\\ &= d(f)\,\alpha^-_v\alpha^-_u(\tilde{\omega})\\ &= d(f)\,\widetilde{B}^{-1}(u\wedge v)\\ &= d(f)\,\tilde{\beta}(u\wedge v)\end{aligned}$$

Thus $(\Lambda_2 f)\tilde{\beta}^{-1}(\Lambda_2 f^*) = d(f)\,\tilde{\beta}$ and this completes the proof. □

12.28 Theorem. *If f, $g \in \Gamma L^*(V)$, then*

$$\kappa(f)\kappa(g) = \begin{cases}\kappa(f\circ g) & \text{if } g : V \to V,\\ d(f)\,\kappa(f\circ g) & \text{if } g : V \to V^*.\end{cases}$$

Proof. This follows directly from the previous lemma and the definition of the product in $\Gamma L^*(V)$ (given in Chapter 7). □

For $f \in \Gamma L^*(V)\setminus\Gamma L(V)$, the definition of $\kappa(f)$ depends on $\tilde{e}$. Changing $\tilde{e}$ to $a\tilde{e}$ replaces β by $a^{-1}\beta$ and κ by $a\kappa$ (Exercise 12.9). Even though the map $\kappa : \Gamma L^*(V) \to \Gamma O(\Lambda_2 V)$ is not a homomorphism, it induces the isomorphism of Theorem 12.18. On the other hand it is possible to obtain a homomorphism from κ as follows.

Let $GL^*(V)$ denote the group of linear transformations in $\Gamma L^*(V)$, and put

$$K^*(V) := \{\, (a, f) \in \mathbb{F}^\times \times GL^*(V) \mid d(f) = a^{-2}\,\}.$$

Make $K^*(V)$ into a group by defining the product of (a, f) and (b, g) to be

$$(a, f)\circ(b, g) := \begin{cases}(ab, fg) & \text{if } g : V \to V\\ (a^{-1}b, \overline{f}g) & \text{if } g : V \to V^*.\end{cases}$$

According to Exercise 12.10 (i), for f, $g \in GL^*(V)$ we have

$$d(f\circ g) = \begin{cases}d(f)d(g) & \text{if } g : V \to V\\ d(f)^{-1}d(g) & \text{if } g : V \to V^*\end{cases}$$

and therefore $K^*(V)$ is closed with respect to this product.

12.29 Theorem. *The map $\kappa : K^*(V) \to O(\Lambda_2 V)$ defined by $\kappa(a,f) := a\kappa(f)$ is a homomorphism onto $O(\Lambda_2 V)$ whose kernel is*

$$\{\,(a^{-2}, a\,\mathbf{1}) \mid a \in \mathbb{F}^\times\,\}.$$

Proof. Exercise 12.10 (*iii*). □

Alternating Forms and Reflections

Suppose that γ is a non-degenerate alternating form on V and let u_1, v_1, u_2, v_2 be a symplectic basis for V with respect to γ. The discriminant of γ with respect to this basis is 1 and therefore, if $u_1 \wedge v_1 \wedge u_2 \wedge v_2 = a\tilde{e}$, then $d(g) = a^{-2}$, where g is the map $v \mapsto \gamma(-, v)$. Thus $(a, g) \in K^*(V)$ and $(a, g)^2 = (1, -\mathbf{1})$.

The map $\tilde{\beta} : \Lambda_2 V \to \Lambda^2 V$ is a bijection and therefore there is a unique bivector γ_0 such that, for all u, $v \in V$,

$$\gamma(u, v) = \beta(u \wedge v, \gamma_0). \tag{12.30}$$

From Theorem 12.14 (*ii*) this is equivalent to

$$\gamma(u, v)\tilde{e} = u \wedge v \wedge \gamma_0.$$

A direct calculation shows that $\gamma_0 = a^{-1}(u_1 \wedge v_1 + u_2 \wedge v_2)$ and therefore $Q(\gamma_0) = \beta(u_1 \wedge v_1, u_2 \wedge v_2)^{-1} = a^{-1} \neq 0$. (We call $Q(\gamma_0)$ the *Pfaffian* of γ.)

Conversely, if $Q(\xi) \neq 0$, then any hyperbolic line through $\langle\,\xi\,\rangle$ has just two singular points and therefore we may write $\xi = e_1 \wedge f_1 + e_2 \wedge f_2$, where e_1, f_1, e_2, f_2 is a symplectic basis for the alternating form $(u, v) \mapsto \beta(u \wedge v, \xi)$. It follows that $\tilde{\beta}$ restricts to a bijection between the non-singular bivectors and the non-degenerate alternating forms on V.

12.31 Theorem. *Let g be the map $v \mapsto \gamma(-, v)$, where γ is the alternating form $\beta(-, \gamma_0)$ corresponding to the non-singular bivector γ_0, and let $a := Q(\gamma_0)^{-1}$. Then $(-a, g) \in K^*(V)$ and $\kappa(-a, g)$ is the reflection in $\langle\,\gamma_0\,\rangle^\perp$. Moreover, the map $\kappa : K^*(V) \to O(\Lambda_2 V)$ restricts to a homomorphism from $Sp(V)$ onto $\Omega(\langle\,\gamma_0\,\rangle^\perp)$ whose kernel is $\{\pm\mathbf{1}\}$.*

Proof. We have $d(g) = (-a)^{-2}$ and therefore $(-a, g) \in K^*(V)$. If u_1, v_1, u_2, v_2 is a symplectic basis for V with respect to γ, then (as above) $u_1 \wedge v_1 \wedge u_2 \wedge v_2 = a\tilde{e}$. A direct calculation, similar to the one of Theorem 12.25, shows that $\kappa(-a, g)$ interchanges $u_1 \wedge v_1$ and $-u_2 \wedge v_2$, and fixes $u_1 \wedge u_2$, $u_1 \wedge v_2$, $v_1 \wedge u_2$ and $v_1 \wedge v_2$. Thus

$$\kappa(-a, g)\xi = \xi - a\beta(\xi, \gamma_0)\gamma_0$$

and therefore $\kappa(-a, g)$ is the reflection in $\langle \gamma_0 \rangle^\perp$.

The symplectic group $Sp(V)$ defined by γ consists of the elements of $GL(V)$ that commute with g. If $f \in Sp(V)$, it follows from (12.30) that $\Lambda_2 f = \kappa(1, f)$ fixes γ_0 and preserves Q, hence $\kappa(1, f) \in O(W)$, where $W := \langle \gamma_0 \rangle^\perp$. In fact, $\kappa(1, f) \in \Omega(W)$, because $Sp(V)' = Sp(V)$.

Conversely, suppose that $h \in \Omega(W)$. Then by Witt's theorem and the fact that $\Omega(W) = O(W)'$, h can be extended to an element of $\Omega(\Lambda_2 V)$ that fixes γ_0. From Theorem 12.20, $h = \kappa(1, f)$ for some $f \in SL(V)$. Then the commutator $[g, f]$ belongs to $SL(V)$ and $\kappa[(-a, g), (1, f)] \in \Omega(\Lambda_2 V)$ fixes every vector of W; hence it must be the identity. It follows that $g \circ f = \pm f \circ g$. That is, for all u, $v \in V$, either $\gamma(f(u), f(v)) = \gamma(u, v)$ or $\gamma(f(u), f(v)) = -\gamma(u, v)$. But if $\gamma(f(u), f(v)) = -\gamma(u, v)$, it follows from (12.30) that $(\Lambda_2 f)\gamma_0 = -\gamma_0$. Hence we must have $f \in Sp(V)$. □

As a corollary we lift another of the order coincidences of p. 166 to an isomorphism of groups.

12.32 Corollary. *$PSp(4, q) \simeq P\Omega(5, q)$, for all prime powers q.* □

Under the Klein correspondence, the totally isotropic lines of $\mathcal{P}(V)$ (with respect to γ) correspond to the singular points of $\mathcal{P}(\langle \gamma_0 \rangle^\perp)$. If P is a point of $\mathcal{P}(V)$, then $\kappa(P) \cap \kappa(P^\perp)$ is the totally singular line of $\mathcal{P}(\langle \gamma_0 \rangle^\perp)$ whose points represent the totally isotropic lines of $\mathcal{P}(V)$ through P. The configuration of points and totally isotropic lines of $\mathcal{P}(V)$ is a (symplectic) generalized quadrangle. Under the Klein correspondence it is dual to the (orthogonal) generalized quadrangle of singular points and totally singular lines of $\mathcal{P}(\langle \gamma_0 \rangle^\perp)$.

Hermitian Forms of Witt Index 2

Let γ be a non-degenerate σ-hermitian form on V and let $g : V \to V^*$ be the associated σ-semilinear map $v \mapsto \gamma(-, v)$. As in Chapter 10, let $\mathbb{F}_0$ be the fixed field of σ, and for $x \in \mathbb{F}$, write $\bar{x}$ as an abbreviation for $\sigma(x)$.

We shall suppose that the Witt index of γ is 2. If $\mathbb{F}$ is finite, this is the only possibility. But if $\mathbb{F}$ is infinite, γ could have Witt index 0 or 1. See Dieudonné (1952) and Ohara (1958) for the details concerning forms of index 0 and 1.

As $\sigma^2 = 1$, we may write $\mathbb{F} = \mathbb{F}_0[\theta]$, where θ satisfies the irreducible quadratic equation $x^2 - ax + b = 0$, with coefficients $a := \theta + \bar{\theta}$ and $b := \theta\bar{\theta}$ in $\mathbb{F}_0$. For future use we let δ be an element of $\mathbb{F}^\times$ such that $\bar{\delta} = -\delta$. If the characteristic of $\mathbb{F}$ is not 2, we may take θ to be δ.

Choose a basis u_1, v_1, u_2, v_2 for V such that $\langle u_1, v_1 \rangle = \langle u_2, v_2 \rangle^\perp$, and (u_1, v_1) and (u_2, v_2) are hyperbolic pairs for γ. The form γ will remain fixed throughout this section and therefore we may as well set $\tilde{e} := u_1 \wedge v_1 \wedge u_2 \wedge v_2$. Then $d(g) = 1$ and $Q(\kappa(g)\xi) = \sigma Q(\xi)$, for all $\xi \in \Lambda_2 V$.

For u, $v \in V$, it follows from Lemma 12.23 that

$$\kappa(g)(u \wedge v) = \alpha^-_{\gamma(-,v)} \alpha^-_{\gamma(-,u)}(\tilde{e}),$$

and hence the action of $\kappa(g)$ on $\Lambda_2 V$ is given by:

$$\begin{aligned} \kappa(g)(u_1 \wedge v_1) &= -u_2 \wedge v_2, & \kappa(g)(u_2 \wedge v_2) &= -u_1 \wedge v_1,\\ \kappa(g)(u_1 \wedge u_2) &= -u_1 \wedge u_2, & \kappa(g)(v_1 \wedge v_2) &= -v_1 \wedge v_2,\\ \kappa(g)(u_1 \wedge v_2) &= u_1 \wedge v_2, & \kappa(g)(v_1 \wedge u_2) &= v_1 \wedge u_2. \end{aligned}$$

The map $\kappa(g)$ is σ-semilinear and therefore its set of fixed points is the vector space W_0 over $\mathbb{F}_0$ with basis

$$\begin{aligned} \xi_1 &:= u_1 \wedge v_1 - u_2 \wedge v_2, & \xi_2 &:= -\theta u_1 \wedge v_1 + \bar{\theta} u_2 \wedge v_2,\\ \xi_3 &:= \delta u_1 \wedge u_2, & \xi_4 &:= \delta v_1 \wedge v_2,\\ \xi_5 &:= u_1 \wedge v_2, & \xi_6 &:= v_1 \wedge u_2. \end{aligned}$$

For $\xi := \sum_{i=1}^6 x_i \xi_i$, we have

$$Q(\xi) = -x_1^2 + a x_1 x_2 - b x_2^2 - \delta^2 x_3 x_4 + x_5 x_6$$

and so the restriction of Q to W_0 is a quadratic form of Witt index 2.

12.33 Lemma. *If $\kappa(g)\xi = \lambda\xi$, for some $\lambda \in \mathbb{F}$ and some non-zero $\xi \in \Lambda_2 V$, then $\mu\xi \in W_0$, for some $\mu \in \mathbb{F}$.*

Proof. We have $\kappa(g)^2 = \mathbf{1}$, and therefore $\xi = \kappa(g)^2\xi = \lambda\bar{\lambda}\xi$. Thus $\lambda\bar{\lambda} = 1$ and from Lemma 10.1 (*iv*), $\lambda = \mu/\bar{\mu}$ for some $\mu \in \mathbb{F}$. Then $\kappa(g)(\mu\xi) = \mu\xi$, as required. □

From Theorem 12.25, a line L of $\mathcal{P}(V)$ is totally isotropic if and only if $\Lambda_2 L$ is fixed by $\kappa(g)$. From the lemma just proved, this is the case if and only if $(\Lambda_2 L) \cap W_0 \neq \emptyset$. Thus the Klein correspondence restricts to a bijection between the totally isotropic lines of $\mathcal{P}(V)$ and the singular points of $\mathcal{P}(W_0)$. Moreover, if P is an isotropic point of $\mathcal{P}(V)$, then $\kappa(P) \cap W_0 = \kappa(P^\perp) \cap W_0$ is the totally singular line of $\mathcal{P}(W_0)$ whose points represent the totally isotropic lines of $\mathcal{P}(V)$ through P. Two isotropic points of $\mathcal{P}(V)$ span a totally isotropic line of $\mathcal{P}(V)$ if and only if the corresponding totally singular lines of

$\mathcal{P}(W_0)$ have a unique common point. Thus, under the Klein correspondence, the (unitary) generalized quadrangle of totally isotropic points and lines of $\mathcal{P}(V)$ is dual to the (orthogonal) generalized quadrangle of totally singular points and lines of $\mathcal{P}(W_0)$.

Let $SU(V)$ be the special unitary group defined by γ. If $f \in SU(V)$, then $\kappa(f)$ preserves Q and from Theorem 12.28 (and because $d(g) = 1$), $\kappa(f)$ commutes with $\kappa(g)$. Hence $\kappa(f)$ acts on W_0. As $SU(V)' = SU(V)$, it follows that $\kappa(f) \in \Omega(W_0)$. We shall show that every element of $\Omega(W_0)$ arises in this way, but first we need an analogue of (12.19).

For $f \in GL(V)$ and $a \in \mathbb{F}^\times$, the condition $[g, f] = a\,\mathbf{1}$ is equivalent to $g^{-1}f^{*-1}gf^{-1} = a\,\mathbf{1}$, which in turn is equivalent to $a\gamma(f(u), f(v)) = \gamma(u, v)$ for all $u, v \in V$. From this last condition it is easy to see that $a \in \mathbb{F}_0$. The appropriate analogue of the group $K(V)$ appearing in (12.19) is the subgroup

$$K_\gamma(V) := \{\, (a, f) \in \mathbb{F}_0^\times \times GL(V) \mid [g, f] = \pm a\,\mathbf{1},\ \det(f) = a^{-2}\,\}.$$

For $(a, f) \in K_\gamma(V)$, we have $\kappa(a, f) \in SO(\Lambda_2 V)$ and as $\kappa(a, f)$ commutes with $\kappa(1, g)$, it follows that $\kappa(a, f) \in SO(W_0)$.

12.34 Theorem. *The map $\kappa\colon K_\gamma(V) \to SO(W_0)$ is a homomorphism onto $SO(W_0)$ whose kernel is $\{(a^{-2}, a\,\mathbf{1}) \mid a \in \mathbb{F}_0^\times$ or $a + \bar{a} = 0\,\}$.*

Proof. Every element of $O(W_0)$ extends uniquely to an element of $O(\Lambda_2 V)$ and so we may identify $O(W_0)$ with a subgroup of $O(\Lambda_2 V)$. The Dickson invariant $D\colon O(W_0) \to \mathbb{Z}_2$ restricts to the Dickson invariant of $O(W_0)$ and therefore we may regard $h \in SO(W_0)$ as an element of $SO(\Lambda_2 V)$. Then by (12.19) we may write $h = \kappa(a, f)$, where $(a, f) \in K(V)$. The commutator $[(1, g), (a, f)]$ belongs to $K(V)$ and is mapped to $\mathbf{1}$ by κ, therefore $[g, f] = b\,\mathbf{1}$ for some $b \in \mathbb{F}_0^\times$ such that $a^2 = b^2$. Thus $a = \pm b$ and therefore $(a, f) \in K_\gamma(V)$. It follows that $\kappa\colon K_\gamma(V) \to SO(W_0)$ is onto.

For $a \in \mathbb{F}$, we have $[g, a\,\mathbf{1}] = (a\bar{a})^{-1}\mathbf{1}$ and therefore $(a^{-2}, a\,\mathbf{1}) \in K_\gamma(V)$ if and only if $a \in \mathbb{F}_0^\times$ or $a + \bar{a} = 0$. □

If we identify $SU(V)$ with the subgroup $\{\, (1, f) \mid f \in SU(V)\,\}$ of $K_\gamma(V)$, then

12.35 Corollary. *The map κ restricts to a homomorphism from $SU(V)$ onto $\Omega(W_0)$ whose kernel is $\{\pm\mathbf{1}\}$.*

Proof. If $(a_1, f_1), (a_2, f_2) \in K_\gamma(V)$, then $[g, [f_1, f_2]] = \mathbf{1}$ and therefore $K_\gamma(V)' = SU(V)$. From Theorem 11.45, $SO(W_0)' = \Omega(W_0)$, thus κ maps $SU(V)$ onto $\Omega(W_0)$. It is clear that the kernel is $\{\pm\mathbf{1}\}$. □

12.36 Corollary. *$PSU(4, q) \simeq P\Omega^-(6, q)$ for all prime powers q.* □

Four-Dimensional Orthogonal Groups

It remains to describe the connections between the four-dimensional orthogonal groups and other linear groups. Rather than continue with the methods of the previous two sections we prefer to obtain the necessary homomorphisms by restricting κ to the stabilizers of various hyperbolic pairs. Other approaches are outlined in the exercises.

Suppose that (ξ, η) is a hyperbolic pair (with respect to Q) in $\Lambda_2 V$. Then there is a basis e_1, e_2, e_3, e_4 for V such that $\xi := e_1 \wedge e_2$, $\eta := e_3 \wedge e_4$ and $e_1 \wedge e_2 \wedge e_3 \wedge e_4 = \tilde{e}$. The restriction of Q to $W := \langle \xi, \eta \rangle^\perp$ is a non-degenerate quadratic form of Witt index 2 and we identify $O(W)$ with the subgroup of $O(\Lambda_2 V)$ that fixes ξ and η. Let $E := \langle e_1, e_2 \rangle$ and $F := \langle e_3, e_4 \rangle$, and identify $GL(E)$ (resp. $GL(F)$) with the subgroup of $GL(V)$ fixing E (resp. F) and acting on F (resp. E) as the identity. Then the stabilizer in $GL(V)$ of the pair (E, F) is $GL(E) \times GL(F)$. If we put

$$K(E,F) := \{ (f, f') \in GL(E) \times GL(F) \mid \det(f) = \det(f') \},$$

then the map $(f, f') \mapsto (\det(f)^{-1}, ff')$ identifies $K(E,F)$ with a subgroup of $K(V)$ and we have

12.37 Theorem. *The map $(f, f') \mapsto \det(f)^{-1}\Lambda_2(ff')$ is a homomorphism from $K(E,F)$ onto $SO(W)$ whose kernel is $\{ (a\,\mathbf{1}_E, a\,\mathbf{1}_F) \mid a \in \mathbb{F}^\times \}$.*

Proof. If $(f, f') \in K(E,F)$, then the transformation $\det(f)^{-1}\Lambda_2(ff') \in SO(\Lambda_2 V)$ fixes ξ and η, and thus belongs to $SO(W)$.

Conversely, if $h \in SO(W)$, then h may be regarded as an element of $SO(\Lambda_2 V)$ that fixes ξ and η. From (12.19), $h = \kappa(a, ff')$, where $f \in GL(E)$ and $f' \in GL(F)$. Then $\kappa(a, ff')\xi = a\det(f)\xi$, whence $\det(f) = a^{-1}$. Similarly, $\det(f') = a^{-1}$ and it follows that $h = \det(f)^{-1}\Lambda_2(ff')$.

It is clear that the kernel is $\{ (a\,\mathbf{1}_E, a\,\mathbf{1}_F) \mid a \in \mathbb{F}^\times \}$. □

12.38 Corollary. *Except when $\mathbb{F} = \mathbb{F}_2$, the map $SL(E) \times SL(F) \to \Omega(W)$ defined by $(f, f') \mapsto \Lambda_2(ff')$ is a homomorphism onto $\Omega(W)$ whose kernel is $\langle (-\mathbf{1}_E, -\mathbf{1}_F) \rangle$.*

Proof. From Theorem 11.45 $SO(W)' = \Omega(W)$ and, except when $\mathbb{F} = \mathbb{F}_2$, it follows from Theorem 4.4 that $K(E,F)' = SL(E) \times SL(F)$. □

12.39 Corollary. *Except for $q = 2$, $P\Omega^+(4,q) \simeq PSL(2,q) \times PSL(2,q)$ for all prime powers q.* □

Let g be the linear transformation $v \mapsto \gamma(-, v)$, where γ is the alternating form for which (e_1, e_3) and (e_2, e_4) are orthogonal hyperbolic pairs. Then E and F are totally isotropic with respect to γ and $\kappa(g)$ fixes ξ and η. From Theorem 12.31, $\kappa(g)$ is the reflection in $\langle \gamma_0 \rangle^\perp$, where $\gamma_0 := -e_1 \wedge e_3 - e_2 \wedge e_4$.

12.40 Lemma. *If $f \in GL(E)$, then $g \circ f \circ g^{-1} \in GL(F)$.*

Proof. Put $f' := g \circ f \circ g^{-1}$ ($= g^{-1} f^{*-1} g$). Then $\gamma(f(u), f'(v)) = \gamma(u, v)$ for all $u, v \in V$. As $E^{\perp} = E$ and $F^{\perp} = F$, it is clear that f' fixes both E and F. Moreover, for $u, u_1 \in E$ and $v_1 \in F$,

$$\begin{aligned}\gamma(f(u_1 + v_1), f'(u)) &= \gamma(f(v_1), f'(u))\\ &= \gamma(v_1, u)\\ &= \gamma(f(v_1), u),\end{aligned}$$

whence $f'(u) = u$ for all $u \in E$. Thus $f' \in GL(F)$. □

It follows from this lemma that g normalizes $K(E, F)$ and therefore we can form the semidirect product $K(E, F)\langle\, g\,\rangle$. The map of Theorem 12.37 extends to a surjective homomorphism

$$K(E, F)\langle\, g\,\rangle \to O(W).$$

The restriction of Q to the three-dimensional subspace $W \cap \langle\, \gamma_0\,\rangle^{\perp}$ is a non-degenerate quadratic form of Witt index 1 and the methods of Theorems 12.31 and 12.34 lead to a new proof of Theorem 11.6 (Exercise 12.16).

For the remainder of this section let γ be a non-degenerate σ-hermitian form of Witt index 2 on V and let $\mathbb{F}_0$ be the fixed field of σ. We shall use the notation introduced at the beginning of the previous section. Thus $g := \gamma(-, v)$ and W_0 is the vector space (over $\mathbb{F}_0$) of the fixed points of $\kappa(g)$. As before, $\tilde{e} := u_1 \wedge v_1 \wedge u_2 \wedge v_2$, where (u_1, v_1) and (u_2, v_2) are orthogonal hyperbolic pairs for γ. Then $E_0 := \langle\, u_1, v_2\,\rangle$ and $F_0 := \langle\, u_2, v_1\,\rangle$ are totally isotropic with respect to γ, $\xi := u_1 \wedge v_2$ and $\eta := v_1 \wedge u_2$ belong to W_0, and (ξ, η) is a hyperbolic pair for the restriction of Q to W_0. (In the previous section ξ and η were labelled ξ_5 and ξ_6, respectively.)

In this case the restriction of Q to $W_1 := W_0 \cap \langle\, \xi, \eta\,\rangle^{\perp}$ is a non-degenerate quadratic form of Witt index 1.

12.41 Theorem. *The map $(\varepsilon, f) \mapsto \det(f)\Lambda_2\big((\varepsilon f) g^{-1} f^{*-1} g\big)$ is a homomorphism from the group $\{\pm 1\} \times \{\, f \in GL(E_0) \mid \det(f) \in \mathbb{F}_0\,\}$ onto $SO(W_1)$ whose kernel is $\{\, (1, a\,\mathbf{1}) \mid a \in \mathbb{F}_0^{\times}\,\}$.*

Proof. Suppose that $h \in SO(W_1)$. As usual we may regard h as an element of $SO(W_0)$ that fixes ξ and η. Then from Theorem 12.34 we may write $h := \kappa(a, ff')$, where $a \in \mathbb{F}_0^{\times}$, $f \in GL(E_0)$, $f' \in GL(F_0)$, and $(a, ff') \in K_{\gamma}(V)$. Thus $[g, ff'] = \varepsilon a\,\mathbf{1}$, where $\varepsilon = \pm 1$, and as in Theorem 12.37 we have $\det(f) = \det(f') = a^{-1}$. The condition $[g, ff'] = \varepsilon a\,\mathbf{1}$ is equivalent to

$f'(v) := \varepsilon \det(f) g^{-1} f^{*-1} g(v)$, for all $v \in F_0$, and thus h has the required form. Conversely, every element of this form is in $SO(W_1)$. $\square$

If the characteristic of $\mathbb{F}$ is not 2, then $(-1, \mathbf{1})$ is mapped to $-\mathbf{1}_{W_1}$ and thus $-\mathbf{1}_{W_1} \notin \Omega(W_1)$—a result that can also be obtained by calculating the spinor norm.

12.42 Corollary. *The map $f \mapsto \Lambda_2(fg^{-1}f^{*-1}g)$ is a homomorphism from $SL(E_0)$ onto $\Omega(W_1)$. Its kernel is $\{\pm\mathbf{1}\}$.*

Proof. This follows immediately from the theorem by passing to the derived groups. $\square$

Finally, we have the last of the isomorphisms corresponding to the order coincidences of p. 166.

12.43 Corollary. *$P\Omega^-(4, q) \simeq PSL(2, q^2)$ for all prime powers q.* $\square$

Generalized Quadrangles and Duality

As pointed out by Todd (1970), one can obtain an isomorphism between $Sp(4, 2^k)$ and $O(5, 2^k)$ either by taking $m = 2$ in Theorem 11.9 or by taking $q = 2^k$ in Corollary 12.32, but the geometric reasons for these isomorphisms are quite different. We shall see that this situation yields simple constructions for an outer automorphism of $Sp(4, 2^k)$ (see also Duncan (1968)), and, when k is odd, a new family of finite simple groups, first discovered by Suzuki (1960). We follow the exposition of Tits (1960) and make extensive use of Lüneburg (1965).

Suppose that V is a four-dimensional vector space over a perfect field $\mathbb{F}$ of characteristic 2 and let γ be a non-degenerate alternating form on V. Let u_1, v_1, u_2, v_2 be a symplectic basis for V and put $\tilde{e} := u_1 \wedge v_1 \wedge u_2 \wedge v_2$. Let Q be the quadratic form defining the Klein quadric on $\Lambda_2 V$ and let β be its polar form. Then $\gamma_0 := u_1 \wedge v_1 + u_2 \wedge v_2$ is the bivector corresponding to γ and the restriction of Q to $W := \langle \gamma_0 \rangle^\perp$ is non-degenerate. In this case, β is an alternating form and the radical of W with respect to β is $\langle \gamma_0 \rangle$.

The alternating form $\bar{\beta}$ induced by β on $\overline{W} := W/\langle \gamma_0 \rangle$ is non-degenerate and therefore $\overline{W}$ is isometrically isomorphic to V. We need an explicit isomorphism and to this end we use the linear transformation $p\colon W \to V$ such that

$$\begin{array}{lll} p(u_1 \wedge u_2) := u_1, & p(v_1 \wedge v_2) := v_1, & \\ & & \text{and} \quad p(\gamma_0) := 0. \\ p(u_1 \wedge v_2) := u_2, & p(v_1 \wedge u_2) := v_2, & \end{array}$$

The map $\bar{p}\colon \overline{W} \to V$ induced by p is the required isomorphism.

Let $\mathcal{Q}$ be the generalized quadrangle of points and totally isotropic lines of $\mathcal{P}(V)$. If $L := \langle\, u, v \,\rangle$ is a totally isotropic line, then $\delta(L) := \langle\, p(u \wedge v) \,\rangle \in \mathcal{Q}$ and the discussion preceding Theorem 11.9 shows that δ is a bijection between the lines and points of $\mathcal{Q}$. Moreover, if P is a point of $\mathcal{Q}$, then $\kappa(P) \cap W = \kappa(P^\perp) \cap W$ is a totally singular line of $\mathcal{P}(W)$ and hence $\delta(P) := p(\kappa(P) \cap W)$ is a line of $\mathcal{Q}$. By construction, the point P is on the line L if and only if $\delta(L)$ is on $\delta(P)$. Thus δ is a bijection between the points and lines of $\mathcal{Q}$ which reverses the incidence relation.

If $f \in Sp(V)$, then $\Lambda_2 f \in O(W)$ and, as shown in Theorem 11.9, there is an induced map $\overline{\Lambda_2 f} \in Sp(\overline{W})$. Let $f' := \bar{p}(\overline{\Lambda_2 f})\bar{p}^{-1}$ be the corresponding element of $Sp(V)$. If $t \in Sp(V)$ is a transvection, then $\Lambda_2 t$ is a Siegel transformation and $\mathrm{im}(\mathbf{1} - \Lambda_2 t)$ is a totally singular line of $\mathcal{P}(W)$. Thus $\mathrm{im}(\mathbf{1} - t')$ is a totally isotropic line of $\mathcal{P}(V)$ and therefore t' is not a transvection. It follows that the automorphism $f \mapsto f'$ of $Sp(V)$ is not inner.

Given $P := \langle\, u \,\rangle \in \mathcal{P}(V)$, we have $\kappa(P) \cap W = \langle\, u \wedge v, u \wedge w \,\rangle$, where $P^\perp = \langle\, u, v, w \,\rangle$. Thus $\delta(P) = \langle\, p(u \wedge v), p(u \wedge w) \,\rangle$ and so

$$\delta^2(P) = \langle\, p\big(p(u \wedge v) \wedge p(u \wedge w)\big) \,\rangle.$$

For $u := x_1 u_1 + y_1 v_1 + x_2 u_2 + y_2 v_2$, a tedious but straightforward calculation shows that

$$p(u \wedge v) \wedge p(u \wedge w) = \\ \gamma(v, w)((x_1 y_1 + x_2 y_2)\gamma_0 + x_1^2 u_1 \wedge u_2 + y_1^2 v_1 \wedge v_2 + x_2^2 u_1 \wedge v_2 + y_2^2 v_1 \wedge u_2). \tag{12.44}$$

and hence

$$\delta^2(x_1 u_1 + y_1 v_1 + x_2 u_2 + y_2 v_2) = \gamma(u, v)(x_1^2 u_1 + y_1^2 v_1 + x_2^2 u_2 + y_2^2 v_2).$$

Thus $\delta^2(P) = \tau(P)$, where τ is the field automorphism $x \mapsto x^2$.

If $\mathbb{F}$ has an automorphism σ such that $\tau = \sigma^2$, then the map $\rho := \sigma^{-1}\delta$ satisfies $\rho^2 = \mathbf{1}$ and plays the rôle of a polarity for the symplectic generalized quadrangle $\mathcal{Q}$.

The Suzuki Groups

We retain the notation of the previous section and from now on suppose that $\mathbb{F}$ has an automorphism σ such that $\sigma^2(x) = x^2$ for all $x \in \mathbb{F}$. If $\mathbb{F}$ is finite, this is the case if and only if $\mathbb{F} = \mathbb{F}_{2^k}$, where k is odd: if $k = 2\ell + 1$, take $\sigma(x) := x^{2^{\ell+1}}$; if k is even, then $\mathbb{F}_{2^k}$ contains $\mathbb{F}_4$ and in this case it is easy to see that σ does not exist.)

The points $P \in \mathcal{Q}$ such that $P \in \rho(P)$ can be thought of as the "isotropic points" for the "polarity" ρ and we put

$$\mathcal{O} := \{\, P \in \mathcal{Q} \mid P \in \rho(P)\,\}.$$

This is the *Suzuki ovoid* introduced by Tits (1960). The group $Sp(V)$ acts faithfully on $\mathcal{Q}$ and we define the *Suzuki group* $Sz(\mathbb{F})$ to be the subgroup of $Sp(V)$ that leaves $\mathcal{O}$ invariant. The Coxeter-Dynkin diagram for $Sp(V)$ is of type B_2 and therefore the groups $Sz(\mathbb{F})$ are sometimes said to be of "twisted type" B_2 and denoted by ${}^2B_2(\mathbb{F})$.

We shall see that $Sz(\mathbb{F})$ acts doubly transitively on $\mathcal{O}$ and that our treatment of $Sz(\mathbb{F})$ bears a certain resemblance to the treatment of the three-dimensional unitary and orthogonal groups given in Chapters 10 and 11—the ovoid $\mathcal{O}$ playing the rôle of the isotropic (resp. singular) points of $\mathcal{P}(V)$.

If $P := \langle\, u\,\rangle$ and if u, v, w is a basis for $P^\perp$, then $P \in \mathcal{O}$ if and only if $\sigma(u) \wedge p(u \wedge v) \wedge p(u \wedge w) = 0$. When $u := x_1u_1 + y_1v_1 + x_2u_2 + y_2v_2$, we see from (12.44) that $P \in \mathcal{O}$ if and only if

$$\begin{aligned}
y_1^2\sigma(x_1) + x_2^2\sigma(y_1) + x_1y_1\sigma(y_2) + x_2y_2\sigma(y_2) &= 0,\\
x_1^2\sigma(y_1) + y_2^2\sigma(x_1) + x_1y_1\sigma(x_2) + x_2y_2\sigma(x_2) &= 0,\\
x_1^2\sigma(y_2) + x_2^2\sigma(x_2) + x_1y_1\sigma(x_1) + x_2y_2\sigma(x_1) &= 0, \quad \text{and}\\
y_1^2\sigma(x_2) + y_2^2\sigma(y_2) + x_1y_1\sigma(y_1) + x_2y_2\sigma(y_1) &= 0.
\end{aligned} \tag{12.45}$$

If $y_1 = 0$, these equations show that $x_2 = y_2 = 0$ and consequently $\langle\, u_1\,\rangle^\perp \cap \mathcal{O} = \langle\, u_1\,\rangle$. If $y_1 \neq 0$, put $x := y_2/y_1$, $y := x_2/y_1$, and $z := x_1/y_1$. Then the equations of (12.45) reduce to the single condition

$$z + xy + \sigma(y) + x^2\sigma(x) = 0. \tag{12.46}$$

Thus

$$\mathcal{O} = \langle\, u_1\,\rangle \cup \{\, \langle\, zu_1 + yu_2 + xv_2 + v_1\,\rangle \mid z + xy + \sigma(y) + x^2\sigma(x) = 0\,\}.$$

In particular, if $\mathbb{F} = \mathbb{F}_q$, then $|\mathcal{O}| = q^2 + 1$.

The matrix of the form γ with respect to the basis u_1, u_2, v_2, v_1, is

$$J := \begin{pmatrix} 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0 \end{pmatrix}$$

and we shall represent the elements of $Sz(\mathbb{F})$ by matrices with respect to this same basis. Our immediate aim is determine the stabilizer of $\langle\, u_1\,\rangle$ in $Sz(\mathbb{F})$. In order to achieve this we need the following lemma of Artin.

12.47 Lemma. *Distinct functions $\chi_1, \chi_2, \dots, \chi_k$ from $\mathbb{F}$ to itself such that $\chi_i(xy) = \chi_i(x)\chi_i(y)$ for all i and for all $x, y \in \mathbb{F}$ are linearly independent.*

Proof. Suppose that i is minimal subject to $a_1\chi_1 + a_2\chi_2 + \cdots + a_i\chi_i = 0$, where $a_1, a_2, \dots\ a_i \in \mathbb{F}$ and $a_i \neq 0$. For every $x \in \mathbb{F}$ we have

$$a_1\chi_1(x)\chi_1 + a_2\chi_2(x)\chi_2 + \cdots + a_i\chi_i(x)\chi_i = 0$$

and therefore

$$a_1(\chi_i(x) - \chi_1(x))\chi_1 + \cdots + a_{i-1}(\chi_i(x) - \chi_{i-1}(x))\chi_{i-1} = 0.$$

If j is the largest integer less than i such that $a_j \neq 0$, choose x so that $\chi_i(x) \neq \chi_j(x)$. This contradicts the minimality of i. $\square$

12.48 Lemma. *Except when $\mathbb{F} = \mathbb{F}_2$, an element of $Sz(\mathbb{F})$ fixes $\langle u_1 \rangle$ and $\langle v_1 \rangle$ if and only if it is represented by a matrix of the form*

$$H(k) := \begin{pmatrix} k\sigma(k) & 0 & 0 & 0 \\ 0 & k & 0 & 0 \\ 0 & 0 & k^{-1} & 0 \\ 0 & 0 & 0 & k^{-1}\sigma(k)^{-1} \end{pmatrix},$$

where $k \in \mathbb{F}^\times$.

Proof. An element $f \in Sp(V)$ that fixes both $\langle u_1 \rangle$ and $\langle v_1 \rangle$ is represented by a matrix of the form

$$\begin{pmatrix} e & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & c & d & 0 \\ 0 & 0 & 0 & e^{-1} \end{pmatrix},$$

where $ad + bc = 1$. The image of $\langle \sigma(y)u_1 + yu_2 + v_1 \rangle \in \mathcal{O}$ under f is

$$\langle e\sigma(y)u_1 + ayu_2 + cyv_2 + e^{-1}v_1 \rangle$$

and therefore, from (12.46), if f fixes $\mathcal{O}$, we have

$$(e^2 + \sigma(ae))\sigma(y) + (ace^2)y^2 + (c^2\sigma(c)e^2\sigma(e))y^2\sigma(y) = 0$$

for all $y \in \mathbb{F}$. If $\mathbb{F} \neq \mathbb{F}_2$, the multiplicative functions σ, $y \mapsto y^2$ and $y \mapsto y^2\sigma(y)$ are distinct and therefore the previous lemma shows that $c = 0$ and $e^2 = \sigma(ae)$.

Similarly, the image of $\langle x^2\sigma(x)u_1 + xv_2 + v_1 \rangle$ under f is

$$\langle ex^2\sigma(x)u_1 + bxu_2 + dxv_2 + e^{-1}v_1 \rangle,$$

and hence

$$\sigma(be)\sigma(x) + (bde^2)x^2 + (e^2 + d^2\sigma(d)e^2\sigma(e))x^2\sigma(x) = 0$$

for all $x \in \mathbb{F}$. Therefore $b = 0$, $d^2\sigma(de) = 1$ and then $d = a^{-1}$ and $\sigma(e) = a^2\sigma(a)$. Applying σ and using the fact that $x \mapsto x^2$ is an automorphism of $\mathbb{F}$ yields $e = a\sigma(a)$ and thus f is represented by $H(a)$.

Conversely, it is a simple matter to check that every element of $Sp(V)$ represented by a matrix of the form $H(k)$ is in $Sz(\mathbb{F})$ and fixes $\langle u_1 \rangle$ and $\langle v_1 \rangle$. $\square$

The next step is to determine the stabilizer of $\langle u_1 \rangle$. For a, $b \in \mathbb{F}$ we put

$$T(a,b) := \begin{pmatrix} 1 & a & a\sigma(a)+b & ab+\sigma(b)+a^2\sigma(a) \\ 0 & 1 & \sigma(a) & b \\ 0 & 0 & 1 & a \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and note that

$$T(a,b)T(c,d) = T(a+c, b+d+c\sigma(a)). \tag{12.49}$$

It is easily checked that $T(a,b)^t JT(a,b) = J$ and therefore $T(a,b)$ represents an element of $Sp(V)$. Let T be the subgroup of $Sp(V)$ whose elements are represented by the matrices $T(a,b)$.

12.50 Lemma. *The group T fixes $\langle u_1 \rangle$ and acts regularly on $\mathcal{O} \setminus \{\langle u_1 \rangle\}$.*

Proof. If $P := \langle zu_1 + yu_2 + xv_2 + v_1 \rangle$ is a point of $\mathcal{O}$, then from (12.46) the image of P under $T(a,b)$ is also a point of $\mathcal{O}$. Moreover, $T(a,b)$ maps $\langle v_1 \rangle$ to $\langle cu_1 + bu_2 + av_2 + v_1 \rangle$, where $c := ab + \sigma(b) + a^2\sigma(a)$, and thus T acts regularly on $\mathcal{O} \setminus \{\langle u_1 \rangle\}$. $\square$

Let H be the subgroup of $Sp(V)$ whose elements are represented by the matrices $H(k)$. Another simple matrix calculation shows that

$$H(k)T(a,b)H(k)^{-1} = T(\sigma(k)a, k^2\sigma(k)b). \tag{12.51}$$

and therefore H normalizes T.

From (12.49) we have $T(a,b)^{-1} = T(a, a\sigma(a)+b)$ and therefore

$$[H(k), T(a,b)] = T((1+\sigma(k))a, (1+k^2\sigma(k))b + (1+k^2)a\sigma(a)).$$

If $\mathbb{F} \neq \mathbb{F}_2$, it follows that every element of T is a commutator and therefore $(TH)' = T$.

12.52 Corollary. $Sz(\mathbb{F})_{\langle u_1 \rangle} = TH$.

Proof. If $f \in Sz(\mathbb{F})$ fixes $\langle u_1 \rangle$, then for some $t \in T$ we have $f(v_1) = t(v_1)$. Thus $t^{-1}f$ fixes $\langle u_1 \rangle$ and $\langle v_1 \rangle$ and by Lemma 12.48, it belongs to H. Therefore, $f \in TH$. □

In order to complete the picture we need an element of $Sz(\mathbb{F})$ that moves $\langle u_1 \rangle$. We let w be the element of $Sp(V)$ represented by the matrix J. Then w maps $x_1u_1+x_2u_2+y_1v_1+y_2v_2$ to $y_1u_1+y_2u_2+x_1v_1+x_2v_2$. Interchanging the symbols x and y in (12.45) swaps the first and second and also the third and fourth equations. Thus w preserves $\mathcal{O}$; i.e., $w \in Sz(\mathbb{F})$. By construction w interchanges $\langle u_1 \rangle$ and $\langle v_1 \rangle$ and for all $k \in \mathbb{F}^\times$ we have

$$wH(k)w^{-1} = H(k)^{-1}. \tag{12.53}$$

12.54 Theorem. *If $\mathbb{F} \neq \mathbb{F}_2$, the group $Sz(\mathbb{F})$ acts doubly transitively on the points of $\mathcal{O}$ and only the identity element fixes more than 2 points.*

Proof. From what we have proved so far it is clear that $Sz(\mathbb{F}) = \langle w, T, H \rangle$ acts doubly transitively on $\mathcal{O}$. Thus if $f \in Sz(\mathbb{F})$ fixes two points, we may suppose the points to be $\langle u_1 \rangle$ and $\langle v_1 \rangle$. Then Lemma 12.48 shows that $f = H(k)$ for some k. Now a matrix calculation shows that f fixes a further point if and only if $k = 1$. □

Suppose that $q := 2^k$, where $k > 1$ is odd. Then $|T| = q^2$, $|H| = q-1$, and therefore

$$|Sz(q)| = q^2(q-1)(q^2+1).$$

12.55 Lemma. *If $\mathbb{F} \neq \mathbb{F}_2$, then $Sz(\mathbb{F})$ has a single conjugacy class of elements of order 2, and these elements generate $Sz(\mathbb{F})$.*

Proof. It follows from (12.53) that $(H(k)w)^2 = \mathbf{1}$ and thus $H(k)$ is the product of two elements of order 2. We noted above that every element of T is a commutator $[H(k), T(a,b)]$ for some k, a, $b \in \mathbb{F}$ and therefore T is contained in the subgroup generated by the elements of order 2. We also have $w^2 = 1$ and therefore $Sz(\mathbb{F})$ is generated by the elements of order 2.

It remains to show that $Sz(\mathbb{F})$ has a single conjugacy class of elements of order 2. So suppose that $\mathbf{1} \neq f \in Sz(\mathbb{F})$ and $f^2 = \mathbf{1}$. Then f interchanges two points, which by double transitivity we may take to be $\langle u_1 \rangle$ and $\langle v_1 \rangle$. Thus $f \in H\langle w \rangle$ and we may suppose that $f := H(k)w$ for some $k \in \mathbb{F}$. The map $x \mapsto x^2$ is an automorphism of $\mathbb{F}$ and therefore there exists $\ell \in \mathbb{F}$ such that $k = \ell^2$ and then $f = H(\ell)wH(\ell)^{-1}$. Thus all elements of order 2 in $Sz(F)$ are conjugate. □

12.56 Theorem. *Except when $\mathbb{F} = \mathbb{F}_2$, the groups $Sz(\mathbb{F})$ are simple.*

Proof. The group $Sz(\mathbb{F})$ is doubly transitive on $\mathcal{O}$ and *a fortiori* it is primitive.

The product formula (12.49) shows that

$$[T(a,b), T(c,d)] = T(0, a\sigma(c) + c\sigma(a))$$

and thus T' is an abelian normal subgroup of $Sz(\mathbb{F})_{\langle u_1 \rangle}$ which contains an element of order 2. It follows from lemma 12.55 that the conjugates of T' generate $Sz(\mathbb{F})$ and that $Sz(\mathbb{F})' = Sz(\mathbb{F})$. The simplicity of $Sz(\mathbb{F})$ is thus a consequence of Iwasawa's criterion. □

If $\mathbb{F} \neq \mathbb{F}_2$, then $T' = \{ T(0,b) \mid b \in \mathbb{F} \}$. Thus every element of $T' \setminus \{\mathbf{1}\}$ has order 2 and every element of T of order 2 belongs to T'.

For $\mathbb{F} \neq \mathbb{F}_2$, the subgroup $Sz(\mathbb{F})$ of $Sp(V)$ that fixes $\mathcal{O}$ coincides with the subgroup that commutes with the action of the "polarity" ρ. We showed in Chapter 8 that $Sp(4,2) \simeq S_6$ and it is not difficult to check that the subgroup of $Sp(4,2)$ fixing $\mathcal{O}$ is S_5. But in this case, it is more usual to define $Sz(2)$ to be the subgroup of order 20 in S_5 that commutes with ρ.

EXERCISES

12.1 Show directly that the varieties of the oriflamme geometry of an orthogonal geometry of dimension 6 and Witt index 3 satisfy the axioms for projective geometry given in Chapter 3.

12.2 Let V be a vector space over a field $\mathbb{F}$ and suppose that there is a linear transformation $\beta : V \to B$, where B is an associative algebra and $\beta(v)^2 = 0$ for all $v \in V$. Suppose that for every associative algebra A and every linear transformation $h : V \to A$ such that $h(v)^2 = 0$ for all $v \in V$, there is a unique algebra homomorphism $\widehat{h} : B \to A$ such that the diagram

$$\begin{tikzcd} V \arrow[r, "\beta"] \arrow[dr, "h"'] & B \arrow[d, "\widehat{h}"] \\ & A \end{tikzcd}$$

commutes. Show that B is isomorphic to $\Lambda_\bullet V$.

12.3 (*i*) Show that for every k-vector $\xi \in \Lambda_k V$, there is a smallest subspace M_ξ such that $\xi \in \Lambda_k M_\xi$. Then show that ξ is decomposable if and only if $\dim M_\xi = k$.

(*ii*) Let N_ξ be the set of vectors v such that $v \wedge \xi = 0$. Show that $N_\xi \subseteq M_\xi$ and that equality holds if and only if ξ is decomposable.

(*iii*) If $\dim V = n$, show that every element of $\Lambda_{n-1} V$ is decomposable.

12.4 If V is a vector space and $\varphi \in V^*$, show that for all $v_1, v_2, \ldots, v_k \in V$,

$$\alpha^-_\varphi(v_1 \wedge \cdots \wedge v_k) = \sum_{i=1}^{k} (-1)^{i+1} \varphi(v_i) v_1 \wedge \cdots \wedge v_{i-1} \wedge v_{i+1} \wedge \cdots \wedge v_k.$$

12.5 Suppose that V is a vector space of dimension $2m$ over $\mathbb{F}$ and that Q is a non-degenerate quadratic form of Witt index m on V. Let E and F be maximal totally singular subspaces such that $V = E \oplus F$ and let A be the algebra of all linear transformations of $\Lambda_\bullet V$. If $x := e + f$, where $e \in E$ and $f \in F$, define

$$\alpha(x) := \alpha^+_e + \alpha^-_{\beta(-,f)},$$

where β is the polar form of Q. Show that $\alpha : V \to A$ is one-to-one and that $\alpha(x)^2 = Q(x)\,\mathbf{1}$ for all $x \in V$.

12.6 Suppose that $\dim V = 4$. Show that

$$\kappa^{-1}(\Omega(\Lambda_2 V)) = \langle\ g^2 \mid g \in K(V)\ \rangle,$$

where $\kappa : K(V) \to SO(\Lambda_2 V)$ is the homomorphism defined at the end of the section 'The Klein Quadric'.

12.7 Let γ be a non-degenerate σ-sesquilinear form on V and let π denote the corresponding correlation of $\mathcal{P}(V)$. Show that the collineation of $\mathcal{P}(\Lambda_2 V)$ induced by π via the Klein correspondence coincides with the collineation induced by the σ-semilinear map $\kappa(g)$ defined in the section 'Correlations'.

12.8 (Pfaffians) Suppose that e_1, e_2, e_3, e_4 is a basis for the vector space V and let $\xi := \sum_{i<j} p_{ij} e_i \wedge e_j$. For $1 \le i \le 4$, define $p_{ii} := 0$, and for $1 \le i < j \le 4$, define $p_{ji} := -p_{ij}$. Show that $Q(\xi)^2 = \det(p_{ij})$.

12.9 Suppose that $\dim V = 4$ and let $\tilde{e}$ be a non-zero element of $\Lambda_4 V$. For $a \in \mathbb{F}^\times$, let $\tilde{e}' := a\tilde{e}$ and let Q', β', d' and κ' be the functions, defined via $\tilde{e}'$, corresponding to the functions Q, β, d and κ, defined using $\tilde{e}$. Show that $Q' = a^{-1}Q$, $\beta' = a^{-1}\beta$ and for every σ-semilinear bijection $f : V \to V^*$ we have $d'(f) = a\sigma(a)\, d(f)$ and $\kappa'(f) = a\,\kappa(f)$.

12.10 (*i*) Show that for all f, $g \in GL^*(V)$,

$$d(f \circ g) = \begin{cases} d(f)d(g) & \text{if } g : V \to V \\ d(f)^{-1}d(g) & \text{if } g : V \to V^*. \end{cases}$$

(*ii*) Check that the multiplication defined on $K^*(V)$ is associative and that the inverse of $(a, f) \in K^*(V)$ is (a^{-1}, f^{-1}) or (a, f^*), according to whether $f : V \to V$ or $f : V \to V^*$.

(*iii*) Check that $\kappa : K^*(V) \to O(\Lambda_2 V)$ is a homomorphism.

12.11 Suppose that V is a vector space of dimension 3 over $\mathbb{F}$ and that σ is an automorphism of $\mathbb{F}$ satisfying $\sigma^2 = 1$. Let β be non-degenerate σ-sesquilinear form such that $\beta(u, v) = \sigma\beta(v, u)$ for all u, $v \in V$. That is, β is either σ-hermitian or symmetric. Let $\tilde{e}$ be a fixed non-zero element of $\Lambda_3 V$. If e_1, e_2, e_3 is a basis for V such that $e_1 \wedge e_2 \wedge e_3 = \tilde{e}$, show that $d := \det\big(\beta(e_i, e_j)\big)$ depends only on β and $\tilde{e}$, not on the basis. Note that $\sigma(d) = d$.

Show that for v, $w \in V$, there is a unique vector $v \times w \in V$ such that

$$u \wedge v \wedge w = \beta(u, v \times w)\,\tilde{e}$$

for all $u \in V$. Then show that the product $v \times w$ has the following properties:

(*i*) $(u + v) \times w = u \times w + v \times w$ and $u \times (v + w) = u \times v + u \times w$.

(*ii*) $(au) \times v = u \times (av) = \sigma(a)(u \times v)$.

(*iii*) $u \times u = 0$.

(*iv*) $\beta(u, u \times v) = \beta(v, u \times v) = 0$.

(*v*) $\alpha^{-}_{\beta(-,v)}\alpha^{-}_{\beta(-,u)}(\tilde{e}) = d\, u \times v$.

(*vi*) $d\,\beta(x \times u, v \times w) = \beta(v, x)\beta(w, u) - \beta(v, u)\beta(w, x)$.

(*vii*) $d\,(u \times v) \times w = \beta(u, w)\, v - \beta(v, w)\, u$.

(*viii*) If $\sigma = 1$, then $(u \times v) \times w + (v \times w) \times u + (w \times u) \times v = 0$.

(*ix*) If $f \in GL(V)$ and $f^\perp$ is defined by $\beta(f^\perp(u), f(v)) = \beta(u, v)$ for all u, $v \in V$, then $f(u) \times f(v) = \sigma(\det(f))\, f^\perp(u \times v)$.

(*x*) If $\sigma \neq 1$, then $SU(V) = \{\, f \in GL(V) \mid f(u \times v) = f(u) \times f(v) \,\}$.

(xi) If $\sigma = 1$ and the characteristic of f is not 2, then

$$SO(V) = \{\, f \in GL(V) \mid f(u \times v) = f(u) \times f(v) \,\}.$$

(xii) If $\sigma = 1$ and $\mathbb{F}$ is a perfect field of characteristic 2, then $W := \{\, v \in V \mid \beta(v,v) = 0 \,\}$ is a non-degenerate subspace of dimension 2, and

$$Sp(W) = \{\, f \in GL(V) \mid f(u \times v) = f(u) \times f(v) \,\}.$$

12.12 Let V be a vector space of dimension 4 over the field $\mathbb{F}$ and let γ be a non-degenerate alternating form on V. Let γ_0 be the non-singular bivector corresponding to γ. Show that, under the Klein correspondence, the symplectic transvections (with respect to γ) correspond to Siegel transformations $\rho_{\xi,\eta}$, where $\langle\, \xi, \eta \,\rangle$ is a totally singular subspace of $\langle\, \gamma_0 \,\rangle^{\perp}$.

12.13 State and prove the result corresponding to Theorem 12.34 for non-degenerate alternating forms.

12.14 Let V be a vector space of dimension 4 over the field $\mathbb{F}$ and let γ be a non-degenerate σ-hermitian form on V. Let $\mathbb{F}_0$ be the fixed field of σ and let W_0 be the $\mathbb{F}_0$-space of fixed points of $\kappa(g)$ in $\Lambda_2 V$, where $g : v \mapsto \gamma(-, v)$. Show that, under the Klein correspondence, the unitary transvections (with respect to γ) correspond to the Siegel transformations $\rho_{\xi,\eta}$, where $\langle\, \xi, \eta \,\rangle$ is a totally singular subspace of W_0.

12.15 Let V be a vector space over $\mathbb{C}$ with basis e_1, e_2, e_3, e_4 and define an hermitian form γ on V by $\gamma(\sum x_i e_i, \sum y_i e_i) := \sum x_i \bar{y}_i$. For $v \in V$, let $g(v) := \gamma(-, v)$ and let W_0 be the (real) subspace of $\Lambda_2 V$ of fixed points of $\kappa(g)$. Show that there is a basis ξ_1, ξ_2, $\ldots$, ξ_6 for W_0 such that $Q(\sum x_i \xi_i) = \sum x_i^2$.

12.16 Use the Klein correspondence to prove Theorem 11.6.

12.17 Verify equation (12.44).

12.18 Using the notation of Theorem 12.37, define $p \in SL(V)$ by $p(e_1) := e_3$, $p(e_2) := e_4$, $p(e_3) := -e_1$, and $p(e_4) := -e_2$. Show that

(i) p interchanges E and F, and $\Lambda_2 p$ interchanges ξ and η.

(ii) The restriction of $\Lambda_2 p$ to W is a reflection.

(iii) $O(W)$ is a homomorphic image of the group $K(E, F)\langle\, p \,\rangle$.

12.19 Suppose that V is four-dimensional vector space over a field $\mathbb{F}$ and that the characteristic of $\mathbb{F}$ is not 2. Let γ be a non-degenerate symmetric bilinear form of Witt index 2 on V and for $v \in V$ define $g(v) := \gamma(-, v)$. Choose a basis u_1, v_1, u_2, v_2 for V such that $\langle u_1, v_1 \rangle = \langle u_2, v_2 \rangle^\perp$, and (u_1, v_1) and (u_2, v_2) are hyperbolic pairs for γ and set $\tilde{e} := u_1 \wedge v_1 \wedge u_2 \wedge v_2$.

(i) Show that $\kappa(g)$ is an element of order 2 and that $\Lambda_2 V$ is the orthogonal sum $W_+ \perp W_-$, where $\kappa(g)$ acts as $\mathbf{1}$ on W_+ and as $-\mathbf{1}$ on W_-. Furthermore, $\dim W_+ = \dim W_- = 3$.

(ii) Show that the restriction of Q to W_+ (and to W_-) is a non-degenerate quadratic form of Witt index 1.

(iii) State and prove the result corresponding to Theorem 12.34.

(iv) Show that, under the Klein correspondence, the dual of the generalized quadrangle of isotropic points and totally isotropic lines of $\mathcal{P}(V)$ is the complete bipartite graph whose two classes of vertices are the singular points of $\mathcal{P}(W_+)$ and $\mathcal{P}(W_-)$, respectively.

12.20 (Quaternion algebras) Let V be a three-dimensional vector space over the field $\mathbb{F}$. Suppose that the characteristic of $\mathbb{F}$ is not 2 and let β be a non-degenerate symmetric bilinear form on V. Put $\tilde{e} := e_1 \wedge e_2 \wedge e_3$, where e_1, e_2, e_3 is a basis for V, and let $d := \det\big(\beta(e_i, e_j)\big)$. For $a \in \mathbb{F}$ and $u \in V$, let $a + u$ denote the pair $(a, u) \in \mathbb{H} := \mathbb{F} \times V$. Define a multiplication on $\mathbb{H}$ by

$$(a + u)(b + v) := ab - d^{-1}\beta(u, v) + av + bu + u \times v,$$

where $u \times v$ is defined in Exercise 12.11.

(i) For $a \in \mathbb{F}$, $u \in V$ and $q := a + u$, define $\bar{q} := a - u$. Show that $\overline{qr} = \bar{r}\bar{q}$ for all q, $r \in \mathbb{H}$.

(ii) Show that $Q(q) := q\bar{q}$ ($= \bar{q}q$) defines a non-degenerate quadratic form on $\mathbb{H}$ such that $Q(qr) = Q(q)Q(r)$ for all q, $r \in \mathbb{H}$. Show also that $q^2 - aq + Q(q) = 0$, where $a := q + \bar{q} \in \mathbb{F}$.

(iii) For $q \in \mathbb{H}$, let $L(q)$ be the map $h \mapsto qh$, and let $R(q)$ be the map $h \mapsto hq$, where $h \in \mathbb{H}$. Show that $\det L(q) = \det R(q) = Q(q)^2$.

(iv) Show that $\widehat{\Omega} := \{\, h \in \mathbb{H} \mid Q(h) \neq 0 \,\}$ is a group and that for $h \in \widehat{\Omega}$, the map $B(h) := L(h)R(h^{-1})$ leaves V invariant and belongs to $SO(V)$.

(v) If $v \in V$ and $Q(v) \neq 0$, show that $-B(v)$ is the reflection in $\langle v \rangle^\perp$. Deduce that $B : \widehat{\Omega} \to SO(V)$ is a homomorphism onto $SO(V)$ whose kernel is $\mathbb{F}^\times$.

(*vi*) If θ is the spinor norm, show that $\theta(B(h)) \equiv Q(h) \pmod{\mathbb{F}^2}$. Then show that B maps $\Omega := \{\, h \in \mathbb{H} \mid Q(h) = 1 \,\}$ onto $\Omega(V)$ and that its kernel is $\{\pm 1\}$.

(*vii*) Show that the map $(q, r) \mapsto L(q)R(r^{-1})$ is a homomorphism from the group $\{\, (q, r) \in \widehat{\Omega} \times \widehat{\Omega} \mid Q(q) = Q(r) \,\}$ onto $SO(\mathbb{H})$ whose kernel is $\{\, (a, a) \mid a \in \mathbb{F}^\times \,\}$.

(*viii*) Suppose that for some $\omega \in V$, $-Q(\omega)$ does not have a square root in $\mathbb{F}$ and show that $K := \mathbb{F} \oplus \mathbb{F}\omega$ is a field. Let $\overline{\mathbb{H}}$ denote $\mathbb{H}$ considered as a two-dimensional vector space over K. Define

$$\gamma(q, r) := q\bar{r} - Q(\omega)^{-1}\omega q \bar{r} \omega$$

and show that γ is an hermitian form on $\mathbb{H}$ such that $\gamma(q, q) = 2Q(q)$. Then show that the map $\Omega \to SU(\overline{\mathbb{H}}) : h \mapsto R(h^{-1})$ is an isomorphism.

(*ix*) If $\mathbb{H}$ contains singular vectors, show that it is isomorphic to the algebra of 2×2 matrices over $\mathbb{F}$.

12.21 Let W be a vector space of dimension 2 over $\mathbb{F}$ and let V be the vector space of linear transformations $h : W \to W$.

(*i*) Show that $Q(h) := \det(h)$ defines a non-degenerate quadratic form of Witt index 2 on V and let β be its polar form.

(*ii*) For f_1, $f_2 \in GL(W)$, let $\lambda(f_1, f_2)$ be the linear transformation $h \mapsto f_1 h f_2^{-1}$. Show that λ defines a homomorphism from the group $K(W, W)$ of Theorem 12.37 onto $SO(V)$.

12.22 Suppose that V is a four-dimensional vector space over a perfect field $\mathbb{F}$ of characteristic 2 and that $\mathbb{F}$ has an automorphism σ such that $\sigma^2(x) = x^2$ for all x. Let ρ be the "polarity" of the symplectic generalized quadrangle derived from V. If $\mathbb{F} \neq \mathbb{F}_2$, show that the subgroup of $Sp(V)$ that commutes with ρ coincides with the subgroup that leaves the Suzuki ovoid invariant.

12.23 Show that the subgroup of $Sp(4, 2)$ that fixes the Suzuki ovoid is isomorphic to the symmetric group S_5 and that the subgroup commuting with the "polarity" ρ has order 20.

12.24 Show that if $\mathbb{F} \neq \mathbb{F}_2$, the subgroups $B := TH$ and $N := H\langle\, w \,\rangle$ form a BN-pair for the Suzuki group $Sz(\mathbb{F})$ with Weyl group $\langle\, w \,\rangle$.

12.25 Show that, except for the Suzuki groups, the order of every finite simple group introduced in this book is divisible by 3.

Bibliography

E. Artin

The orders of the linear groups, *Comm. Pure Appl. Math.* **8**, (1955), 355–365. MR 17–12

The orders of the classical simple groups, *Comm. Pure Appl. Math.* **8** (1955), 455–472. MR 17–1457

Geometric algebra, Interscience 1957. MR 18–553

M. Aschbacher

Finite Group Theory, Cambridge Studies in Advanced Mathematics 10, Cambridge University Press 1986. MR 89b:20001

R. Baer

Linear algebra and projective geometry, Academic Press 1952. MR 14–675

G. Birkhoff and J. von Neumann

The logic of quantum mechanics, *Ann. of Math.* (2) **37** (1936), 823–843.

N. Bourbaki

Algèbre. Chapitre 9. Formes sesquilinéaires et formes quadratiques, Actualités Sci. et Ind. No 1272, Hermann 1959.

Groupes et algèbres de Lie. Chapitres IV, V, VI., Actualités Sci. et Ind. No 1337, Hermann 1968. MR 39#1590

Algèbre. Chapitres 1 à 3, Hermann 1970. MR 43#2

R. Brauer

A characterization of null systems in projective space, *Bull. Amer. Math. Soc.* **42** (1936), 247–254.

K. S. Brown

Buildings, Springer-Verlag 1989.

F. Bruhat

Représentations induites des groupes de Lie semi-simples connexes, *C. R. Acad. Sci. Paris* **238** (1954), 437–439. MR 15–504

F. Buekenhout

Une charactérisation des espaces affins basée sur la notion de droite, *Math. Z.* **111** (1969), 367–371. MR 42#8383

A characterization of polar spaces, *Simon Stevin* **53** (1979), 3–7. MR 82e:51025

The geometry of diagrams, pp. 69–75 of *Proc. Sympos. Pure Math.* XXXIV, Amer. Math. Soc., Providence R.I. 1979. MR 80b:51005

The geometry of diagrams, *Geom. Dedicata* **8** (1979), 253–257. MR 83c:51020

Diagrams for geometries and groups, *J. Combin. Theory Ser. A* **27** (1979), 121–151. MR 83f:51003

The basic diagram of a geometry, pp. 1–29 of *Geometries and Groups (Proceedings, Berlin 1981)* (ed. M. Aigner and D. Jungnickel), Lecture Notes in Mathematics, **893**. Springer-Verlag 1981. MR 83f:51010

An approach to buildings based on points, lines and convexity, *European J. Combin.* **3** (1982), 103–118. MR 83j:51011

The geometry of the finite simple groups, pp. 1–78 of *Buildings and the Geometry of Diagrams* (ed. L. A. Rosati), Lecture Notes in Mathematics, **1181**. Springer-Verlag 1986. MR 87h:20032

F. Buekenhout and E. E. Shult
On the foundations of polar geometry, *Geom. Dedicata* **3** (1974), 155–170. MR 50#3091

W. Burnside
Theory of groups of finite order, Dover 1955; 2nd edition. MR 16–1086

D. Callan
The generation of $Sp(\mathbb{F}_2)$ by transvections, *J. Algebra* **42**, (1976), 378–390. MR 54#7652

P. J. Cameron
Parallelisms of complete designs, London Math. Soc. Lecture Note Series, **23**. Cambridge University Press 1976. MR 54#7269

P. J. Cameron and W. M. Kantor
2-transitive and antiflag transitive collineation groups of finite projective spaces, *J. Algebra* **60** (1979), 384–422. MR 81c:20032

E. Cartan
La géométrie des groupes simples, *Ann. Mat. Pura Appl.* (4) (1927), 209–256.

R. W. Carter
Simple groups of Lie type, Pure and Applied Mathematics, Vol. 28. Wiley 1972. MR 53#10946

Some aspects of the representation theory of finite groups of Lie type, Oxford Lecture Notes 1977.

Finite groups of Lie type (*Conjugacy classes and complex characters*), Wiley 1985.

R. W. Carter and P. Fong
The Sylow 2-subgroups of the finite classical groups, *J. Algebra* **1** (1964), 139–151. MR 29#3548

C. Chevalley
The algebraic theory of spinors, Columbia University Press 1954. MR 15–678
Sur certain groupes simples, *Tôhoku Math. J.* (2) **7** (1955), 14–66. MR 17–457

J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker and R. A. Wilson
Atlas of Finite Groups. Maximal subgroups and ordinary characters for simple groups, Oxford University Press 1985. MR 889:20025

H. S. M. Coxeter
The complete enumeration of finite groups of the form $R_i^2 = (R_iR_j)^{k_{ij}} = 1$, *J. London Math. Soc.* **10** (1935), 21–25.

A. Crumeyrolle

Orthogonal and Symplectic Clifford Algebras – Spinor Structures, Mathematics and its Applications, Volume 57. Kluwer Academic Publishers 1990.

P. Dembowski

Finite geometries, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 44. Springer-Verlag 1968. MR 38#1597

J. B. Derr

Stabilizers of isotropic subspaces in classical groups, *Arch. Math. (Basel)* **34** (1980), 100–107. MR 82e:20003

L. E. Dickson

Linear groups with an exposition of the Galois field theory, Teubner 1901.

J. Dieudonné

Sur les groupes classiques, Actualités Sci. et Ind. No. 1040, Hermann 1948. MR 9–494

Les extensions quadratiques des corps non commutatifs et leurs applications, *Acta Math.* **87** (1952), 175–242. MR 14–239

Sur les generateurs des groupes classiques, *Summa Brasil. Math.* **3** (1955), 149–179. MR 18–217

La géométrie des groupes classiques, Troisième édition. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 5. Springer-Verlag 1971. MR 46#9186

A. Duncan

An automorphism of the symplectic group $Sp_4(2^n)$, *Proc. Cambridge Phil. Soc.* **64** (1968), 5–9. MR 36#1548

R. H. Dye

On the transitivity of the orthogonal and symplectic groups in projective space, *Proc. Cambridge Philos. Soc.* **68** (1970), 33–43. MR 41#4363

A geometric characterization of the special orthogonal groups and the Dickson invariant, *J. London Math. Soc.* (2) **15** (1977), 472–476. MR 56#11899

M. Dyer

Hecke Algebras and Reflections in Coxeter Groups, Ph. D. thesis. The University of Sydney 1987.

Reflection subgroups of Coxeter systems, *J. Algebra* **135** (1990), 57–73.

M. Eichler

Quadratische Formen und orthogonale Gruppen, Grundlehren der mathematischen Wissenschaften, Band 63. Springer-Verlag 1952. MR 14–450; Zweite Auflage MR 50#4484

H. Freudental

Une interprétation géométrique des automorphismes extérieurs du group symétrique S_6, *Rend. Sem. Mat. Fis. Milano* **42** (1972), 47–56. MR 49#2908

Une étude de quelques quadrangles généralisés, *Ann. Mat. Pura Appl.* (4) **102** (1975), 109–133. MR 50#11011

D. Goldschmidt

Abstract reflections and Coxeter groups, *Proc. Amer. Math. Soc.* **67** (1977), 209–214. MR 57#3281

H. Gross
Quadratic forms in infinite dimensional vector spaces, Progress in Mathematics, Vol. 1. Birkhäuser 1979. MR 81f:10027

L. C. Grove and C. T. Benson
Finite Reflection Groups, Graduate Texts in Mathematics **99**. Springer-Verlag 1985; 2nd edition. MR 85m:20001

A. J. Hahn
Unipotent elements and the spinor norms of Wall and Zassenhaus, *Arch. Math. (Basel)* **32** (1979), 114–122. MR 80d:15030

A. J. Hahn and O. T. O'Meara
The Classical Groups and K-Theory, Grundlehren der mathematischen Wissenschaften, 291. Springer 1990.

M. Hazewinkel, W. Hesselink, D. Siersma and F. D. Veldkamp
The ubiquity of Coxeter-Dynkin diagrams (an introduction to the A-D-E problem), *Nieuw Arch. Wisk.* (3) **25** (1977), 257–307. MR 58#10879

D. Hertzig
Forms of algebraic groups, *Proc. Amer. Math. Soc.* **12** (1961), 657–660. MR 23#A2486

J. W. P. Hirschfeld
Projective Geometries over Finite Fields, Oxford Mathematical Monographs. The Clarendon Press. Oxford University Press 1979. MR 81h:51007

D. G. Higman
Classical groups with an appendix by D. E. Taylor, T.H. Report 78–WSK–04 Technological University Eindhoven 1978.

A. J. Hoffman and R. R. Singleton
On Moore graphs with diameters 2 and 3, *IBM J. Research and Development* **4** (1960), 497–504. MR 25#3857

B. Huppert
Endliche Gruppen. I, Grundlehren der mathematischen Wissenschaften, 134. Springer-Verlag 1967. MR 37#302
Geometric algebra, U. of Illinois at Chicago Circle 1968/69.
Isometrien von Vektorräumen. I, II, *Arch. Math. (Basel)* **35** (1980), 164–176; Math. Z. **175** (1980), 5–20. MR 84d:15002

B. Huppert and N. Blackburn
Finite Groups. III, Grundlehren der mathematischen Wissenschaften, 243. Springer-Verlag 1982. MR 84i:20001b

K. Iwasawa
Uber die Einfachheit der speziellen projektiven Gruppen, *Proc. Imp. Acad. Tokyo* **17** (1941), 57–59. MR 2–308

C. Jordan
Traité des substitutions et des équations algébriques, Gauthier-Villars 1857.

W. M. Kantor

Rank 3 characterizations of classical geometries, *J. Algebra* **36** (1975), 309–313. MR 52#8229

Classical groups from a non-classical standpoint, (Oxford lectures 1978) Mathematical Institute, Oxford 1979. MR 81h:20056

Some geometries that are almost buildings, *Europ. J. Combinatorics* **2** (1981), 239–247. MR 83a:51020

Generalized polygons, SCABS and GABS, pp. 79–158 of *Buildings and the Geometry of Diagrams* (ed. L. A. Rosati), Lecture Notes in Mathematics, **1181**. Springer-Verlag 1986. MR 87k:51014

P. Kleidman and M. Liebeck

The Subgroup Structure of the Finite Classical Groups, London Math. Soc. Lecture Note Series **129**, Cambridge University Press 1990.

F. Klein

Gesammelte Mathematische Abhandlungen. Erster Band, Springer 1921.

R. Lipschitz

Analyse française du memoire précédent, *Bull. Sci. Math.* (2) **10** (1886), 163–183.

H. Lüneburg

Die Suzukigruppen und ihre Geometrien, Lecture Notes in Mathematics **10**, Springer-Verlag 1965. MR 34#7634

G. Mason and E. E. Shult

The Klein correspondence and the ubiquity of certain translation planes, *Geometriae Dedicata* **21** (1986), 29–50. MR 87g:51002

J. McLaughlin

Some groups generated by transvections, *Arch. Math. (Basel)* **18** (1967), 364–368. MR 96–1448

Some subgroups of $SL_n(\mathbb{F}_2)$, *Illinois J. Math.* **13** (1969), 108-115. MR 96–1449

A. Neumaier

Some sporadic geometries related to PG(3,2), *Arch. Math. (Basel)* **42** (1984), 89–96. MR 86a:51028

A. Ohara

La structure du groupe des similitudes directes $GO_6^+(Q)$ sur un corps de caractéristique 2, *Osaka Math. J.* **10** (1958), 239–257. MR 20#7067

O. T. O'Meara

Introduction to quadratic forms, Grundlehren der mathematischen Wissenschaften, 117. Springer-Verlag 1963. MR 27#2485; second printing MR 50#269

Lectures on linear groups, CBMS Regional Conference Series No. 22. Amer. Math. Soc. 1974. MR 50#2355

M. O'Nan

Automorphism groups of unitary block designs, *J. Algebra* **20** (1973), 495–511. MR 45#4995

S. E. Payne and J. A. Thas
Finite Generalized Quadrangles, Pitman 1984.

R. Ree
On some simple groups defined by Chevalley, *Trans. Amer. Math. Soc.* **84** (1957), 392–400.

M. Ronan
Lectures on Buildings, Academic Press 1989.

M. Ronan and J. Tits
Building buildings, *Math. Ann.* **278** (1987), 291–306. MR 89e:51005

P. Scherk
On the decomposition of orthogonalities into symmetries, *Proc. Amer. Math. Soc.* **1** (1950), 481–491. MR 12–157

E. E. Shult
Groups, polar spaces and related structures, pp. 130-161 of *Combinatorics. Part 3: Combinatorial Group Theory* (ed. M. Hall, Jr. and J.H. van Lint), Math. Centre Tracts, 57. Math. Centrum, Amsterdam 1974. MR 52#176
Characterizations of the Lie incidence geometries, pp. 157–186 of *Surveys in Combinatorics* (ed. E. Keith Lloyd), London Math. Soc. Lecture Note Series, **82**. Cambridge University Press 1983. MR 85g:05003

C. L. Siegel
Uber die Zetafunktionen indefiniter Quadratischer Formen. II, *Math. Z.* **44** (1938), 398–426.

R. Steinberg
Variations on a theme of Chevalley, *Pacific J. Math.* **9** (1959), 875–891. MR 22#79
Lectures on Chevalley groups, Yale University 1967.
Endomorphisms of Linear Algebraic Groups, *Mem. Amer. Math. Soc.* **80** (1968), 1–108. MR 37#6288

M. Suzuki
A new type of simple groups of finite order, *Proc. Nat. Acad. Sci. U.S.A.* **46** (1960), 868–870. MR 22#11038
Group Theory I, Grundlehren der mathematischen Wissenschaften, 247. Springer-Verlag 1982. MR 82k:20001c

J. J. Sylvester
Elementary researches in the analysis of combinatorial aggregation, *Philos. Mag.* **24** (1844), 285–296.

T. Tamagawa
On the structure of orthogonal groups, *Amer. J. Math.* **80** (1958), 191–197. MR 20#71

D. E. Taylor
Some classical theorems on division rings, *Enseignement Math.* (2) **20** (1974), 293–298. MR 50#13134
Unitary block designs, *J. Combin. Theory* **16** (1974), 51–56. MR 48#12270
Regular 2-graphs, *Proc. London Math. Soc. (3)* **35** (1977), 257–274. MR 57#16147

J. A. Thas

Combinatorics of finite quadrangles: a survey, *Ann. of Discrete Math.* **14** (1982), 57–76. MR 83g:51018

J. A. Thas and S. E. Payne

Classical generalized quadrangles: a combinatorial study, *Ars Combinatoria* **2** (1976), 57–110. MR 55#140

J. Tits

Sur la trialité et certains groupes qui s'en déduisent, *Inst. Hautes Études Sci. Publ. Math., no. 2*, (1959) 14–60. MR 21#3517

Les groupes simples de Suzuki et Ree, *Séminaire Bourbaki*, déc. 1960, exp. 210.

Ovoïdes et groupes de Suzuki, *Arch. Math. (Basel)* **13** (1962), 187–198. MR 25#3990

Théorème de Bruhat et sous-groupes paraboliques, *C.R. Acad. Sci. Paris* **254** (1962), 2910–2912. MR 25#2149

Algebraic and abstract simple groups, *Ann. of Math.* (2) **80** (1964), 313–329. MR 29#2259

Une propriété caracteristique des ovoïdes associés aux groupes de Suzuki, *Arch. Math. (Basel)* **17** (1966), 136–153. MR 98–1482

Formes quadratiques, groupes orthogonaux et algèbres de Clifford, *Invent. Math.* **5** (1968), 19–41. MR 37#6307

Le problème des mots dans les groupes de Coxeter, pp. 175–185 of *Symposia Mathematica* (INDAM, Rome 1967/68) vol. 1. Academic Press 1969. MR 40#7339

Buildings of spherical type and finite BN-pairs, Lecture Notes in Mathematics, **386**. Springer-Verlag 1974. MR 57#9866

Classification of buildings of spherical type and Moufang polygons: a survey, pp. 229–246 of *Teorie combinatorie* (ed. B. Segre), Roma Accad. Naz. Lincei 1976. MR 56#3140

Buildings and Buekenhout geometries, pp. 309–320 of *Finite Simple Groups II* (ed. M. J. Collins). Academic Press 1980. MR 82h:20021

A local approach to buildings, pp. 519–547 of *The Geometric Vein. The Coxeter Festschrift* (ed. C. Davis, B. Grünbaum and F. A. Sherk). Springer-Verlag 1981. MR 83k:51014

Immeubles de type affine, pp. 159–190 of *Buildings and the Geometry of Diagrams* (ed. L. Rosati), Lecture Notes Mathematics, **1181**. Springer-Verlag 1986. MR 87a:22032

Buildings and group amalgamations, pp. 110–127 of *Proceedings of Groups–St. Andrews 1985*, London Math. Soc. Lecture Ser. **121** Cambridge University Press 1986. MR 89i:20001

J. A. Todd

As it might have been, *Bull. London Math. Soc.* **2** (1970), 1–4. MR 52#3353

O. Veblen and J. W. Young

Projective geometry. I, Ginn 1910.

F. D. Veldkamp

Polar geometry. I, II, III, IV, V, *Nederl. Akad. Wetensch. Proc. Ser. A* **62**; **63** = *Indag. Math.* **21** (1959), 512–551; **22** (1960), 207–212. MR 23#A2773

B. L. van der Waerden
Gruppen von Linearen Transformationen, Springer 1935. (Reprinted by Chelsea Publishing Co. 1948.)

A. Wagner,
On collineation groups of projective spaces. I, *Math. Z.* **76** (1961), 411–426. MR 24#A1661

G. E. Wall
The structure of a unitary factor group, *Inst. Hautes Études Sci. Publ. Math., no. 1*, (1959). MR 21#3517
On the conjugacy classes in the unitary, symplectic and orthogonal groups, *J. Austral. Math. Soc.* **3** (1963), 1–63. MR 27#212

J. H. M. Wedderburn
A theorem on finite algebras, *Trans. Amer. Math. Soc.* **6** (1905), 349–352.

R. Lipschitz
Correspondence, *Ann. of Math.* **69** (1959), 247–251.

H. Weyl
The Classical Groups. Their Invariants and Representations, Princeton Math. Series 1, Princeton University Press 1939. MR 1–42
The theory of groups and quantum mechanics, Dover 1950.

A. N. Whitehead
The axioms of projective geometry, Cambridge Transactions in Mathematics and Mathematical Physics 4. Cambridge University Press 1906.

H. Wielandt
Finite permutation groups, Academic Press 1964. MR 32#1252

E. Witt
Theorie der quadratischen Formen in beliebigen Korpern, *J. reine angew. Math.* **176** (1937), 31–44.
Spiegelungsgruppen und Aufzählung halbeinfacher Liescher Ringe, *Abh. Math. Sem. Univ. Hamburg* **14** (1941), 289–322.

A. E. Zalesskii
Linear groups, *Uspekhi Mat. Nauk* **36** (1981), 57–107; *Russian Math. Surveys* **36** (1981), 63–128. MR 83d:20032

H. Zassenhaus
Kennzeichnung endlicher linearer Gruppen als Permutationsgruppen, *Abh. Math. Sem. Univ. Hamburg* **11** (1936), 17–44.
On the spinor norm, *Arch. Math. (Basel)* **13** (1962), 434–451. MR 26#6266

Index of Symbols

Index of Names

Index