CLASSICAL GROUPS FROM A NON-CLASSICAL VIEWPOINT

BY

W.M. KANTOR

Classical groups from a non-classical viewpoint

PREFACE

How well-known are the well-known finite simple groups? There are several approaches to this question. One leads to characterization theorems, and another to representation theory. A third approach, taken in these notes, concerns subgroups and permutation representations.

These notes record a series of lectures given at Oxford University in 1978. Part 1 is devoted to an extremely quick summary of the standard, basic facts concerning the classical groups, their geometries and their BN-pairs; it is essentially a long collection of exercises. Part 2 is the main part, and presents some recent results and methods concerning these groups. Finally, Part 3 is devoted to the geometry of the not-so-classical groups $E_8(q)$.

The background of the audience included the basic, elementary properties of groups, characters and BN-pairs.

The goal was to present ideas, rather than complete proofs of the most general known results. Consequently, only partial proofs of significant special cases were given, while more general results were merely described. References to the latter are given here. However, no attempt has been made to give a comprehensive bibliography;

therefore, apologies are owed to those people whose important contributions have not been mentioned.

I am grateful to C. Ronse for his excellent notes, upon which the present account is based; to the Mathematics Institute at Oxford University for inviting me to give these lectures; and to Joyce E. Falkenberg for her excellent typing.

Table of Contents

Part 1. The classical theory of classical groups

(A) The special linear group SL(n,q)

All fields, vectors spaces and groups will be finite.

Let $F = GF(q)$, and let V be an n-dimensional vector space over F with $n \geq 2$.

$SL(V) = SL(n,q)$ denotes the group of all linear transformations on V of determinant 1, while $GL(V) = GL(n,q)$ denotes the group of all nonsingular linear transformations on V. Both groups are transitive on the set of i-spaces for each $i \leq n-1$.

BN-pair. Fix a basis $e_1,\ldots,e_n$ of V. Set

B = stabilizer of $\{\langle e_1,\ldots,e_i\rangle \mid i = 1,\ldots,n-1\}$,

N = stabilizer of $\{\langle e_1\rangle,\ldots,\langle e_n\rangle\}$,

with both stabilizers taken in $SL(n,q)$. Then $N \cong (F^*)^{n-1} \times S_n$, where S_n is the Weyl group.

Maximal parabolic subgroups. Each of those containing B is just the stabilizer of some $\langle e_1,\ldots,e_i\rangle$.

Root groups = transvection groups. These are the conjugates of $\left\{\begin{pmatrix} 1 & 0 & \alpha \\ & 1 & \\ 0 & & 1 \end{pmatrix} \mid \alpha \in F\right\}$, and are isomorphic to F^+. Thus, each consists of all transformations

$v \to v + \alpha f(v)a$ for all $\alpha \in F$ and some $a \in V - \{0\}$, $f \in V^* - \{0\}$. Notice that each such transformation (called a transvection) induces the identity on both $\ker f$ and $V/\langle a \rangle$.

Lemma. $SL(V)$ is generated by its transvection groups.

Theorem. $SL(n,q)$/scalars is simple, except for $SL(2,2)$ and $SL(2,3)$.

See Artin [1] or Carter [7].

(1.B) Bilinear forms

We will now assume that V is also equipped with a form $(\ ,\ ) : V \times V \to F$; for the time being the form will be assumed bilinear, and subject to one of the following two conditions $\forall u, v \in V$:

symplectic	orthogonal, q odd
$(v,v) = 0$, $(u,v) = -(v,u)$	$(u,v) = (v,u)$.

If $X \subseteq V$, set $X^\perp = \{v \in V \mid (v,x) = 0\}$. We then further assume that $V^\perp = 0$ (so V is nonsingular).

Lemma. If $W \leq V$ then $\dim V = \dim W + \dim W^\perp$ and $(W^\perp)^\perp = W$.

Thus, $W \cap W^{\perp} = 0$ iff $V = W \oplus W^{\perp}$; in this case W is called nonsingular, and we write $V = W \perp W^{\perp}$.

Isometry: $g \in GL(V)$ satisfying $(u^g, v^g) = (u,v)$ $\forall u, v \in V$. The set of isometries forms a group, called and denoted as follows in our two cases.

symplectic group $Sp(V) = Sp(n,q)$ orthogonal group $O(V)$.

(1.C) Symplectic geometry

Assume that V is symplectic.

Basis. Take any $e_1 \neq 0$ in V, and any $f_1 \in V - e_1^{\perp}$. Then $(e_1, f_1/(e_1,f_1)) = 1$. Replace f_1 by $f_1/(e_1,f_1)$ and deduce that

$$(e_1,e_1) = 0 = (f_1,f_1) \quad \text{and} \quad (e_1,f_1) = 1 = -(f_1,e_1).$$

Set $W_1 = \langle e_1, f_1 \rangle$, and compute that $W_1 \cap W_1^{\perp} = 0$. Thus, $V = W_1 \perp W_1^{\perp}$, and $(\ ,\ )$ induces a nonsingular symplectic "geometry" on $W_1^{\perp}$. This yields the following.

Theorem. There is a basis (called a symplectic basis) $e_1, \ldots, e_m, f_1, \ldots, f_m$ (where $m = n/2$) such that $\forall i,j$:

$$(e_i,e_j) = 0 = (f_i,f_j) \quad \text{and} \quad (e_i,f_j) = \delta_{ij} = -(f_j,e_i).$$

If $\alpha_i, \beta_i, \gamma_i, \delta_i \in F$ then

$$(\Sigma \alpha_i e_i + \Sigma \beta_i f_i,\ \Sigma \gamma_i e_i + \Sigma \delta_i f_i) = \Sigma(\alpha_i \delta_i - \beta_i \gamma_i).$$

Corollary. The dimension n of V is even, and there is a unique symplectic geometry for each q and each even n. (Here, "unique" refers to uniqueness up to an obvious notion of equivalence of spaces and forms.)

Example. $n = 2m = 2$. $(\alpha e_1 + \beta f_1, \gamma e_1 + \delta f_1) = \det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$.

Here, $g \in GL(V)$ is an isometry iff $\det g = 1$, so $Sp(2,q) = SL(2,q)$.

The proof of the theorem also yields several transitivity properties:

Corollary. $Sp(V)$ is transitive on (the set of) symplectic bases.

Corollary. $Sp(V)$ is transitive on the ordered pairs of non-perpendicular 1-spaces, as well as on the ordered pairs of district perpendicular 1-spaces.

A subspace W of V is called totally isotropic (t.i.) if $(W,W) = 0$.

Corollary. $Sp(V)$ is transitive on the t.i. subspaces of dimension j for each $j = 1,\ldots,m$. The maximal t.i. subspaces have dimension m ($\langle e_1,\ldots,e_m \rangle$ is an example of one).

Such transitivity properties are all contained in the following basic result.

Witt's Theorem. Let $W, W' \leq V$, and assume that $h : W \to W'$ is an invertible linear transformation such that $(v^h, w^h) = (v,w)$ $\forall v, w \in W$. Then there exists some $g \in Sp(V)$ such that $g|_W = h$. (Artin [1].)

Lemma. If $W \leq V$ then $W/W \cap W^\perp$ naturally inherits a nonsingular symplectic geometry. (Compute, using $w + W \cap W^\perp, v + W \cap W^\perp) = (w,v)$ for $w, v \in W$.)

BN-pair. B = stabilizer of $\{\langle e_1,\ldots,e_i\rangle \mid i=1,\ldots,m\}$,

N = stabilizer of $\{\langle e_1\rangle,\ldots,\langle e_m\rangle,\langle f_1\rangle,\ldots,\langle f_m\rangle\}$.

This time $N \cong (K^*)^m \times (2^m \rtimes S_m)$, where 2^m denotes an elementary abelian group of that order and $2^m \rtimes S_m$ is the Weyl group, of type C_m.

The maximal parabolic subgroups containing B are precisely the stabilizers of the t.i. subspaces $\langle e_1,\ldots,e_i\rangle$.

Each long root group is a transvection group $\{v \to v + \alpha(v,a)a \mid \alpha \in F\}$, one for each 1-space $\langle a\rangle$. (Here and elsewhere, short root groups will usually be omitted from our discussion.)

Theorem. $Sp(V)$ is generated by its transvection groups.

This follows from the BN structure, but can also be easily obtained directly by starting with $Sp(2,q) = SL(2,q)$ and applying induction.

Theorem. $Sp(2m,q)/\langle -1\rangle$ is simple for $2m > 2$, with the single exception $Sp(4,2) \cong S_6$. (Artin [1].)

Example. Let W be a 6-dimensional vector space over $F = GF(2)$, and let $v_1,..,v_6$ be a basis. Note that S_6 acts as a subgroup of $GL(W)$ permuting $\{v_1,\ldots,v_6\}$. Define $(\ ,\ ): W \times W \to F$ by $(v_i,v_j) = 1 + \delta_{ij}$. This yields a nonsingular symplectic geometry (compute!). Clearly, $W_1 = \langle \Sigma v_i \rangle$ is S_6-invariant, so S_6 acts on the space $V = W_1^{\perp}/W_1$, which is itself nonsingular by a previous lemma. The transposition $(1,2) \in S_6$ induces a transvection $v \to v + (v,v_1 + v_2)(v_1 + v_2)$ on W, and hence also on V. This yields all $\binom{6}{2} = 2^4 - 1$ transvections in $Sp(V)$. Thus, $S_6 \cong Sp(4,2)$.

An alternative approach to part of this identification of $Sp(4,2)$ is provided by the following

Theorem. $|Sp(2m,q)| = q^{m^2} \prod_1^m (q^{2i}-1)$.

Proof. There are $q^{2m} - 1$ choices for e_1, and then $(q^{2m} - q^{2m-1})/(q-1) = q^{2m-1}$ choices for f_1 with $(e_1,f_1) = 1$. Now induction shows that the number of symplectic bases equals the stated product.

(1.D) Unitary geometry

Since unitary geometry resembles symplectic geometry, we will discuss it before returning to the orthogonal case. Here, F will denote $GF(q^2)$, and hence admits the the involutory antomorphism $\bar{\alpha} = \alpha^q$, $\alpha \in F$. This time, the (hermitian) form $(\ , \ ) : V \times V \to F$ will satisfy the following ($\forall u, v, w \in V$, $\forall \alpha \in F$):

$(u+v,w) = (u,w) + (v,w)$

$(\alpha u, v) = \alpha(u,v)$ and $(u, \alpha v) = \bar{\alpha}(u,v)$

$(u,v) = \overline{(v,u)}$

$V^\perp = 0$ (nonsingularity).

Note that $(v,v) = \overline{(v,v)}$, so $(v,v) \in GF(q)$.

Once again, if $W \leq V$ then $\dim V = \dim W + \dim W^\perp$ and $(W^\perp)^\perp = W$, and W is called nonsingular if $W \cap W^\perp = 0$. But this time, many 1-spaces are not t.i.:

Theorem. There is an orthonormal basis.

Proof. 1) $\exists v : (v,v) \neq 0$ if $n > 1$. For, let $(v,v) = 0 = (w,w)$ with $w \notin v^\perp$. Then we may assume that $(v,w) = 1$, so that $(v + \alpha w, v + \alpha w) = (v,v) + \alpha(w,v) + \bar{\alpha}(v,w) + (\alpha w, \alpha w) = \alpha + \bar{\alpha} \neq 0$ for some $\alpha \in F$.

2) $V = \langle v \rangle \perp v^\perp$ and $(\ , \ )$ induces a unitary geometry on $v^\perp$. Also, $(v,v) \in GF(q)$, so $(v,v) = \alpha^{-(q+1)} = 1/\alpha\bar{\alpha}$

for some $\alpha \in F$, and then $(\alpha v, \alpha v) = 1$. Induction now completes the proof.

Corollary. Each vector space over $GF(q^2)$ has a unique unitary geometry. If $v_1,\ldots,v_n$ is an orthonormal basis then $(\Sigma\alpha_i v_i, \Sigma\beta_i v_i) = \Sigma\alpha_i\bar{\beta}_i$.

Theorem. There is another basis of one of the following types:

$n = 2m : e_1,\ldots,e_m, f_1,\ldots,f_m,$

$n = 2m+1 : d,e_1,\ldots,e_m, f_1,\ldots,f_m,$

where $(e_i,e_j) = 0 = (f_i,f_j)$, $(e_i,f_j) = \delta_{ij} = (f_j,e_i)$, $(d,d) = 1$, $(d,e_i) = 0 = (d,f_i)$.

Proof. We may assume that $n = 2m$. There is a vector $e_1 \neq 0$ with $(e_1,e_1) = 0$ (such as $\alpha v_1 + v_2$, where $\alpha^{q+1} = -1$), and some $f_1 \notin e_1^{\perp}$ with $(e_1,f_1) = 1$. Now proceed as in the symplectic case.

This time the group of isometries (the unitary group) is denoted by $GU(n,q)$, while $SU(n,q) = GU(n,q) \cap SL(n,q^2)$.

Example. If $\alpha_i\bar{\alpha}_i = 1$ for $i = 1,\ldots,n$, then with respect to an orthonormal basis $\{v_i\}$ the diagonal matrix $\mathrm{diag}(\alpha_1,\ldots,\alpha_n)$ is in $GU(n,q)$; its determinant is $\Pi\alpha_i$, and hence is an element of F of order $q+1$. More generally, $GU(n,q)$ can be identified with the group

of all $n \times n$ (unitary) matrices (α_{ij}) satisfying $(\alpha_{ij})(\bar{\alpha}_{ij}) = 1$. Thus, $(\det g)^{q+1} = 1$ for all $g \in GU(n,q)$. If S consists of all scalar transformations $v \to \alpha v$ with $\alpha^{q+1} = 1$, then $GU(n,q) = SU(n,q) \cdot S$.

Lemma. $SU(n,q)$ is transitive on nonsingular (resp. t.i.) j-spaces for each $j \leq n/2$. A maximal t.i. subspace has dimension $[n/2]$.

BN structure. Let $n = 2m$ or $2m+1$ as above and set

B = stabilizer of $\{\langle e_1,\ldots,e_i\rangle \mid i = 1,\ldots,m\}$,

N = stabilizer of $\{\langle e_1\rangle,\ldots,\langle e_m\rangle, \langle f_1\rangle,\ldots,\langle f_m\rangle\}$,

with both stabilizers taken in $SU(n,q)$ or in $GU(n,q)$.

Maximal parabolic subgroups behave as in the symplectic case, the Weyl group is as before, but the root system is of type B_m or "BC_m" (the union of a system of type B_m with one of type C_m).

Long root groups again consist of transvections, and can be described as follows. Fix $\sigma \in F$ with $\bar{\sigma} = -\sigma \neq 0$. Set $(u,v)' = \sigma(u,v)$, so $(v,u)' = \sigma(v,u) = -\bar{\sigma}(\overline{u,v}) = -(\overline{u,v})'$. Then $\{v \to v + \alpha(v,a)'a \mid \alpha \in GF(q)\}$ is a long root group for each t.i. 1-space $\langle a\rangle$. These transvections generate $SU(n,q)$.

Theorem. $SU(2m,q) \geq Sp(2m,q)$.

Proof. Write $f_i' = -f_i/\sigma$. Then

$(\Sigma\alpha_i e_i + \Sigma\beta_i f_i', \Sigma\gamma_i e_i + \Sigma\delta_i f_i')' = \Sigma(\alpha_i\bar{\delta}_i - \beta_i\bar{\gamma}_i)$,

which implies the result.

Note that the long root groups of $Sp(2m,q)$ are also long root groups for $SU(2m,q)$. Also, $SU(2,q) = Sp(2,q) = SL(2,q)$ (compute!).

This time, $SU(n,q)/\text{scalars}$ is simple for $n > 2$, with the exception of $SU(3,2)$ (Huppert [14] or Dieudonné [13]). Also, $|SU(n,q)|$ is computed as before.

(1.E) Orthogonal geometry

The orthogonal case is somewhat harder. Since we want to include characteristic 2, we will need both a bilinear form and a quadratic form.

Let V be equipped with a nonsingular symmetric bilinear form $(\ ,\ )$. Thus, V is an orthogonal geometry in odd characteristic. A quadratic form associated with $(\ ,\ )$ is a function $Q: V \to F$ satisfying ($\forall u, v \in V$, $\forall \alpha \in F$):

$Q(\alpha v) = \alpha^2 Q(v)$ and $Q(u+v) = Q(u) + Q(v) + (u,v)$.

Remarks. $4Q(v) = Q(2v) = Q(v) + Q(v) + (v,v)$, so $(v,v) = 2Q(v)$. Thus, for q odd, $(\ ,\ )$ and Q determine one another. For q even, this is false; however,

$(v,v) = 0$, so $(\ ,\ )$ defines an underlying symplectic geometry. In either case, V, equipped with Q, will be called an orthogonal geometry.

Note also that $\dim V$ is even if q is even. A modification of this definition is needed when $\dim V$ is odd but q is even; this will be discussed later in (1.F).

An isometry is a linear transformation $g \in GL(V)$ such that $Q(v^g) = Q(v)$ (and hence $(u^g, v^g) = (u,v)$) for all $u,v \in V$. The group of isometries is denoted by $O(V)$, while $SO(V) = O(V) \cap SL(V)$. There is also a normal subgroup $\Omega(V)$ of index 2 in $SO(V)$ which has yet to be constructed (and only will be in characteristic 2); here, $\Omega(V) = O(V)'$ with only one exception.

A singular vector is a vector $v \neq 0$ satisfying $Q(v) = 0$. Any other nonzero vector is called nonsingular. A totally singular (t.s.) subspace is a subspace W satisfying $Q(W) = 0$, and hence also $(W,W) = 0$. Thus, if q is even then t.s. implies t.i., but the converse is false.

As before, a subspace W is called nonsingular if $W \cap W^{\perp} = 0$. (Note that Q is irrelevant here.) Warning: If q is even and v is a nonsingular vector, then $\langle v \rangle$ is not a nonsingular subspace.

Lemma. If $W \leq V$ then $W^{\perp}/W \cap W^{\perp}$ acquires an orthogonal geometry via the definition $\overline{Q}(v+W) = Q(v)$.

Following the pattern of (1.C) and (1.D), we have to discuss the following topics for V and $O(V)$: bases, uniqueness of geometry types, transitivity properties, generation, simplicity, and BN structure.

Bases. If q is odd, then V has an orthogonal basis (whose existence is proved as in (1.D)). However, we will aim at a different type of basis, which exists in each characteristic and is related to the BN structure.

The orthogonal geometry V is called anisotropic if $Q(v) \neq 0$ for all $v \neq 0$.

Lemma. If V is anisotropic, then $\dim V \leq 2$. If, moreover, $\dim V = 2$, then V is of a unique type, and has a basis d,d' satisfying $Q(d') = 1 = (d,d')$.

Proof. Assume that $n \geq 2$, take any $e \neq 0$ and any $d \notin e^{\perp}$, and consider $W = \langle d,e \rangle$. Set $Q(e) = \varepsilon$, and adjust d so that $(d,e) = \varepsilon$. Set $Q(d) = \sigma\varepsilon$ for some σ. Then $Q(\alpha e + d) = \alpha^2\varepsilon + \sigma\varepsilon + \alpha\varepsilon = \varepsilon(\alpha^2 + \alpha + \sigma) \neq 0$ for all $\alpha \in F$, so $x^2 - x + \sigma$ is an irreducible polynomial. Let θ be a root of this polynomial in $GF(q^2)$, and let bar denote the involutory automorphism of $F(\theta) = GF(q^2)$. Then

$Q(\alpha e + \beta d) = \varepsilon(\alpha + \beta\theta)(\alpha + \beta\bar{\theta})$. This proves first that $Q(W) = F$, so we may take $\varepsilon = 1$; and then that W has a unique type.

Finally, if $n \geq 3$ and $v \in \langle d,e\rangle^{\perp} - \{0\}$, then $Q(w) = -Q(v)$ for some $w \in W$. But then $Q(v+w) = 0$, contradicting the fact that $\dot{V}$ is anisotropic.

Theorem. There is a basis of one of the following types:

(i) $n = 2m$: $e_1,\ldots,e_m$, $f_1,\ldots,f_m$ with $Q(e_i) = Q(f_i) = 0$ and $(e_i,f_j) = \delta_{ij}$;

(ii) $n = 2m+2$: d,d', $e_1,\ldots,e_m$, $f_1,\ldots f_m$, with e_i and f_i as above, $(d,e_i) = (d,f_i) = (d',e_i) = (d',f_i) = 0$, $Q(d') = 1 = (d,d')$, $Q(d) \neq 0$, and $\langle d,d'\rangle$ anisotropic; or

(iii) $n = 2m+1$: $d,e_1,\ldots,e_m$, $f_1,\ldots,f_m$, behaving as in (ii).

Proof. By the preceeding lemma, we may assume that there is a singular vector e_1. As usual, there exists f with $(e_1,f) = 1$. Then $Q(\alpha e_1 + f) = Q(f) + \alpha$, so $\langle e_1,f\rangle$ has a unique t.s. 1-space other that $\langle e_1\rangle$. We may then assume that $Q(f) = 0$. Now set $f_1 = f$, observe that $V = \langle e_1 f_1\rangle \perp \langle e_1,f_1\rangle^{\perp}$, and apply induction.

Corollary. Q has one of the following shapes:

(i) $Q(\Sigma\alpha_i e_i + \Sigma\beta_i f_i) = \Sigma\alpha_i\beta_i$;

(ii) $Q(\gamma d + \gamma' d' + \Sigma\alpha_i e_i + \Sigma\beta_i f_i) = \Sigma\alpha_i\beta_i + \gamma^2\sigma + \gamma\gamma' + \gamma^2$; or

(iii) $Q(\gamma d + \Sigma\alpha_i e_i + \Sigma\beta_i f_i) = \Sigma\alpha_i\beta_i + \gamma^2\sigma$.

Corollary. For each even n there are at most two types of orthogonal geometries. For each odd n and odd q there is essentially just one such type.

Proof. Defining Q by (i), (ii) or (iii) produces a quadratic form, and $V^\perp = 0$ by a simple computation. Uniqueness for (ii) is clear from the lemma. If $n = 2m+1$, replacing $Q(v)$ by $Q(v)/\sigma$ for all $v \in V$ does not change anything except in a trivial manner, and hence the geometry is essentially unique.

Definitions. In (i), (ii) and (iii), V is said to have type $O^+(2m,q)$, $O^-(2m+2,q)$ and $O(2m+1,q)$, respectively. The corresponding groups are denoted as follows (where the last column remains to be defined):

$O(V)$	$SO(V)$	$\Omega(V)$
$O^+(2m,q)$	$SO^+(2m,q)$	$\Omega^+(2m,q)$
$O^-(2m+2,q)$	$SO^-(2m+2,q)$	$\Omega^-(2m+2,q)$
$O(2m+1,q)$	$SO(2m+1,q)$	$\Omega(2m+1,q)$.

Remark. The geometries of types $0^+(2m,q)$ and $0^-(2m,q)$ are, in fact, distinct. This follows, for example, from the fact that they have different numbers of t.s. 1-spaces (see below).

Corollary. $O(V)$ is transitive on t.s. i-spaces for each $i \leq \frac{1}{2} \dim V$ (or $i \leq \frac{1}{2} \dim V - 1$ in the case of $0^-(2m,q)$). Also, $O(V)$ is transitive on the set of ordered pairs of distinct perpendicular (resp. non-perpendicular) t.s. 1-spaces.

The number of t.s. subspaces of each dimension is easily counted (as is also true in symplectic and unitary geometries). The method is elementary; two examples will be given for future use.

Example. The number of t.s. 1-spaces is $(q^m-1)(q^{m-1}+1)/(q-1)$ for a geometry of type $0^+(2m,q)$. (This number is $(q^m+1)(q^{m-1}-1)/(q-1)$ for an $0^-(2m,q)$-space.)

Proof. Let φ_m denote the number of vectors v such that $Q(v) = 0$; this is also the number of solutions to $\Sigma \alpha_i \beta_i = 0$. Clearly, $\varphi_0 = 1$. Let $m \geq 1$. If $\alpha_1 \neq 0$ then $\beta_1 = -\alpha_1^{-1} \sum_2^m \alpha_i \beta_i$, and this provides $(q-1)q^{2m-2}$ solutions. If $\alpha_1 = 0$ then β_1 is arbitrary, and we obtain $q\varphi_{m-1}$

solutions. Thus, $\varphi_m = q\varphi_{m-1} + (q-1)q^{2m-2}$. Since $(\varphi_m - 1)/(q-1)$ is the number of t.s. 1-spaces, our assertion is proved. (The case $0^-(2m,q)$ is very similar.)

Example. For an $0^+(2m,q)$ geometry, there are exactly $2\prod_{i=1}^{m-1}(q^i+1)$ t.s. m-spaces.

Proof. We will count the pairs $(\langle v\rangle, M)$ with $0 \neq v \in M$ and M a t.s. m-space. If θ_m denotes the number of such subspaces M, then there are $\theta_m(q^m-1)/(q-1)$ such pairs. But each $\langle v\rangle$ determines an $0^+(2m-2,q)$ space $v^\perp/\langle v\rangle$, whose t.s. m-1-spaces correspond to those t.s. m-spaces containing $\langle v\rangle$. By the preceding example, the number of pairs is also $\theta_{m-1}(q^m-1)(q^{m-1}+1)/(q-1)$. Since $\theta_1 = \varphi_1 = 2$, the result follows.

Generation. If $Q(a) \neq 0$, define r_a by

$$r_a : v \to v - \frac{(v,a)}{Q(a)}\, a.$$

Then r_a induces the identity on $a^\perp$ and sends a to -a. If q is even, r_a is a transvection. If q is odd, then r_a is a reflection.

Theorem. $O(V) = \langle r_a \mid Q(a) \neq 0\rangle$ except for $0^+(4,2) = S_3 \wr S_2$.

This is proved by induction; see Artin [1] or Dieudonné [13].

$\Omega(V)$.

Theorem. $O(V)$ has a normal subgroup $\Omega(V)$ having index 2 in $SO(V)$.

Again see Artin [1] (for q odd) or Dieudonné [13]. They use the preceding generation result, together with Clifford algebras.

Proof for q even.

Case $O^+(2m,q)$. Let $\mathcal{S}$ denote the set of all t.s. m-spaces, so $|\mathcal{S}| = 2\prod_1^{m-1}(q^i+1)$. Let $Q(a) \neq 0$, and consider the action of $r = r_a$ on $\mathcal{S}$. If $M^r = M$ for $M \in \mathcal{S}$, then the definition of r_a implies that either $a \in M$ or $M \leq a^\perp$. Since $Q(a) \neq 0 = Q(M)$, the first possibility cannot occur; if $M \leq a^\perp$ then $a \in M^\perp = M$ since $M \leq M^\perp$ and $\dim M^\perp = \dim V - \dim M = m$. Thus, r has $|\mathcal{S}|/2$ transpositions on $\mathcal{S}$, and hence induces an odd permutation there.

Now simply define $\Omega^+(2m,q)$ to be the set of elements of $O^+(2m,q)$ inducing even permutations on $\mathcal{S}$.

Case $O^-(2m,q)$. Set $\tilde{V} = V \otimes_{GF(q)} GF(q^2)$, and define $\tilde{Q}(v \otimes 1) = Q(v)$. Then $\tilde{Q}$ extends to a quadratic form on $\tilde{V}$, and turns $\tilde{V}$ into an $O^+(2m,q^2)$ space. (Think in terms of the basis and the irreducible polynomial used to define $O^-(2m,q)$, but let $\gamma, \gamma', \alpha_i, \beta_i \in GF(q^2)$.)

Then $O(V) \cong O(V) \otimes 1 \leq O(\tilde{V})$. Define $\Omega(V)$ by $\Omega(V) \otimes 1 = (O(V) \otimes 1) \cap \Omega(\tilde{V})$. If $Q(a) \neq 0$ then $r_{a\otimes 1} = r_a \otimes 1 \notin \Omega(\tilde{V})$, and hence $r_a \notin \Omega(V)$. Thus, $|O(V) : \Omega(V)| = 2$.

Simplicity. See the table in (I.G). There, $P\Omega(V) = \Omega(V)/\text{scalars}$, $PSp(V) = Sp(V)/\text{scalars}$, and $PSU(V) = SU(V)/\text{scalars}$.

BN-pair. Let $e_1,\ldots,e_m$, $f_1,\ldots,f_m$ be as before; d and d' are irrelevant for now. As in (1.C) and (1.D), define the following subgroups of $O(V)$:

B = stabilizer of $\{\langle e_1,\ldots,e_i\rangle \mid i = 1,\ldots,m\}$,

N = stabilizer of $\{\langle e_1\rangle,\ldots,\langle e_m\rangle, \langle f_1\rangle,\ldots,\langle f_m\rangle\}$.

This provides a BN-pair for $O(V)$, with Weyl group of type B_m. The maximal parabolic subgroups are simply the stabilizers of t.s. subspaces.

Similarly, $\Omega^-(2m,q)$ and $\Omega(2m+1,q)$ have BN-pairs of type B_m.

Case $\Omega^+(2m,q)$. Here $\langle e_1,\ldots,e_{m-1}\rangle$ is in exactly two t.s. m-spaces: $\langle e_1,\ldots,e_m\rangle$ and $\langle e_1,\ldots,e_{m-1}, f_m\rangle$; and these are in different $\Omega^+(2m,q)$-orbits. Thus, the stabilizer of $\langle e_1,\ldots,e_{m-1}\rangle$ is no longer a maximal parabolic subgroup. When B and N are defined as above, the Weyl group has type D_m, of index 2 in B_m.

"Long" root groups. Assume that t.s. 2-spaces $L = \langle a,b \rangle$ exist. Then long root groups $T(L)$ are defined as $\{g \in O(V) \mid g$ induces the identity on $L^{\perp}\}$. A computation shows that $T(L)$ consists of the transformations

$$v \to v - \alpha(v,a)b + \alpha(v,b)a \quad \text{with} \quad \alpha \in F.$$

A long root element is defined to be an element of some such $T(L)$.

Theorem. If t.s. lines exist then $\Omega(V) = \langle T(L) \mid$ L is a t.s. line$\rangle$.

"Short" root groups are defined only when the Weyl group is B_m, C_m or BC_m. However, as groups of linear transformations these also exist for D_m. Namely, let $\dim \langle a,b \rangle = 2$ with $Q(a) = 0$, $Q(b) = 1$ and $(a,b) = 0$. Then the desired subgroup of $O(V)$ consists of all

$$v \to v + \alpha(v,b-a)(a+b) - \alpha(v,b)b \quad \text{with} \quad \alpha \in K.$$

Moreover, $\Omega(V)$ is always generated by all long and short root groups.

(1.F) Further remarks concerning orthogonal groups

(i) Each r_a is in $O(V) - \Omega(V)$. The group $\Omega^+(2m,q)$ has 2 orbits on the set of t.s. m-spaces of V. Using

the transformations r_a, it is not hard to show that two t.s. m-spaces E and F are in the same orbit iff dim $E/E \cap F$ is even.

(ii) We have seen that $Sp(2m,q)$ is a subgroup of $SU(2m,q)$, generated by suitable long root groups of the latter.

Theorem. $SU(2m,q) \leq \Omega^+(4m,q)$ and $SU(2m+1,q) \leq \Omega^-(4m+2,q)$, with unitary transvection groups being long root groups of the orthogonal groups.

Proof. Start with a unitary space V over $GF(q^2)$, regard V as a $GF(q)$-space, and define $Q(v) = (v,v)$. This yields a quadratic form (over $GF(q)$), and turns V into an orthogonal space. Thus, $SU(n,q) \leq \Omega^{\pm}(2n,q)$. A 1-dimensional t.i. unitary space over $GF(q^2)$ becomes becomes a 2-dimensional t.s. subspace; this implies the last part of the result.

(iii) What about $\Omega(2m+1,q)$ when $q = 2^i$?

Let V be a $2m+2$-dimensional orthogonal $GF(q)$-space. Take any nonsingular vector w, and set $W = \langle w \rangle$.

Definition. $O(2m+1,q) = \Omega(2m+1,q)$ is the stabilizer of W in $\Omega(V)$.

Theorem. $\Omega(2m+1,q) \cong Sp(2m,q)$.

Proof. Since $(W,W) = 0$, the space $\overline{V} = W^{\perp}/W$ inherits a symplectic geometry from that of V. Each $g \in \Omega(2m+1,q)$ induces a symplectic transformation $\overline{g}$ on $\overline{V}$. We will show that $g \to \overline{g}$ is the desired isomorphism.

1-1. If $\overline{g} = 1$ then g induces the identity on $W^{\perp}/W$ and preserves Q. It follows that $g \in \langle r_w \rangle$, and hence that $g = 1$ since $r_w \notin \Omega(V)$.

Onto. Each $T(\langle w,a \rangle)$ with singular $a \in w^{\perp}$ (defined at the end of (1.E)) induces a transvection on $W^{\perp}$, and hence also on $\overline{V}$. All transvection groups of $Sp(\overline{V})$ arise in this manner, and generate $Sp(\overline{V})$.

Note: $\Omega(2m+1,q)$ naturally has a BN-pair of type B_m. Its long and short root groups correspond, respectively, to the short and long root groups of $Sp(2m,q)$. Moreover, the underlying space $W^{\perp}$ (equipped with Q) is readily seen to be essentially independent of the choice of the original space V.

(1.G) Summary

Some properties of symplectic, unitary and orthogonal groups are summarized in the accompanying table and the first 6 items on the following list. Many of these

properties have been discussed, and many will reappear later in these notes. The reader may wish to regard these as a formidable exercise, with Artin [1], Dieudonné [13] and Carter [7] available if rescue is needed from the hardest parts of them.

Notation. G is the relevant group. A point is a t.i. or t.s. 1-space, and a line is a t.i. or t.s. 2-space. (In the orthogonal case, only t.s. subspaces are considered now.)

(1) If points exist but not lines, then G is 2-transitive on the set of points (with the exception of $\Omega^+(2,q)$).

If lines exist and geometries of type $\Omega^+(4,q)$ are excluded, then G is transitive on the set of lines and has rank 3 on the set of points. If x is a point then the 3 point-orbits of G_x are $\{x\}$, the set of points in $x^\perp$ other than x, and the set of points not in $x^\perp$; their lengths are given in the last column of the table.

Two points are perpendicular iff they are collinear.

(2) $G/Z(G)$ is usually simple (see the table).

(3) G has a BN-pair, with corresponding root system in the table. The groups B and N are obtained as follows. There are linearly independent vectors

$e_1,\ldots,f_1,\ldots$, with $\langle e_i\rangle$ and $\langle f_i\rangle$ points and $(e_i,f_j)=\delta_{ij}$, such that $\langle e_1,\ldots,f_1,\ldots\rangle^\perp$ contains no points. Then B is the stabilizer of $\{\langle e_1\rangle,\langle e_1,e_2\rangle,\ldots\}$ and N is the stabilizer of $\{\langle e_1\rangle,\ldots,\langle f_1\rangle,\ldots\}$. The stabilizer of a t.i. or t.s. subspace is a maximal parabolic subgroup of G, except in the case of t.s. m-1-spaces when $G=\Omega^+(2m,q)$.

(4) G is transitive on the t.i. or t.s. subspaces of each dimension, except that $\Omega^+(2m,q)$ has 2 orbits on the set of t.s. m-spaces, which are interchanged by $O^+(2m,q)$.

(5) Symplectic and unitary cases. Each point $x=\langle a\rangle$ yields a long root group T(x), which consists of the q transvections

$$v\to v+\alpha(v,a)a,\ \alpha\in GF(q) \qquad \text{(symplectic case)}$$
$$v\to v+\alpha(v,a)a,\ \bar{\alpha}=-\alpha\in GF(q^2) \qquad \text{(unitary case)}.$$

Clearly, $T(x)\cong GF(q)^+$. Also, G permutes the set of T(x)'s as it does the set of x's, as a 2-transitive or rank 3 group (cf. (1)).

If x and y are distinct perpendicular points, then $\langle T(x),T(y)\rangle = T(x)\times T(y)$; if x and y are non-perpendicular points, then $\langle T(x),T(y)\rangle\cong SL(2,q)$.

(6) Orthogonal cases, when lines exist: each line $L = \langle a,b \rangle$ yields a long root group

$$T(L) = \{v \to v - \alpha(v,a)b + \alpha(v,b)a \mid \alpha \in GF(q)\} \cong GF(q)^+,$$

which centralizes $L^\perp$.

Let L and L' be distinct lines. If $L' < L^\perp$, then $\langle T(L), T(L') \rangle = T(L) \times T(L')$; if $\dim L' \cap L^\perp = 1$, then $\langle T(L), T(L') \rangle$ has order q^3 and is isomorphic to a Sylow subgroup of $SL(3,q)$; and if $L' \cap L^\perp = 0$ then $\langle T(L), T(L') \rangle \cong SL(2,q)$. (This is proved by a straightforward calculation.)

(7) "B-S property". If a point x is not on a line L, then x is collinear with either exactly one point of L or with all points of L. (For, $\dim x^\perp \cap L = 1$ or 2.)

Group	# Singular points	# t.s. m-spaces	Exceptions to simplicity (mod scalars)	Root system	Order of "long" root groups	Order of "short" root groups	Rank 3 subdegrees
$Sp(2m,q)$	$\frac{q^{2m}-1}{q-1}$	$\prod_1^m (q^i+1)$	$Sp(2,2) \cong S_3$ $Sp(4,2) \cong S_6$ $Sp(2,3) = SL(2,3)$	C_m	q §	q $(m>1)$	$1, q\frac{q^{2m-2}-1}{q-1}, q^{2m-1}$
$\Omega(2m+1,q)$			$\Omega(3,3) \cong A_4$	B_m	q $(m>1)$	q	
$SU(2m,q)$ $(m>1)$	$\frac{(q^{2m}-1)(q^{2m-1}+1)}{q^2-1}$	$\prod_1^m (q^{2i-1}+1)$	None	C_m	q §	q^2	$1, q^2\frac{(q^{2m-2}-1)(q^{2m-3}+1)}{q^2-1}, q^{4m-3}$
$SU(2m+1,q)$	$\frac{(q^{2m+1}+1)(q^{2m}-1)}{q^2-1}$	$\prod_1^m (q^{2i+1}+1)$	$SU(3,2) \cong 3^{1+2} \times Q_8$ (3^{1+2} denotes the extraspecial group of order 27 and exponent 3.)	BC_m **	q §	q^2 $(m>1)$	$1, q^2\frac{(q^{2m-2}-1)(q^{2m-1}+1)}{q^2-1}, q^{4m-1}$ $(m>1)$
$\Omega^+(2m,q)$	$\frac{(q^m-1)(q^{m-1}+1)}{q-1}$	$2\prod_1^m (q^i+1)$ (2 orbits of equal size)	$\Omega^+(2,q) \cong Z_{(q-1)/d}$ where $d=(2,q-1)$ $\Omega^+(4,q)$*	D_m	q $(m>1)$	q^{***} $(m>1)$	$1, q\frac{(q^{m-1}-1)(q^{m-2}+1)}{q-1}, q^{2m-2}$
$\Omega^-(2m,q)$	$\frac{(q^m+1)(q^{m-1}-1)}{q-1}$	$\prod_2^m (q^i+1)$ t.s. m-1-spaces	$\Omega^-(2,q) \cong Z_{(q+1)/d}$	B_{m-1} $(m>1)$	q $(m>2)$	q $(m>1)$	$1, q\frac{(q^{m-1}+1)(q^{m-2}-1)}{q-1}, q^{2m-2}$ $(m>2)$

* Other "natural" isomorphisms between classical groups:

$\Omega(3,q) \cong PSL(2,q)$

$P\Omega^+(4,q) \cong PSL(2,q) \times PSL(2,q)$

$P\Omega^-(4,q) \cong PSL(2,q^2)$

$\Omega(5,q) \cong PSp(4,q)$

$P\Omega^-(6,q) \cong PSU(4,q)$

$P\Omega^+(6,q) \cong PSL(4,q)$

$\Omega(2m+1,2^i) \cong Sp(2m,2^i)$

** Union of B_m and C_m in, e.g., Bourbaki's appendix [5]

*** These are not root groups from the BN viewpoint; they merely exist as linear transformations

§ Transvections

Part 2 Recent results

(2.A) The Buekenhout-Shult Theorem.

Consider a geometry consisting of a finite set of points, together with a family of distinguished subsets called lines. Assume that the following axioms hold:

(1) The set of lines is nonempty; each line has at least 3 points;

(2) No point is collinear with all remaining points;

(3) If x if a point not on the line L, then x is collinear with either one or with all points of L.

Buekenhout-Shult Theorem. If (1) - (3) hold, then either

(I) The geometry is isomorphic to the geometry consisting of all t.i. or t.s. points and lines of a symplectic, unitary or orthogonal space; or

(II) If x and L are as in (3) then x is collinear with exactly one point of L (in which case the geometry is precisely the same as what is called a generalized quadrangle).

For the proof, see Buekenhout-Shult [6], Shult [29] and Tits [30].

Note that axiom (3) is the basic one. That it holds in (I) was already noted in (1.G(7)).

Applications.

1. Characterizing generalized hexagons (Yanushka [33], Ronan [26]).

2. Characterizing classical groups via rank 3 subdegrees (cf. (2.B)).

3. Aschbacher [2] used this theorem in his characterization of (more or less) simple groups G having an involution t such that $C_G(t)$ contains $SL(2,q)$ as a subnormal subgroup for some odd q.

4. Characterizing E_6 and E_7 geometries (Cooperstein [8]).

5. Further characterizations of classical geometries.

(2.B) Rank 3 characterizations.

Let G be a transitive group of permutations of the finite set X. If $x \in X$ and G_x has exactly r orbits on X, then G is said to have rank r on X; the lengths of these orbits are the subdegrees of G.

Theorem. Assume that the subdegrees of a rank 3 group are as in the last column of the table in (1.G), for some prime power q. Then there is a natural way to introduce lines so that the Buekenhout-Shult Theorem applies. (In particular, if m is not too small then

X can be identified with the set of points of a classical geometry, and G is a group of automorphisms of the geometry.)

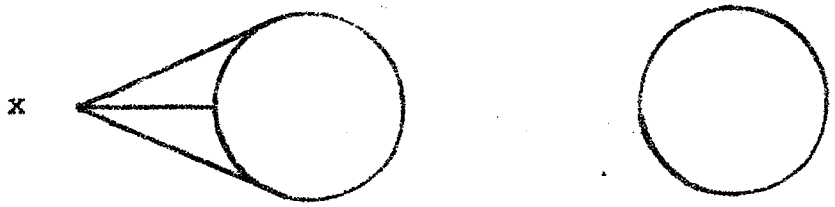
Remarks. 1. In each case where X corresponds to a classical geometry, the automorphism group of the geometry is known to be generated by all isometries, all field automorphisms, and suitable additional diagonal matrices.

2. The theorem says nothing about G itself (cf. (2.C)).

3. There are general numerical conditions which guarantee that the Buekenhout-Shult Theorem applies to a given rank 3 situation (cf. [16]). However, only one special case will be proved here:

Theorem. Suppose G has rank 3 on X, with subdegrees 1, $q(q^{n-2}-1)/(q-1)$, q^{n-1} for some prime q and some integer $n \geq 5$. Then n is even, and X can be identified with the set of all points of an $Sp(n,q)$ or an $O(n+1,q)$ geometry, in such a way that G acts on the geometry as a group of automorphisms.

Proof. Step 1. Graph.



$k = q(q^{n-2}-1)/(q-1)$ $\qquad$ $\ell = q^{n-1}$

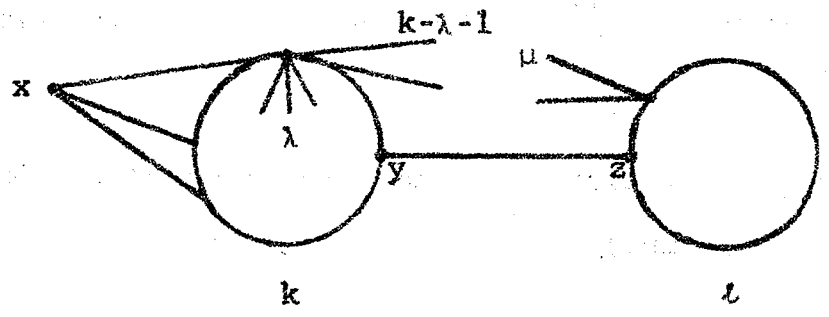
There is an orbit $(x,y)^G \subset X \times X$ of length $|X|k$. Call $x_1 \sim y_1$ iff $(x_1,y_1) \in (x,y)^G$. Call $x \not\sim z$ if $x \neq z$ and $x \sim z$ is false.

Note that G_x is transitive on the set of points $y \sim x$, as well as on the set of points $y \not\sim x$. Thus, since $k \neq \ell$, the relation $\sim$ is symmetric.

Step 2. Parameters.

If $x \sim y$, let λ denote the number of points $z \sim x,y$.

If $x \not\sim y$, let μ denote the number of points $z \sim x,y$.



The number of indicated pairs (y,z) is $k(k-\lambda-1) = \ell\mu$. Thus,

$$q\,\frac{q^{n-2}-1}{q-1}\,(k-\lambda-1) = q^{n-1}\mu, \quad \text{or}$$

$$\frac{q^{n-2}-1}{q-1}\,(k-\lambda-1) = q^{n-2}\mu, \quad \text{so}$$

$$q^{n-2} \mid k-\lambda-1 < k < 2q^{n-2}, \quad \text{and hence}$$

$$k-\lambda-1 = q^{n-2} \quad \text{and} \quad \mu = (q^{n-2}-1)/(q-1).$$

Step 3. Lines. (This step is independent of the particular k,ℓ,λ,μ.)

Call $x^{\perp} = \{y \mid y = x \text{ or } y \sim x\}$.

Clearly $y \in x^{\perp}$ iff $x \in y^{\perp}$. Let $x \sim y$, and note that $|x^{\perp} \cap y^{\perp}| = 2+\lambda$ (the number 2 counting the set $\{x,y\}$). Now define the line xy by

$$xy = \cap\{w^{\perp} \mid x,y \in w^{\perp}\} = \cap\{w^{\perp} \mid w \in x^{\perp} \cap y^{\perp}\}.$$

Claim: If $x' \neq y'$ and $x',y' \in xy$, then $xy = x'y'$.

Proof. $w \in x^{\perp} \cap y^{\perp} \Rightarrow x',y' \in xy \subseteq w^{\perp} \Rightarrow w \in x'^{\perp} \cap y'^{\perp}$. Use $w = x'$ to see that $x' \in y'^{\perp}$. Then $|x^{\perp} \cap y^{\perp}| = |x'^{\perp} \cap y'^{\perp}|$. But $x^{\perp} \cap y^{\perp} \subseteq x'^{\perp} \cap y'^{\perp}$. Thus, $x^{\perp} \cap y^{\perp} = x'^{\perp} \cap y'^{\perp}$, and hence $xy = x'y'$.

Claim: All lines xy have the same size $h+1$, say. (Use transitivity.)

Claim: The number of lines on x is k/h. (For, these lines, with x removed, partition $x^\perp - \{x\}$.)

Step 4. q-Groups. Let $x \sim y$. Let $P \in \mathrm{Syl}_q\, G_{xy}$ and $P < Q \in \mathrm{Syl}_p G_x$.

Since $|y^{G_x}| = q(q^{n-2}-1)/(q-1)$, we have $|Q : P| = q$. Since G_x is transitive on the q^{n-1} points not joined to x, so is its Sylow subgroup Q.

Claim: P is transitive on the set of $k - \lambda - 1$ points z such that $z \sim y$ and $z \neq x$.

Proof. Note that

$$q^{n-2} = k - \lambda - 1 \geq |z^P| = \frac{|P|}{|P_z|} = \frac{|Q|/q}{|P_z|} \geq \frac{|Q|}{|Q_z|}/q = q^{n-1}/q.$$

Step 5. Main Step. $h \geq q$. (Thus, lines are "big". In fact, it is unusual for rank 3 groups to have lines of size greater than 2.)

Proof. Assume that $h < q$.

Then P acts on the set $x^\perp \cap y^\perp - xy$, whose cardinality satisfies the condition $(\lambda + 2) - (h + 1) = \lambda + 1 - h \neq 0 \pmod q$. Thus, P fixes some $y' \in x^\perp \cap y^\perp - xy$.

If $y^\perp \cap y'^\perp \subseteq x^\perp \cap y^\perp$ then $y^\perp \cap y'^\perp = x^\perp \cap y^\perp$ (as both set have size $\lambda + 2$), and then $y' \in yy' = xy$.

Thus, there exists a point z in $y^\perp \cap y'^\perp$ but not in $x^\perp$. Clearly, P acts on the set of all such points z. By Step 4, $|z^P| = k - \lambda - 1$. This yields the contradiction

$$q^{n-2} = |z^P| \leq |y^\perp \cap y'^\perp| = \lambda + 2.$$

Step 6. B-S axioms. The first two axioms in (2.A) are obvious. Consider the third axiom: if $z \notin L$, we must show that $z^\perp$ contains one or all points of L.

We may assume that $x \in L$ and $z \notin x^\perp$. We must then prove that $|z^\perp \cap L| = 1$.

Firstly, $|z^\perp \cap L| \leq 1$. For, if $y_1, y_2 \in z^\perp \cap L$ with $y_1 \neq y_2$, then $x \in L = y_1 y_2 \subset z^\perp$, which is not the case.

Next, there are $\mu = (q^{n-2}-1)/(q-1) = k/q \geq k/h$ points of $x^\perp$ which lie on lines through z. (Use the definition of μ, together with Steps 2 and 5.) But each of the k/h lines on x has at most one such point (cf. Step 3). Thus, each such line meets $z^\perp$, as required.

Step 7. Completion. Since $\lambda + 2 \geq h + 1$, the set $x^\perp \cap y^\perp - xy$ is nonempty, and hence case (II) of the Buekenhout-Shult Theorem cannot occur. Comparison of the given k and ℓ with the table in (1.G) now completes the proof of the theorem.

Further remarks. 4. Only a little more care is needed to handle the case where q is a prime power.

5. Clearly, the precise parameters k and ℓ were not needed. What was needed was a large enough prime power divisor of ℓ, suitably related to k (cf. [16]).

6. The theorem is also true for $n=4$ when q is prime, but the prime power case remains open.

7. The above method of proof originated in the study of symmetric designs. It has very recently also been used by Cameron in the study of distance-transitive graphs.

(2.C) Perin's Method.

The results in (2.B) still leave open the question of determining G.

Theorem. Let G be an automorphism group of a symplectic, unitary or orthogonal geometry V. Assume that t.i. or t.s. lines exist, and that G acts as a rank 3 permutation group on the set of all points. Then G contains the corresponding group $PSp(V)$, $PSU(V)$ or $P\Omega(V)$, except for the case $G = A_6 < S_6 \cong Sp(4,2) \cong \Omega(5,2)$.

The symplectic and unitary cases of this theorem are due to Perin [23], except for the case $Sp(2m,2)$; the orthogonal case was handled by Kantor & Liebler [19], except for the case $\Omega(2m+1,2)$, using Perin's method; and the excluded cases were settled very recently by Cameron & Kantor using a different approach. More generally, Cameron & Kantor determined all automorphism groups which are transitive on the ordered pairs of non-perpendicular points.

We will only prove a special case of the theorem, concentrating primarily on Perin's method for proving such results. Further variations will then be discussed.

We begin with a basic number-theoretic result.

Theorem (Zsigmondy, 1892; see Dickson [12]). If q and n are integers greater than 1, then there is a prime p dividing q^n-1 but not dividing q^i-1 for $1 \le i < n$, except when either $n=2$ and $q+1=2^i$, or $n=6$ and $q=2$.

Lemma. Let q, n and p be as above (and not an exceptional case). Suppose that $A \le GL(k,q) = GL(V)$ with $|A| = p$. Then

(i) Each nontrivial irreducible constitutent of A on V has dimension divisible by n; and

(ii) If $k = n$ then

$$N_{GL(n,q)}(A) \cong GF(q^n)^* \rtimes Gal(GF(q^n)/GF(q)).$$

Proof. (i) If $k = Qn + R$ with $0 \leq R < n$, then $GL(k,q)$ and $GL(k - R,q)$ have the same p-Sylow orders. Thus, A must centralize an R-space and act on a complementary Qn-space. (N.B. - In fact, each nontrivial irreducible constituent has dimension n, as is seen by an examination of eigenvalues.)

(ii) By Schur's Lemma, $C_{GL(V)}(A) = GF(q^n)^*$, so we may identify V with $GF(q^n)$.

Let $g \in N_{GL(V)}(A)$, where we may assume that $1^g = 1$ (by modifying g using an element of $C_{GL(V)}(A)$). Write $A = \langle a \rangle$. Then every element of $GF(q^n)$ has the form $f(a)$ with $f(x) \in GF(q)[x]$. Since $1^g = 1$, $f(a)^g = f(a^g)$ and g is in the indicated Galois group.

Theorem. Let $G \leq Sp(2m,q)$, where $m \geq 3$ and $q > 3$. If G has rank 3 on the set of points, then $G = Sp(2m,q)$.

Proof. Step 1. Since $k = q(q^{2m-2}-1)/(q-1)$, there is a p dividing $q^{2m-2}-1$ as in Zsigmondy's Theorem. Fix $A < G$ with $|A| = p$.

Step 2. A is completely reducible. By the lemma, since $2m - 2 > 2$ we must have $\dim C_V(A) = 2$. Set $T = C_V(A)$. Now complete reducibility implies that $T^\perp$ is nonsingular and an irreducible A-space.

By the lemma (part (ii)), the group $N_G(A)^{T^\perp}$ induced on $T^\perp$ is metacyclic. But also, $N_G(A)^T \leq Sp(T) = SL(2,q)$.

Step 3. $N_G(T)^T = Sp(2,q)$. For $N_G(T)$ acts 2-transitively on the set of points of T. (Any two such points are not perpendicular, and span T. An element of G exists mapping such a pair to any given pair of this type, and must send T to itself.) Then $N_G(T)^T$ contains one Sylow group of order q, and hence all of them.

Also, $N_G(T) = C_G(T)(N_G(A) \cap N_G(T))$ by the Frattini argument. Thus, $N_G(A)^T = N_G(T)^T$ is $SL(2,q)$.

Step 4. Completion. The second commutator group $N_G(A)''$ induces the identity on $T^\perp$ and $SL(2,q)$ on T. Since $V = T \perp T^\perp$, $N_G(A)''$ contains a transvection group of order q. From the transitivity of G it follows that G contains all transvection groups. Thus, $G \geq Sp(2m,q)$.

Remark. The case $q = 3$ requires just a little more care, since $N_G(A)''$ merely contains an involution centralizing $T^\perp$ and inducing -1 on T.

Definition. The procedure in the above proof will be called Perin's Method.

We will outline several other uses for this method.

Example 1. If $G \leq SU(n,q)$, $n > 6$, and if G is transitive on both the set of all points and the set of all lines, then $G = SU(n,q)$.

Proof. The number of lines is $(q^n-\epsilon)(q^{n-1}+\epsilon)(q^{n-2}-\epsilon)(q^{n-3}+\epsilon)/(q-1)(q^2-1)$, where $\epsilon = (-1)^n$. Use a prime $p|q^{n-3}+\epsilon$, as in Zsigmondy's Theorem. (More precisely, if n is even, use $p|q^{2n-6}-1$.) Then $C_V(A)$ is a nonsingular 3-space, and $N_G(T)^T$ is transitive on points. Except for some difficulties with small values of q, Perin's Method yields the result.

Example 2. Subgroups of orthogonal groups having rank 3 on the set of points are handled similarly.

Example 3. If G is inside $Sp(V)$, $SU(V)$ or $\Omega(V)$ with $\dim V \geq 9$, and if G is transitive on points, lines and planes (t.i. or t.s. 3-spaces), then G is $Sp(V)$, $SU(V)$ or $\Omega(V)$, respectively. This and many similar results are found in Kantor & Liebler [19].

Example 4 (Perin [23]). Let $G \leq GL(n,q)$ with $q > 2$ and $n > 4$. If G is transitive on the ordered triples of independent points (i.e., 1-spaces), then $G \geq SL(n,q)$.

This time, use $p|q^{n-2}-1$, where $q^{n-2}-1$ divides the number of such triples. Then $V = C_V(A) \oplus [V,A]$ by complete reducibility (where $[V,A] = \langle v^a - v | v \in V,\ a \in A \rangle$). Here, $\dim C_V(A) = 2$ and $N_G(A)''$ contains a transvection group (unless $q = 3$, where a further argument is needed). It follows readily that G contains all transvection groups, and hence contains $SL(n,q)$.

Remark. Stronger results are now known. Suppose that $G \leq GL(n,q)$ and that G is transitive on i-spaces for some i with $1 \leq i \leq n-1$. The cases $i = 1$ or $n-1$ remain open. Suppose that $2 \leq i \leq n-2$. Then it can be shown that G is 2-transitive on points, unless $G \cong Z_{31} \rtimes Z_5$, inside $GL(5,2)$.

These 2-transitive groups have been studied by Wagner, Higman, Kornya, Orchel, and Cameron & Kantor. They were recently all determined. More generally, in unpublished work Cameron & Kantor determined all subgroups of $GL(n,q)$ which are transitive on the pairs consisting of an $n-1$-space and a point not in it; although the basic ideas are quite different from those presented here, Perin's method is eventually required.

Related results. 1. Suppose that G is a Chevalley group, and B is a Borel subgroup. Then Seitz [27]

has determined all subgroups K transitive on the conjugates of B. (In the case of classical groups, this means that the group K is assumed transitive on the sequences $V_1 < \ldots < V_m$ of t.i. or t.s. subspaces, where $\dim V_i = i$ and m is as large as possible. Thus, in the case of classical groups of sufficiently large dimension, Example 3 contains this result; however Seitz's result was required for the stronger result.)

2. Let $G \leq GL(n,q)$, $n > 3$. Assume that G induces a primitive rank 3 group on the set of points of the underlying vector space. In unpublished work, Perin showed that G must preserve a symplectic geometry (and hence $G \geq Sp(n,q)$ or $G = A_6 < Sp(4,2)$, by the first theorem stated in this section).

(2.D) Generation

Problem (*). Determine all irreducible subgroups K of $SL(n,q) = SL(V)$ generated by transvections.

Notation. Let $a = \langle w \rangle$ be a point (1-space) and $A = \ker f$ a hyperplane ($n-1$-space), where $0 \neq f \in V^*$. Set $T(a,A) = \{v \to v + \alpha f(v)w \mid \alpha \in GF(q)\}$ whenever $a \leq A$. Then $T(a,A) \cong GF(q)^+$, and $T(a,A)$ is precisely one of root groups for $SL(V)$.

Remark. If $t \in T(a,A) - \{1\}$ and if W is a t-invariant subspace of V, then either $a \leq W$ or $W \leq A$. (This follows immediately from the definition of $T(a,A)$.)

McLaughlin [20,21] answered (*) when K is generated by full transvection groups $T(a,A)$. He even handled infinite fields.

Theorem (McLaughlin). Suppose that G is an irreducible subgroup of $SL(n,F)$ generated by root groups. Then one of the following holds:

(i) $G = SL(n,F)$;

(ii) $G = Sp(n,F)$; or

(iii) $|F| = 2$ and G is $O^{\pm}(n,2)$ or S_{n+2} with n even, or S_{n+1} with n odd. In each case, G is embedded in a "natural" manner.

We have seen each of these examples: (ii) in (1.C) and $O^{\pm}(n,2)$ in (1.E), while the case of symmetric groups occurs as in the example of S_6 in (1.C).

More can be said in the finite case.

Theorem (Piper [24,25], Wagner [31]). Assume that G is an irreducible subgroup of $SL(n,q)$, $n \geq 3$, and that G is generated by transvections. Then G is one of the following:

(i) $SL(n,q')$, $Sp(n,q')$, $SU(n,q')$, $q = q'^i$;

(ii) $O^{\pm}(n,q')$, $q = q'^i$, q and n even;

(iii) $3 \cdot A_6 < SL(3,4^i)$ (a central extension of A_6 by Z_3);

(iv) $S_{n+\varepsilon} < SL(n,2^i)$, $\varepsilon = (2,n)$; or

(v) $3 \cdot PSU(4,3) \cdot 2 < SL(6,4^i)$.

(vi) $(Z_a)^{n-1} \rtimes S_n < SL(n,2^i)$ with $1 \neq a | 2^i - 1$.

In each case, G is embedded in a "natural" manner.

We will not prove this, but will give one application of it in (2.E). Applications of McLaughlin's result are given in (2.F,G). Some versions were implicitly used in (2.C).

Remarks. 1. $O_p(G) = Z(G)$ and $C_V(G) = 0$ can be substituted for irreducibility here, with only slight modifications [17].

2. Similar results exist for groups generated by long root elements of the remaining classical groups [17].

3. The exceptional examples (iii) and (v) are intimately related to the existence of some sporadic groups.

4. For $n = 2$, further examples occur: $SL(2,5) < SL(2,9^i)$, and dihedral groups if q is even.

5. If G is merely an irreducible group containing a nontrivial transvection, let G^* denote the subgroup generated by all transvections in G. Then $N_{GL(n,q)}(G^*)$ is simply a wreathed product, by Clifford's Theorem.

(2.E) Permutation representations: degrees

In 1832, Galois stated the following result.

Theorem. If $K < G = SL(2,q)$ and q is prime, then $|G:K| \geq q+1$ except when $q \leq 11$; and equality holds iff K is a Sylow q-normalizer.

The corresponding result was proved for q a prime power more than 50 years later (the only further exception occurring when $q = 9$; cf. Dickson [11, Ch. 12]). So were the analogous results for $SL(3,q)$, $SU(3,q)$, and (when q is odd) $Sp(4,q)$, by enumerations of all maximal subgroups. 150 years after Galois, $SL(n,q)$ was handled in an unpublished Ph.d. thesis by W. Patton [22]:

Theorem 1. If $K < G = SL(n,q)$, $n \geq 3$, then $|G : K| \geq (q^n-1)/(q-1)$, except for the case $K = A_7 < A_8 = SL(4,2)$. Moreover, equality holds iff K is the stabilizer of a point or a hyperplane.

Proof. Assume that $|G : K| \leq (q^n-1)/(q-1)$. Note that the latter number is both the number of points and the number of hyperplanes, and is greater than the number of i-spaces whenever $1 < i < n-1$. We may thus assume that K is irreducible.

If K contains a nontrivial transvection, then the results of the last section determine all possibilities for K, and none satisfies the desired bound.

We may thus assume that K has no nontrivial transvections. We must show that K is A_7 inside $SL(4,2)$.

Take any hyperplane H. Set $P = C_G(V/H)$ and $Q = C_G(H) \cap C_G(V/H)$.

$$P: \begin{pmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & SL(n,q) & \\ 0 & & & \end{pmatrix} \qquad Q: \begin{pmatrix} 1 & * & \cdots & * \\ & 1 & & \\ & & \ddots & 0 \\ & & 0 & \ddots \\ & & & 1 \end{pmatrix}$$

Thus, Q consists entirely of transvections, and $|Q| = q^{n-1}$.

Since $Q^g \cap K = 1$ for all $g \in G$, Q acts semiregularly on the cosets of K. Thus, $|G:K| \equiv 0 \pmod{q^{n-1}}$. But $|G:K| \leq (q^n-1)/(q-1) < 2q^{n-1}$. Thus, $|G:K| = |Q|$, so $|G| = |QK|$ and hence $G = QK$. Since Q fixes an i-space for each i, it follows that K is transitive on i-spaces for each i! Results in (2.C) now apply, and we deduce that K is either $SL(n,q)$ or A_7 inside $SL(4,2)$.

The same general idea produces significant improvements if n is not too small.

Theorem 2. Let $K < SL(n,q)$, where q is odd and $q > 11$. If K is irreducible, then either $|G : K| > q^{\frac{1}{2}n(n-1)}$ or $G \geq Sp(n,q)$.

We will use induction on n. However, when we reduce to a smaller dimensional situation, we will no longer be able to guarantee the irreducibility of our group. For example, if $K = Sp(n,q)$ then the stabilizer of any hyperplane acts reducibly on the hyperplane. Therefore, we will prove a modification of Theorem 2. Recall that, by McLaughlin's results (cf. (2.D)), it suffices to proves that K contains a full transvection group.

Theorem 2'. Let $K \leq SL(n,q) = SL(V)$, where q is odd and $q > 11$. If $|G : K| \leq q^{\frac{1}{2}n(n-1)}$, then K has a subgroup $S \cong SL(2,q)$ generated by two transvection groups; moreover, $V = L \oplus C_V(S)$, where L is a 2-space on which S acts in the natural manner.

Proof. The case $n = 2$ is handled using Dickson [11, Ch. 12], so suppose that $n \geq 3$. Let H, P and Q be as before. Let $(K \cap P)^H$ denote the group induced by $K \cap P$ on H.

Case 1. $K \cap Q = 1$ for some choice of H.

Note that

$$q^{\frac{1}{2}n(n-1)} \geq |G:K| \geq |P:K\cap P| = |P:(K\cap P)Q|\cdot|(P\cap K)Q:P\cap K|$$
$$= |P/Q:(K\cap P)Q/Q|\cdot|Q:P\cap K\cap Q|$$
$$= |P^H:(K\cap P)^H|q^{n-1}$$

since $K\cap Q=1$. Thus, $|P^H:(K\cap P)^H| \leq q^{\frac{1}{2}(n-1)(n-2)}$. By induction, $(K\cap P)^H$ contains a subgroup S^H of the desired type, where $S \leq K\cap P$.

This group S may not meet our requirements, but we will show that a suitable subgroup of S does. We may assume that $S=S'$. Then, since S centralizes both V/H and H/L, it centralizes V/L.

Let z be an involution in S. Then $S=N_S(\langle z\rangle)\,C_S(H)$ by the Frattini argument. We may thus replace S by $C_S(z)$, and assume that $z\in Z(S)$. Since S centralizes V/L, $L=C_V(-z)$. Now S preserves the decomposition $V=L\oplus C_V(z)$, and centralizes V/L, so $V=L\oplus C_V(S)$, as required.

Case 2. $K\cap Q\neq 1$ for every choice of H.

Here, induction does not apply, but there are enough transvections to generate the desired $SL(2,q)$. Let W be a minimal K-invariant subspace. If $H\not\geq W$, then $K\cap Q$ does not move W, so that $\dim W>1$ and K^W contains a nontrivial transvection centralizing $W\cap H$.

The results of the preceding section imply that $K^W \geq SL(W)$ or $Sp(W)$. There are thus transvection groups $T(a,A)$ and $T(b,B)$ inside K such that the group $S = \langle T(a,A), T(b,B) \rangle$ induces $SL(2,q)$ on W. Then $V = L \oplus C_V(S)$ with $L = \langle a,b \rangle$ and $C_V(S) = A \cap B$. This produces the desired S, and completes the proof of Theorems 2' and 2.

Remarks. 1. There are analogues of Theorem 2 for arbitrary q [18]. The main difficulty is not the case $n = 2$, but rather the last paragraph of Case 1. For q even, there might not be a subgroup S generated by transvections of V. This is tied up with the first cohomology groups of the classical groups.

2. Cooperstein [9] has obtained the precise minimum value of $|G : K|$ for proper subgroups K of $Sp(n,q)$, $SU(n,q)$ and $\Omega^{\pm}(n,q)$. The methods are similar to the preceding ones. Analogues of Theorem 2 have also been obtained for all the classical groups [18].

3. Suitable modifications of all of these results exist when G is obtained from a classical group by adjoining outer automorphisms.

4. The case of the remaining Chevalley groups remains open. This would, however, be settled if all their

subgroups K were determined such that $O_p(K) = 1$ and K is generated by (long) root elements.

5. It seems very likely that $|G : K| > |G : B|$ in Theorem 2. The corresponding type of results for the other classical groups should also be true.

(2.F) Permutation Representations: Arbitrary Rank

The next two topics concern permutation representations of classical (and Chevalley) groups. In this section the permutation rank will be arbitrary; in the next one, ranks 2 and 3 will be considered. The following result was conjectured by Peter M. Neumann in 1973.

Theorem. Given an integer r, only finitely many presently known finite simple groups have presently unknown primitive rank r permutation representations.

Here, "unknown" essentially means that the one-point stabilizer is transitive for A_n, irreducible for classical groups, and non-parabolic for the remaining Chevalley groups. (However, a little more care is needed in the case of symplectic groups in characteristic 2.) This theorem follows immediately from the next three theorems.

Theorem (Bannai [4]). Given $r>1$, if $G=S_n$ or A_n has a (faithful) primitive rank r permutation representation on G/K, where K is transitive on the original n points, then $n \leq 6r+2$.

Theorem (Seitz [28]). Given $r>1$ and $\ell>1$, there there is a number $Q(r,\ell)$ with the following property: if G is a rank ℓ Chevalley (or twisted) group defined over $GF(q)$ with $q>Q(r,\ell)$, and if G has a rank r permutation representation on G/K, then $\langle U^P \rangle \trianglelefteq K^g \trianglelefteq P$ for some parabolic subgroup P and some $g \in G$ (where $U \in Syl_p G$ for the prime $p|q$).

Theorem [18]. Given $r>1$, if an n-dimensional classical group G (other then $Sp(2n,q)$ with q even) has a faithful primitive rank r permutation representation on G/K with K irreducible, then $n \leq 16r$.

A modification of this result also holds for the excluded symplectic groups. The problem is that $Sp(2n,q)$ has primitive rank r permutation representations with r independent of n. (For example, we will see in the next section that $Sp(2n,2)$ has 2-transitive representations.) Such representations arise because

$Sp(2n,q) \cong \Omega(2n+1,q)$ for q even, and K can leave invariant a subspace other than $V^{\perp}$, where V is the $2n+1$-dimensional space for $\Omega(2n+1,q)$ (cf. (1.F)).

We will prove versions of these results in the case $G = SL(n,q)$.

Theorem 1. Given $r > 1$ and an odd q with $q > 11$, if G has a primitive rank r representation on G/K with K irreducible, then $n \leq 4r - 3$.

Lemma 1. If G is a primitive rank r permutation group on X, and if n_i is the length of an orbit of G_x on $X - \{x\}$ (where $x \in X$), then $|X| \leq 2n_i^{r-1}$.

Proof. Let G_x have orbit lengths $1 = n_1 \leq n_2 \leq \ldots \leq n_r$. Then $n_{i+1} \leq n_2 n_i$ (Wielandt [32, 17.4]), so $|X| = \Sigma n_i \leq \sum_0^{r-1} n_2{}^i \leq 2n_2{}^{r-1} \leq 2n_i{}^{r-1}$.

Lemma 2 (Maillet, 1895, and Bannai [4]). If G is a primitive rank r permutation group on X, and if $1 \neq g \in G$, then $|X| \leq 2|G : C_G(g)|^{r-1}$.

Proof. Let $x^g \neq x$, and let $\Gamma = (x^g)^{G_x}$. If $y \in \Gamma$ then some conjugate of g sends x to y. Thus, $|\Gamma| \leq |G : C_G(g)|$, so $|X| \leq 2|\Gamma|^{r-1} \leq 2|G : C_G(g)|^{r-1}$ by Lemma 1.

Proof of Theorem 1. In (2.E) it was shown that $|G : K| \geq q^{\frac{1}{2}n(n-1)}$ or $K \geq Sp(n,q)$. In the latter case, the rank r is easily found to be large. Thus, by Lemma 2 we must have $q^{\frac{1}{2}n(n-1)} \leq 2|G : C_G(g)|^{r-1}$ for any g not in the kernel of the permutation representation. Choose g to be a nontrivial transvection. Then $|G : C_G(g)|$ is the number of pairs (v,H) with H a hyperplane and v a nonzero vector in H, so $|G : C_G(g)| = (q^n - 1)(q^{n-1} - 1)/(q-1)$. Use of arithmetic now yields $n \leq 4r - 3$.

Remark. The proof of Bannai's bound $n \leq 6r+2$ proceeds in a very similar manner. In his case, G is S_n or A_n, and the bound $|G:K| \geq \frac{1}{2}[\frac{1}{2}(n+1)]!$ was used; this is essentially just a familiar bound due to Bochert (in 1889).

Theorem 2 (Seitz). Given $r>1$ and n, suppose that that $q > 16\{r(n-1)! + (r-1)((n-1)!)^2\}^2$. If $G = SL(n,q)$ has a rank r permutation representation on G/K, then K contains a full root group (i.e., a transvection group).

Remarks. 1. Theorem 2 is much harder than Theorem 1.

2. McLaughlin's result (mentioned in (2.D)) determines all irreducible K here. None meets the stated bound for r; hence, the subgroup K must be reducible. If the permutation representation is primitive, then K must be a maximal parabolic subgroup.

3. Our proof is part of Seitz's proof of the much more general theorem; only slight modifications are needed to prove the general result.

The proof of Theorem 2 consists of two parts. Part I uses elementary character theory to bound the number of U-orbits on G/K. Part II is geometric and inductive.

Notation. $G = SL(n,q)$.

$W \cong S_{n-1}$ is the Weyl group.

$U \in Syl_p G$ (where $p|q$).

$B = N_G(U) = UH$ with H abelian of order $(q-1)^{n-1}$.

Part I. We will show that the number of U-orbits on $\Omega = G/K$ is at most $r|W| + (r-1)|W|^2$; or, equivalently, that

$$(1_K^G, 1_U^G) \leq r|W| + (r-1)|W|^2.$$

Recall that B consists of upper triangular matrices of determinant 1, while U consists of those having 1's on the diagonal. Since we are assuming that $q > 2$, we have $B' = U$.

Write $1_K^G|_B = m1_B + \sum_{j=1}^{d} \varphi_j + \Sigma\tau_i$,

where the φ_j's and τ_i's are non-principal irreducible characters of B, the φ_j's are linear, and the τ_i's are non-linear; the φ_j's need not all be different, nor need the τ_i's be. Now $1_K^G|_U = m1_U + \sum_{j=1}^{d} \varphi_j|_U + \Sigma\tau_i|_U$

so

$$(1_K^G|_U, 1_U) = m + \sum_{j=1}^{d} (\varphi_U|_U, 1_U) + \Sigma(\tau_i|_U, 1_U).$$

The desired inequality for $(1_K^G|_U, 1_U)$ is then an immediate consequence of the following four assertions:

(1) $(\varphi_j|_U, 1_U) = 1$

(2) $(\tau_i|_U, 1_U) = 0$

(3) $m \le r|W|$, and

(4) $d = \sum_{j=1}^{d} (\varphi_U|_U, 1_U) \le (r-1)|W|^2$.

Proof of (1). Since φ_j is linear and $B' = U$, we have $\varphi_j|_U = 1_U$.

Proof of (2). $(\tau_i|_U, 1_U) = (\tau_i, 1_U{}^B)$, where $1_U{}^B$ is the regular character of $H \cong B/U$. Since H is abelian, each irreducible constituent of $1_U{}^B$ is linear, and hence $(\tau_i, 1_U{}^B) = 0$.

Notation. $1_K^G = 1_G + \Sigma a_i \chi_i$, where the χ_i's are distinct non-principal irreducible characters of G.

Since $r = (1_K^G, 1_K^G) = 1 + \Sigma a_i{}^2$, we have $r - 1 \ge \Sigma a_i$.

Proof of (3). By definition, $m = (1_K^G|_B, 1_B) = (1_K^G, 1_B^G)$ $= 1 + \Sigma a_i (\chi_i, 1_B^G) \le 1 + \Sigma a_i (1_B^G, 1_B^G) = 1 + \Sigma a_i |W| \le r|W|$. (Note that $(\chi_i, 1_B^G)$ is at most the norm of 1_B^G, which is in turn the number of (B,B) double cosets.)

Statement (4) is harder. We will need the following fact.

Lemma. Suppose that φ and ψ are linear characters of B. If φ and ψ are both constituents of $\chi|_B$

for some irreducible character χ of G, then $(\psi|_H)^w = \varphi|_H$ for some $w \in W$. (As usual, we can identify W with $N_G(H)/H$.)

Proof. $0 < (\varphi, \chi|_B) = (\varphi^G, \chi)$. By Mackey's subgroup theorem,

$$0 < (\chi,\chi) \le (\varphi^G, \psi^G) = \sum_{w\in W} (\varphi^{w^{-1}}|_{B\cap B^w}, \psi|_{B\cap B^w}).$$

But $B \cap B^w = (U \cap U^w)H$ since $B = UH$ and $H^w = H$. Since $0 \neq (\varphi^{w^{-1}}|_{B\cap B^w}, \psi|_{B\cap B^w})$ for some $w \in W$, it follows that

$$0 \neq (\varphi^{w^{-1}}|_{(U\cap U^w)H}, \psi|_{(U\cap U^w)H}) = (\varphi^{w^{-1}}|_H, \psi|_H).$$

(Note that U is in the kernels of φ and ψ, so $U \cap U^w$ is in the kernels of $\varphi^{w^{-1}}$ and ψ.) But $\varphi^{w^{-1}}|_H$ and $\psi|_H$ are linear characters, and w normalizes H, so the lemma follows.

Proof of (4). $1_K^G = 1_G + \Sigma a_i \chi_i$. Write $\chi_i|_B = \Sigma m_j \varphi_{ji}$ + (some linear combination of 1_G and τ_i's), where now these φ_{ji}'s are distinct and each $m_j > 0$. We will prove that

(4') For each i, the number of φ_{ji} here is at most $|W|$, and

(4") Each $m_j \le |W|$.

For then the total number of non-principal linear constituents of $\chi_i|_B$ will be at most $|W|^2$, and hence the total number d of φ_j's will be at most $\Sigma a_i |W|^2 \leq (r-1)|W|^2$, so (4) will hold.

Proof of (4'). The lemma implies that $\chi_i|_B$ involves at most $|W|$ linear characters.

Proof of (4''). By definition, $m_j = (\chi_i|_B, \varphi_{ji}) = (\chi_i, \varphi_{ji}{}^G) \leq (\varphi_{ji}{}^G, \varphi_{ji}{}^G)$. But

$$(\varphi_{ji}{}^G, \varphi_{ji}{}^G) = \sum_{w \in W} (\varphi_{ji}{}^{w-1}|_{B\cap B^w}, \varphi_{ji}|_{B\cap B^w}) \leq |W|$$

since $\varphi_{ji}{}^{w-1}|_{B\cap B^w}$ and $\varphi_{ji}|_{B\cap B^w}$ are linear. Thus, $m \leq |W|$.

This ends Parts I.

Part II. Suppose that $G = SL(n,q) > K$, and that U has at most ℓ orbits on G/K. If $q > 16\ell^2$, then K contains a full root group (i.e., transvection group).

Note that this, together with Part I, will complete the proof of Theorem 2.

Proof. We will use induction on n, the case $n = 1$ being vacuous.

Let E be a hyperplane. Set $P = C_G(V/E)$ and $Q = C_P(E)$. Then Q consists of transvections, and $P = QR$ with $R \cong SL(n-1,q)$. We may assume that $Q \leq U \leq P$ and $N_G(U) = UH$.

With respect to a suitable basis, P, Q and R are as follows.

P = all $\begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix}$ with $\det A = 1$, $v \in GF(q)^{n-1}$, and $0 = (0,\ldots,0)$.

Q = all such matrices with $A = 1$.

R = all such matrices with $v = 0$.

U = all lower triangular matrices.

H = all diagonal matrices.

Write $\Omega = G/K$, and let $\alpha \in \Omega$ with $G_\alpha = K$.

Write $\alpha^{QR} = \bigcup_{i=1}^{m} \Phi_i$, where the Φ_i's are Q-orbits in α^{QR}.

Clearly, R is transitive on $\{\Phi_1,\ldots,\Phi_m\}$. Also, $P = QR$ implies that $U = Q(U \cap R)$, so every $U \cap R$-orbit on $\{\Phi_1,\ldots,\Phi_m\}$ arises from a U-orbit on α^{QR}. Thus, $U \cap R$ has at most ℓ orbits on $\{\Phi_1,\ldots,\Phi_m\}$.

By induction, there is a root group $X \leq R$ stabilizing some Φ_i. (If $n = 2$, take $X = 1$.) We may assume that $X \leq Z(U \cap R)$.

Let $\beta \in \Phi_i \subseteq \beta^U = \Lambda_1$, and write $\beta^{UH} = \bigcup_{i=1}^{k} \Lambda_i$ with each Λ_i a U-orbit. (Here, H is defined as above.) By hypothesis, $k \leq \ell < (\sqrt{q}+1)/4$. Let H_0 be the stabilizer of Λ_1 in H. Then $|H : H_0| = k < (\sqrt{q}+1)/4$.

Note that H_0 fixes some point $\gamma = \beta^u \in \Lambda_1$, where $u \in U$. For, both U and UH_0 are transitive on Λ_1, so $UH_0 = U(UH_0)_\beta$. Since H_0 is a Hall subgroup of the solvable group UH_0, it must be conjugate to a subgroup of $(UH_0)_\beta$, and our assertion follows.

Now $\langle H_0, U_\beta^u \rangle \leq G_\gamma$, so H_0 normalizes $U_\gamma = U_\beta^u$.

At this point we must stop to deal with the case $n = 2$. Here, Part II requires that $q \mid |K|$. To see that this holds observe that $2(\sqrt{q} - 1) < |H_0| \mid |G_\gamma| = |K|$. It is then easy to use the list of subgroups in Dickson [11, Ch. 12], along with the assumed bound on the number of U-orbits, in order to obtain $q \mid |K|$.

Now suppose that $n \geq 3$, so $|X| = q$. Since both Q and QX are transitive on Φ_i, we have $QX = Q(QX)_\beta$ with $1 \neq (QX)_\beta \leq U_\beta$. Thus, $U_\gamma \neq 1$.

Now U_γ is a nontrivial p-group normalized by H_0. Let Y be a minimal H_0– invariant subgroup of U_γ.

The group U is a product of root groups, each normalized by H. Moreover, H acts irreducibly on each such root group, inducing a fixed-point-free group of order $q - 1$. Since $|H : H_0| < (\sqrt{q} + 1)/4$, H_0 acts irreducibly on each such root group.

Thus, H_0 acts on Y as it does on some H-invariant root group $X^* = T(a, A)$. (As in (2.D), A is a hyperplane and a is a 1-space of A.) But $C_H(X^*)$ decomposes A/a into the direct sum of $n - 2$ inequivalent $C_H(X^*)$-modules, while a and V/A are $C_H(X^*)$-isomorphic. The same must also hold for $C_{H_0}(Y)$. Since $C_{H_0}(Y)$ is

a group of diagonal transformations, Y induces the identity on some $n-2$-space of V and acts on a complementary 2-space. Thus, $\dim C_V(Y) = n-1$, and Y is a full transvection group (in fact, $Y = X^*$).

This completes the proof of Theorem 2.

Open Problem. Significantly decrease the bounds on n and q in Theorems 1 and 2.

(2.G) Permutation Representations: Ranks 2 and 3.

The crudeness of the bounds in Theorems 1 and 2 of (2.F) is already seen when $r = 2$ or 3. For, in those cases, if $n = 2$ then $q \leq 11$, while if $n > 2$ then $n \leq 4$ and $q \leq 3$. More precise answers are, in fact, known. Curtis-Kantor-Seitz [10] contains the determination of all the 2-transitive permutation representations of all the Chevalley groups G; Kantor-Liebler [19] does the same for the rank 3 representations of the classical groups. (The corresponding result remains open in the case of the exceptional Chevalley groups.) Moreover, in both papers the group G is replaced by any subgroup of $\mathrm{Aut}(G/Z(G))$ containing $G/Z(G)$.

There are three basic tools required for these results:

(I) Properties of 1_B^G (compare the last half of (2.F)).

(II) Characterizations of "large" subgroups.

(III) The Pigeon-Hole Principle.

We will only discuss the first two of these.

(I) 1_B^G. Let p be the characteristic of G.

Theorem (Steinberg, Green, Howlett, Hoefsmit, Liebler, Benson, Grove, Surowski; cf. [10, pp. 57-58] and [19, §4]). Every irreducible constituent of $1_B^G - 1_G$ has degree divisible by p, except when G has type $Sp(2m,2)$, $G_2(2)$, $G_2(3)$, ${}^2F_4(2)$ or $F_4(2)$. (Moreover, in each of these cases all offending characters have been determined.)

Example 1. The permutation character for the action of $Sp(2m,q)$ on points splits as $1_G + \rho + \sigma$ with

$$\rho(1) = \frac{q(q^n-1)(q^{n-1}+1)}{2(q-1)} \quad \text{and} \quad \sigma(1) = \frac{q(q^n+1)(q^{n-2}-1)}{2(q-1)}.$$

In particular, $\rho(1)$ and $\sigma(1)$ are odd in the case of $Sp(2m,2)$; there is exactly one further offending character in this case.

Example 2. Let θ_1 denote the permutation character for the action of $G = SL(n,q)$ on i-spaces. Then $\theta_1 = 1_G + \chi_1$ with χ_1 irreducible, since G is 2-transitive on points.

Claim. If $2 \le i \le n/2$ then $\theta_i - \theta_{i-1} = \chi_i$ is irreducible and $\theta_i = 1 + \chi_1 + \ldots + \chi_i$. Moreover, the characters $1, \chi_1, \chi_2, \ldots$ are all distinct.

Proof. Let $1 \le j \le i \le n/2$. Then (θ_i, θ_j) is the number of orbits of G on ordered pairs (V_i, W_j) with $\dim V_i = i$ and $\dim W_j = j$. Since $i+j \le n$, $\dim V_i \cap W_j$ can be any number from 0 to j. Thus, $(\theta_i, \theta_j) = j+1$. In particular, $(\theta_i - \theta_{i-1}, \theta_i - \theta_{i-1}) = i+1-2(i-1+1) + (i-1+1) = 1$. Also, $\theta_i(1) - \theta_{i-1}(1)$ is the number of i-spaces minus the number of $i-1$-spaces, which is positive (and a multiple of q). This proves both the claim and the following

Corollary. $q \mid \chi_i(1)$ for $1 \le i \le n/2$.

(II) Characterizations.

(a) Seitz [27] determined all $K < G$ for which $(1_B^G, 1_K^G) = 1$. This, and related results, were mentioned in (2.C).

(b) Tits' Lemma. If $U \le K < G$ (where $U \in \mathrm{Syl}_p G$), then $\langle U^P \rangle \le K \not\ni P$ for some parabolic subgroup P.

Remark. In the cases $SL(n,q)$, $Sp(n,q)$ and $SU(n,q)$, U contains transvection groups, and much stronger results are available (cf. (2.D)). However,

this lemma holds for all Chevalley groups, and even remains valid in the case of infinite fields.

Proof. In the usual BN notation, $G = BNB = UHNHU = UHNU$. Let $hn \in K$ $(h \in H, \ n \in N)$, so $n = n_w$ represents some $w \in W = N/H$. Suppose that $w = w_1 s_1$ with $\ell(w_1) < \ell(w)$ and s_1 a fundamental reflection. If α_1 is the root corresponding to s_1, then $\alpha_1^{w_1^{-1}} > 0$, and hence

$$U_{\alpha_1}^{s_1} = (U_{\alpha_1}^{w_1^{-1}})^w \leq U^w = U^{hw} = U^{hn} \leq K.$$

Thus, $\langle U_{\alpha_1}, U_{\alpha_1}^{s_1} \rangle \leq K$. But there is a coset representative $\tilde{s}_1$ of s_1 in $\langle U_{\alpha_1}, U_{\alpha_1}^{s_1} \rangle$. (That is, $\tilde{s}_1 \in \langle U_{\alpha_1}, U_{\alpha_1}^{s_1} \rangle$ and $H\tilde{s}_1$ is s_1.) Thus, $\tilde{s}_1 \in K$, and $hn\tilde{s}_1 \in K$. This process can clearly be continued, gradually reducing $\ell(w)$. If P is generated by B and all elements $\tilde{s}_1$ arising as above, then P is parabolic and $KH = P$.

2-Transitive Representations. Suppose that G acts 2-transitively on the set of cosets of K, where $K < G$. Write $1_K^G = 1 + \chi$ with χ irreducible.

Clearly $(1_B^G, 1_K^G) = 1 + (1_B^G, \chi)$.

If $\chi \neq 1_B^G$, then Seitz's theorem lists all potential subgroups K.

If $\chi \in 1_B^G$, then $|G:K| = 1 + \chi(1)$, where $p|\chi(1)$ for most of our groups G (cf. (I)). Therefore, assume that $p|\chi(1)$. Then $|G:K| \equiv 1 \pmod{p}$, so $K^g \geq U$ for some $g \in G$, and Tits' Lemma implies that K^g is parabolic.

Thus, the determination of all pairs (G,K) is reduced to the checking of specific potential pairs, except in the case of specific groups G defined over small fields.

Case $Sp(2m,2)$. This is the most interesting case in which offending characters appear: 2-transitive representations occur, and fit nicely into the framework of the above argument. We first prove the existence of these permutation representations, and then prove a version of uniqueness.

Existence. $G = Sp(2m,2) \cong \Omega(2m+1,2)$ acts on a $2m+1$-dimensional vector space $r^\perp$, fixing a nonzero vector r, as in (1.F). There is a quadratic form Q available, and we may assume that $Q(r) = 1$.

If $v \in r^\perp - \langle r \rangle$, then $Q(v) = 1$ iff $Q(v+r) = 0$. Thus, the permutation character θ of the action of G on vectors $v \neq r$ with $Q(v) = 1$ coincides with that of G on the points of the symplectic space $r^\perp/\langle r \rangle$.

Moreover, the character of G on $r^\perp - \langle r \rangle$ is 2θ, so that of G on the 1-spaces of $r^\perp$ is $1+2\theta$.

Then the character of the action of G on hyperplanes of $r^\perp$ is also $1+2\theta$. (In Example 2, $\theta_1 = \theta_{n-1}$ since $(\theta_1,\theta_{n-1}) = (\theta_1,\theta_1) = (\theta_{n-1},\theta_{n-1}) = 2$.) The character on the set of hyperplanes containing $\langle r \rangle$ is also θ.

Hence, the character on the set of hyperplanes not containing $\langle r \rangle$ is $1+\theta = (1+\rho)+(1+\sigma)$. Each such hyperplane has stabilizer $O^{\pm}(2m,2)$. Hence, G must act 2-transitively on the hyperplanes of each type, with permutation character $1+\rho$ or $1+\sigma$.

Uniqueness. Suppose that $\varphi = 1_K^G = 1+\chi$ is $1+\rho$ or $1+\sigma$. We must show that K is $O^{\pm}(2m,2)$, embedded in G as above. We will regard G as $Sp(2m,2)$.

Let g be a nontrivial transvection in G. Count the pairs (K_1,g_1) for which K_1 is a conjugate of K containing the conjugate g_1 of g:

$$\varphi(1)\,|g^G \cap K| = |G : C_G(g)|\,\varphi(g).$$

(Here, $\varphi(1)$ is the number of conjugates K_1, and $\varphi(g)$ is the number of such conjugates containing g.)

The numbers $\varphi(1)$ and $\varphi(g)$ are computable using the known 2-transitive representation with

character φ (whose existence was proved above). Thus, $|g^G \cap K|$ is the number of transvections in $0^{\pm}(2m,2)$ (for a suitable choice of sign). The result of McLaughlin [2]] mentioned in (2.D) now yields the required determination of both K and its embedding.

Remarks. 1. All the other offending characters arising in (I) must be dealt with separately, by ad hoc methods. Only one yields a 2-transitive representation; this arises because of the isomorphism $G_2(2)' \cong PSU(3,3)$.

2. The bulk of the 2-transitive case was handled using a very short, simple argument, based upon the difficult results in (I). It is this simplicity which suggested the possibility of attacking rank 3 permutation representations.

Rank 3 representations. Suppose that G acts as a primitive rank 3 group on the cosets of K. Set $\varphi = 1_K^G$. Then $\varphi = 1 + \chi + \zeta$ for irreducible characters χ and ζ.

If $\chi, \zeta \notin 1_B^G$, then Seitz's theorem again applies.

If $\chi, \zeta \in 1_B^G$, then Tits' lemma applies, if we ignore a few possible groups G.

This leaves the possibility $\chi \in 1_B^G$ but $\zeta \notin 1_B^G$. In [19], all subgroups K of classical groups are determined for which $1_K^G = 1 + \chi + \zeta$ with χ irreducible, $\chi \in 1_B^G$, and $(\zeta, 1_B^G) = 0$; that is, the irreducibility of ζ can be dispensed with. For example:

Theorem. Let $G = SL(n,q)$ with $n \geq 8$ and $q > 2$, and let $K < G$. Suppose that $1_K^G = 1 + \chi + \zeta$ with χ irreducible, $\chi \in 1_B^G$ and $(\zeta, 1_B^G) = 0$. Then K fixes a subspace of dimension 1 or $n-1$.

Partial Proof. Recall that θ_i denotes the permutation character of G on i-spaces. Let θ_{ij} denote the permutation character on the pairs (V_i, V_j) of subspaces with $\dim V_i = i$, $\dim V_j = j$ and $V_i \subset V_j$ (of course, $i < j$ here).

Step 1. If $\chi \notin \theta_3$ then $(1_K^G, \theta_3) = 1$, so K is transitive on 3-spaces. By one of Perin's results (2.C), $K = G$.

Thus, in the notation of Example 1, χ must be χ_1, χ_2 or χ_3.

Step 2. Suppose that $\chi = \chi_1$.

Compute $(\theta_1, \varphi) = (\theta_1, \theta_1) = 2$

$(\theta_4, \varphi) = (\theta_4, \theta_1) = 2$

$(\theta_{14}, \varphi) = (\theta_{14}, \theta_1) = 3$.

(The stabilizer of the pair (V_1,V_4) has 3 point-orbits: $\{V_1\}$, the remaining points of V_4, and the points outside V_4.) There are thus 2 orbits each of points and hyperplanes, and 3 orbits of pairs (V_1,V_4).

By the Pigeon-Hole Principle, there is a point x such that K_x is transitive on the 4-spaces through x. Then K_x is transitive on the 3-spaces of V/x. By Perin's results, K_x induces at least $SL(n-1,q)$ on V/x.

If K is irreducible, it is now easy to show that $K=G$. Thus, K is reducible, and fixes either a point or a hyperplane, as required.

Step 3. Suppose that $\chi=\chi_3$.

This time $(\theta_{12},\varphi)=1+(\theta_{12},\theta_3-\theta_2)=1+4-4$, so K is transitive on the pairs (V_1,V_2). Also, $(\theta_3,\varphi)=1+(\theta_3,\chi_3)=2$, while $(\theta_{34},\varphi)=1+(\theta_{34},\theta_3-\theta_2)=1+7-5=3$. By the Pigeon-Hole Principle, there is thus a 3-space E such that K_E is transitive on the $(q^{n-3}-1)/(q-1)$ 4-spaces containing E. Now Perin's Method (2.C) can be applied, and produces the contradiction $K=G$.

Step 4. If $\chi=\chi_2$, the argument is similar, but somewhat more involved; cf. [19].

Remarks. 1. The omitted cases ($n \leq 7$ or $q = 2$) are handled similarly. However, examples arise when $n = 4$: K can contain $Sp(4,q)$, $SL(2,q^2)$ or (if $q = 2$) A_6 as a normal subgroup.

2. The determination of the rank 3 permutation representations of the classical groups follows the same pattern, but is much more involved. In particular, much more information is required concerning 1_G^B: there is no longer a nice nesting of characters as in Example 1.

Part 3. The root group geometry of $E_8(q)$

The prerequisites for this part are Carter [7, §§3.4, 3.6, 5.2, 6.3, 8.5, 12.1].

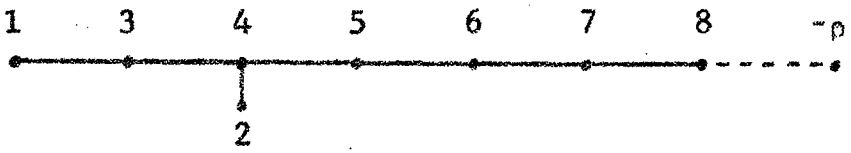
(3.A) The root system.

Notation. Fundamental base $\alpha_1, \ldots, \alpha_8$.

Fundamental reflections $s_1, \ldots, s_8$.

Height of $\Sigma a_i \alpha_i$ is Σa_i.

Highest root: ρ. This is related to the root system as in the following extended Dynkin diagram.



For all roots α, β, we have $\alpha \cdot \alpha = 2$, and $\alpha \cdot \beta = 0$ or ± 1, when $\beta \neq \pm \alpha$.

$\alpha_i \cdot \alpha_j = 0$ or -1.

Since s_i is the reflection in $\alpha_i^{\perp}$,

$v^{s_i} = v - 2\frac{v \cdot \alpha_i}{\alpha_i \cdot \alpha_i}\alpha_i = v - (v \cdot \alpha_i)\alpha_i$. Thus,

$$\alpha_j^{s_i} = \begin{cases} -\alpha_i & \text{if } j = i \\ \alpha_j & \text{if } \alpha_i \cdot \alpha_j = 0 \\ \alpha_j + \alpha_i & \text{if } \alpha_i \cdot \alpha_j = -1 \end{cases}$$

Computations. Write $\Sigma a_i\alpha_i = {}^{a_1a_3a_4a_5a_6a_7a_8}_{\quad a_2}$. When applying s_i to this, only a_i is changed; namely, to

$-a_i$ + (sum of a_j with i and j "adjacent" in the Dynkin diagram).

Starting from α_1, the entire root system can be computed; fortunately, however, lists in Bourbaki [5, pp. 268-270] and Aschbacher-Seitz [3, pp. 5-8] are available in order to frequently save time.

Examples.

$$0\ 1\ \underset{1}{1}\ 1\ 0\ 0\ 0 \xrightarrow{s_1} 1\ 1\ \underset{1}{1}\ 1\ 0\ 0\ 0 \xrightarrow{s_4} 1\ 1\ \underset{1}{2}\ 1\ 0\ 0\ 0$$

$$\rho = 2\ 4\ \underset{3}{6}\ 5\ 4\ 3\ 2 \xrightarrow{s_8} 2\ 4\ \underset{3}{6}\ 5\ 4\ 3\ 1 \xrightarrow{s_7} 2\ 4\ \underset{3}{6}\ 5\ 4\ 2\ 1$$

The coordinates of ρ are taken from [5]. Note that $\rho^{s_8} = \rho - \alpha_8$ is the next-to-the-highest root.

(3.B) Commutator relations.

Each root α yields a root group $X_\alpha \cong GF(q)^+$. The Chevalley commutator relations assert that, if $\beta \neq \pm\alpha$, then

$$[X_\alpha, X_\beta] = \begin{cases} 1 & \text{if } \alpha+\beta \text{ is not a root} \\ X_{\alpha+\beta} & \text{if } \alpha+\beta \text{ is a root.} \end{cases}$$

Moreover, $\langle X_\alpha, X_{-\alpha}\rangle \cong SL(2,q)$.

$G = E_8(q)$ is generated by all these X_α.

$U = \langle X_\alpha \mid \alpha > 0\rangle$ is a Sylow subgroup of G.

$|U| = q^{120}$ since there are 120 positive roots.

$B = U \rtimes H$ with $H \cong (GF(q)^*)^8$.

$X_\rho \leq Z(U)$; for, since ρ is the highest root, $\rho + \alpha$ cannot be a root for any root $\alpha > 0$.

We will say that α involves α_j (or s_j) if $\alpha = \sum_1^8 a_i\alpha_i$ with $a_j \neq 0$.

More notation. If $I \subseteq \{1,\ldots,8\}$, then $s_I = \{s_i \mid i \in I\}$, and

$W_I = \langle s_j \mid s_j \in s_I\rangle$

$P_I = BW_IB$

$Q_I = \langle X_\alpha \mid \alpha > 0$ and α involves some member of $s_I\rangle$

$L_I = \langle X_{\alpha_j}, X_{-\alpha_j} \mid j \notin I\rangle$.

We will write $L_{ij}\ldots$ in place of $L_{\{i,j,\ldots\}}$, and so on.

Levi decomposition. $P_I = Q_I \rtimes L_IH$. The structure of L_I is determined by deleting the nodes in I from the Dynkin diagram.

Double cosets. The number of W_I, W_J double cosets in W equals the number of P_I, P_J double cosets in G,

for all $I,J \subseteq \{1,\ldots,8\}$ (Bourbaki [5, p. 83]). Thus, the number of orbits of W_I on the set of W_J-cosets equals the number of orbits of P_I on the set of P_J-cosets.

Example. $Q_8 = \langle X_\rho, X_\alpha \mid \alpha$ has α_8-coefficient $1\rangle$, while L_8 has type $E_7(q)$. For the following discussion, a list of roots α is helpful [5; 3].

$|Q| = q^{1+56}$. If α has α_8-coefficient 1 then $\rho - \alpha$ is also a root and $\langle X_\alpha, X_{\rho-\alpha}\rangle$ is special of order q^3 with center X_ρ. If β has α_8-coefficient 1, but $\beta \neq \rho - \alpha$, then $\alpha + \beta$ is not a root. Thus, Q_8 is special of order q^{1+56}, with $Z(Q) = Q' = X_\rho$.

This produces a 56-dimensional module for L_8H over GF(q). (Scalars are obtained from a cyclic subgroup H_0 of H which centralizes L_8 and acts fixed-point-freely on Q_8/X_ρ.)

L_8H preserves the alternating bilinear form on Q_8/X_ρ defined by $(uX_\rho, vX_\rho) = [u,v]$, where X_ρ is identified with GF(q).

Remark. Here, and elsewhere in Part 3, the finiteness of the field is irrelevant: analogous results exist for infinite fields.

(3.C) Root groups

Definition. $\Omega = \{X_\rho{}^g \mid g \in G\}$ is the set of root groups of G. This will be used as the set of points of a geometry.

First Suborbit Lemma. G has rank 5 on Ω. Representatives of orbits of G on $\Omega \times \Omega$ and the groups they generate are as in the following table.

Pair X_α, X_β	$\langle X_\alpha, X_\beta \rangle$	Remark
X_ρ, X_ρ	X_ρ	$\alpha = \beta$
$X_\rho, X_{\rho - \alpha_8}$	$X_\rho \times X_{\rho - \alpha_8}$	$\alpha - \beta$ = root
X_ρ, X_{α_1}	$X_\rho \times X_{\alpha_1}$	$\alpha + \beta, \alpha - \beta$ not roots
$X_\rho, X_{-\rho + \alpha_8}$	Sylow in SL(3,q)	$\alpha + \beta$ = root
$X_\rho, X_{-\rho}$	SL(2,q)	$\alpha = -\beta$

Definitions. If $x, y \in \Omega$ and $\langle x, y \rangle$ is conjugate to $\langle X_\rho, X_{\rho - \alpha_8} \rangle$, we will write $y \in \Delta(x)$. If $\langle x, y \rangle \cong SL(2,q)$, x and y are called opposite.

Proof. $s_1, \ldots, s_7$ fix ρ, so $N_G(X_\rho) \geq BW_8B = P_8$. Since P_8 is a maximal parabolic subgroup, we have $N_G(X_\rho) = P_8$. Thus, G acts on Ω precisely as it does on the cosets of P_8. Since the number of

double cosets in W equals the number of P_8, P_8 double cosets in G, it suffices to consider W and W_8.

Thus, consider the action of W_8 on roots. Its orbits are as follows:

ρ

$\{\alpha \mid \alpha_8\text{-coefficient is } 1\}$

$\{\alpha \mid \alpha_8\text{-coefficient is } 0\}$

$\{\alpha \mid \alpha_8\text{-coefficient is } -1\}$

$-\rho$.

The third set is an orbit because the Weyl group of type E_7 is transitive on the roots for E_7. The second set arose at the end of (3.B); that it is an orbit can be readily proved by computing as in (3.A).

This proves the lemma, except for the generation part. But that follows from the commutator relations, the standard fact that $\langle X_\alpha, X_{-\alpha} \rangle \cong SL(2,q)$, and the fact that ρ and α_8 determine an A_2 subsystem.

Remark. The Weyl group $W(E_8) \cong 2 \cdot O^+(8,2)$ (a non-split central extension of $O^+(8,2)$ by Z_2). It acts on the 120 pairs $\{\alpha, -\alpha\}$ as a rank 3 group, just as on the $(2^8 - 1) - (2^4 - 1)(2^3 + 1) = 120$ nonsingular vectors

of an $O^+(8,2)$-space. The stabilizer of ρ is $W(E_7) = \Omega(7,2) \times Z_2$ (compare (I.E)); the stabilizer of $\{\rho, -\rho\}$ is $W(E_7) \times Z_2$.

Lemma. $N_G(\langle X_\rho, X_{\rho-\alpha_8}\rangle) = P_7$. Moreover, the group $L = \langle X_\rho, X_{\rho-\alpha_8}\rangle$ contains $q+1$ root groups, each nontrivial element of L lying in exactly one of these root groups.

Proof. Since $\rho^{s_8} = \rho - \alpha_8$, $W_7 = \langle s_1, \ldots, s_6, s_8\rangle$ normalizes L. Thus, $N_G(L)$ contains P_7, and hence equals P_7. Also, L is normal in the special group $\langle X_{\alpha_8}, X_{\rho-\alpha_8}\rangle$. Thus, $X_{\rho-\alpha_8}$ has q conjugates under X_{α_8}. Together with X_ρ, these provide us with the desired $q+1$ root groups. (Note that the last part of this argument simply takes place in an $SL(3,q)$.)

Definition. Point: element of Ω.

Line: conjugate of $\langle X_\rho, X_{\rho-\alpha_8}\rangle$ in G.

We will identify lines with their $q+1$ points. Points will be denoted $x, y, \ldots$, and lines by $L, M, \ldots$. The stabilizer of the point x is $G_x = N_G(x)$, and so on.

Second Suborbit Lemma. $P_8 = N_G(X_\rho)$ induces a rank 4 group on the set of lines through X_ρ. The stabilizer

of the line $\langle X_\rho, X_{\rho-\alpha_8}\rangle$ has the following orbit representatives on the lines through X_ρ.

$\langle X_\rho, X_{\rho-\alpha_8}\rangle$

$\langle X_\rho, X_{\rho-\alpha_7-\alpha_8}\rangle$ where $X_{\rho-\alpha_8}$ and $X_{\rho-\alpha_7-\alpha_8}$ are collinear

$\langle X_\rho, X_\tau\rangle$ where $\tau = \overset{2}{2}\,3\,4\,3\,2\,1\,1$, so $X_{\rho-\alpha_8}$ and X_τ commute but are not collinear

$\langle X_\rho, X_{\alpha_8}\rangle$ where $[X_{\rho-\alpha_8}, X_{\alpha_8}] = X_\rho$.

Proof. Once again, we reduce to W_{78} and $W_8 = W(E_7)$. Clearly, W_{78} leaves invariant the following sets of roots:

$\{\alpha > 0 \mid \alpha_8\text{-coefficient } 1,\ \alpha_7\text{-coefficient } 3\} = \{\rho - \alpha_8\}$ 1 root

$\{\alpha > 0 \mid \alpha_8\text{-coefficient } 1,\ \alpha_7\text{-coefficient } 2\}$ 27 roots

$\{\alpha > 0 \mid \alpha_8\text{-coefficient } 1,\ \alpha_7\text{-coefficient } 1\}$ 27 roots

$\{\alpha > 0 \mid \alpha_8\text{-coefficient } 1,\ \alpha_7\text{-coefficient } 0\} = \{\alpha_8\}$ 1 root

Each of these 56 roots yields an X_α collinear with X_ρ (that is, $\rho - \alpha$ is a root). (See [5; 3] for lists of roots, which show that these are all the roots X_α with $\rho - \alpha$ a root, and that the stated numbers 1,27,27,1 are correct.) Each of these sets is found to be a W_{78}-orbit.

Remarks. 1. W_8 is 2-transitive on the pairs $\{\alpha, \rho - \alpha\}$ with α as above. This corresponds to the 2-transitive representation of $Sp(6,2)$ on the 28 conjugates of $O^-(6,2)$ (cf. (2.F)).

2. The roots just considered also occurred in our discussion of Q_8 in (3.B). In fact, it is clear that the lines through X_ρ generate the group Q_8.

3. By a unpublished lemma of Borel and Tits [10,(2.4)], transitivity assertions such as the ones we have been considering can be proved effortlessly, without computation.

(3.D) Root group geometry.

We now present some properties of the geometry of points and lines just introduced. These are special cases of work of Tits, Stensholt and Cooperstein.

Recall that X_α and X_β are collinear iff $\alpha - \beta$ is 0 or a root.

(I) If $z = [x,y]$ is a point, then z is the unique point collinear with x and y.

Proof. If z' is a point collinear with x and y then we may assume that $z' = X_\rho$. Then $\langle x,z'\rangle, \langle y,z'\rangle \leq Q_8$, so $[x,y] = 1$ or X_ρ (since Q_8 is special). Consequently, $z = X_\rho = z'$. (Remark: $\langle x,z\rangle$ is a line since $\langle x,y\rangle$ is special; cf. (3.C).)

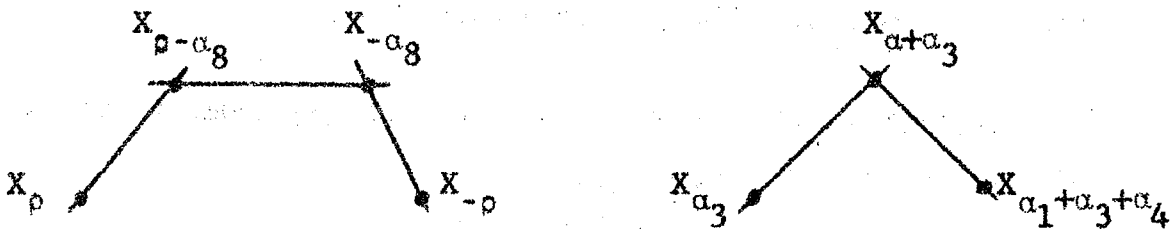
(II) If x and x' are opposite points, then $G_{xx'}$ acts on the lines through x exactly as G_x does.

Proof. We may assume that $x = X_\rho$, $x' = X_{-\rho}$. Then $G_x = Q_8 L_8 H$, where L_8 and H both fix $X_{-\rho}$. If L is any line through X_ρ then $[L, Q_8] \leq [Q_8, Q_8] = X_\rho$, so Q_8 fixes every line through X_ρ. This proves our assertion.

There is a natural graph on Ω: two points are adjacent iff they are collinear and distinct.

(III) This graph has diameter 3.

Proof. By the First Suborbit lemma, it suffices to consider the following pictures.

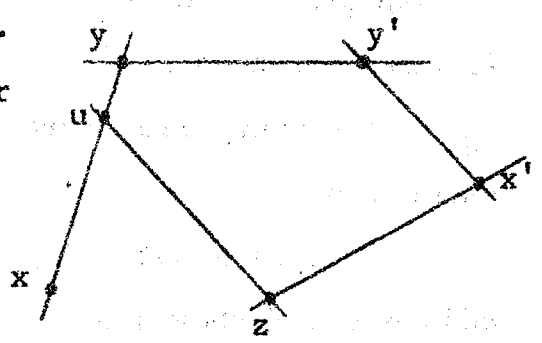


(Recall that X_α and X_β are collinear iff $\alpha - \beta$ is a root. Recall also that, if x and y are collinear with z then $[x,y] = 1$ or z, so that X_ρ and $X_{-\rho}$ cannot be distant <3 from one another.)

(IV) If x and x' are opposite points, and if $x \in L$, then there is a unique shortest path from x' ending inside L.

Proof. By (II), we may assume that $x = X_\rho$, $x' = X_{-\rho}$ and $L = \langle x, y \rangle$ with $y = X_{\rho-\alpha_8}$. The above picture then gives us a path x', $y' = X_{-\alpha_8}$, y. It remains to prove uniqueness.

Consider a path x', z, u with x' and z, resp. z and u, collinear, and $u \in L$. If $u = y$ then z is collinear with both x' and y, so $z = [x', y]$ is uniquely determined by (I).



So suppose that $u \neq y$. Since $\langle x, y' \rangle$ is isomorphic to a Sylow subgroup of $SL(3,q)$ (cf. (3.C)) we can find $g \in y'$ with $x = u^g$. But g certainly fixes x', so now z^g is collinear with both x and x', which we saw (in the proof of (III)) is impossible.

Remarks. Property (IV) is strongly reminiscent of of the B - S property (1.G).

The picture  now sets

up a bijection between the lines L through x and the lines L' through x'. We will show that this is, in some sense, an isomorphism (cf. (VIII)).

Definition. A subspace is a set of points which contains each line meeting it twice. An abelian subspace is a subspace in which any two points commute.

General example. If K is a subgroup of G generated by root groups, then $\Omega \cap K$ is a subspace. (For, if $x,y \in K$ and $\langle x,y \rangle$ is a line then $\langle x,y \rangle \leq K$.)

Example. Set $E = \langle X_{\rho}, X_{\rho-\alpha_7}, X_{\rho-\alpha_7-\alpha_8} \rangle$. From the extended Dynkin diagram we see that E lies in the subgroup of type $SL(4,q)$ generated by the groups $X_{\pm\alpha}$ with $\alpha = \rho$, $\rho - \alpha_7$ and $\rho-\alpha_7-\alpha_8$. It follows that E is elementary abelian of order q^3, and contains q^2+q+1 root groups (which are transvection groups for the $SL(4,q)$). There are also q^2+q+1 lines inside E, and hence $E \cap \Omega$ has the natural structure of a projective plane.

The usual proof shows that $N_G(E) = P_6$. Moreover, $\langle X_{\pm\alpha_7}, X_{\pm\alpha_8} \rangle \cong SL(3,q)$, and acts on E in the natural manner.

Definition. A plane is any subspace of the form $(E \cap \Omega)^g$, $g \in G$. Of course, we will identify $(E \cap \Omega)^g$ with E^g.

(V) Three pairwise collinear points are coplanar.

Proof. This follows immediately from the Second Suborbit Lemma.

The entire building of $E_8(q)$ can be seen in terms of abelian subspaces. First, note that, if P_I is any parabolic subgroup, then $\langle Z(Q_I)^{P_I}\rangle \cap \Omega$ is an abelian subspace.

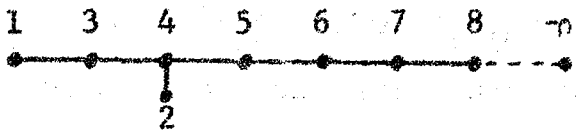
Examples. Each maximal parabolic subgroup P_i occurs as $N_G(S_i)$ for a suitable abelian subspace S_i on which P_i acts "nicely". Specifically, the following table lists each P_i, the desired S_i, and the group induced by the Levi factor L_i and S_i.

P_i	S_i		$\lvert S_i\rvert$	Induced on S_i
P_8	X_ρ	(point)	q	1
P_7	$\langle X_\rho, X_{\rho-\alpha_8}\rangle$	(line)	q^2	$SL(2,q)$
P_6	$\langle X_\rho, X_{\rho-\alpha_8}, X_{\rho-\alpha_7-\alpha_8}\rangle$	(plane)	q^3	$SL(3,q)$
P_5	$\langle X_\rho, X_{\rho-\alpha_8}, X_{\rho-\alpha_7-\alpha_8}, X_{\rho-\alpha_6-\alpha_7-\alpha_8}\rangle$		q^4	$SL(4,q)$
P_4	$\langle X_\rho, X_{\rho-\alpha_8}, \ldots, X_{\rho-\alpha_5-\alpha_6-\alpha_7-\alpha_8}\rangle$		q^5	$SL(5,q)$
P_3	$\langle X_\alpha \mid \alpha$ involves $4\alpha_3\rangle$		q^7	$SL(7,q)$
P_2	$\langle X_\alpha \mid \alpha$ involves $3\alpha_2\rangle$		q^8	$SL(8,q)$
P_1	$\langle X_\alpha \mid \alpha$ involves $2\alpha_2\rangle$		q^{14}	$\Omega^+(14,q)$

This table is readily checked using the roots listed in Bourbaki [5] or Aschbacher-Seitz [3]; we will return to the P_1 case soon.

Example. If $D = \langle X_{\pm\alpha_2}, X_{\pm\alpha_3}, \ldots, X_{\pm\alpha_8}, X_{\pm\rho} \rangle$, then $D = \langle X_{\pm\epsilon_i \pm \epsilon_j} \mid i \neq j \rangle$ and D has type $D_8(q)$.

Proof. The equality follows by forming various sums $\alpha_k + \alpha_{k+1} + \ldots$ in order to obtain the roots $\epsilon_i \pm \epsilon_j$. The structure of D is proven by looking at the extended Dynkin diagram



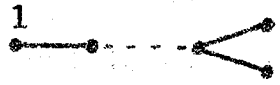
and using generators and relations (Carter [7, 12.1]).

Digression concerning orthogonal groups. Let V be an orthogonal vector space of type $\Omega^+(2m,q)$, $m \geq 2$, and let x be a singular point. Call $K = \Omega^+(2m,q)$ and $Q = O_p(K_x)$. Then Q is abelian and $K_x = Q \rtimes L$ with $L \cong \Omega^+(2m-2,q)$, where L centralizes x and acts the same on $x^\perp/x$ and Q.

Proof. Use a basis $e_1, \ldots, e_m, f_1, \ldots, f_m$ as in (1.E), where $x = \langle e_m \rangle$. Then Q consists of all transformations g_u, $u \in \langle e_m, f_m \rangle^\perp$, where $u = \sum_{i=1}^{m-1} \gamma_i e_i + \sum_{i=1}^{m-1} \delta_i f_i$ and

$$g^u : \begin{cases} e_m \to e_m \\ f_m \to f_m + u \\ e_i \to e_i - \beta_i e_m \\ f_i \to f_i - \gamma_i e_m \end{cases}$$

for $i < m$. Call $L = C_K(\langle e_m, f_m \rangle)$. If $\ell \in L$, compute that $g_u{}^{\ell} = g_{u^{\ell}}$.

Remark. K_x is a parabolic subgroup (see (1.D)). If  is the Dynkin diagram of K, then $K_x = P_1 = Q_1 L_1 H$ with $Q_1 = Q$ and $L_1 = L$.

We now return to the geometry of $E_8(q)$.

(VI) Let $x, y \in \Omega$, with x and y not collinear but $\langle x,y \rangle$ abelian. Define

$$\Delta(x,y) = \Delta(x) \cap \Delta(y)$$

$$\Sigma(x,y) = \Omega \cap \langle x, y, \Delta(x,y) \rangle.$$

Then $\Sigma(x,y)$ is an abelian subspace. It can be identified with the set of points of an $O^+(14,q)$ space $\langle \Sigma(x,y) \rangle$. The group induced on it by its stabilizer has a normal subgroup $\Omega^+(14,q)$, acting on it in the natural manner.

Proof. By the First Suborbit Lemma, we may assume that $x = X_\rho$ and $y = X_\sigma$, where $\sigma = 2\,3\,4\,\underset{2}{3}\,2\,1\,0$. Then $\langle \Sigma(x,y) \rangle$ contains $Y = \langle X_\sigma | \alpha \text{ involves } 2\alpha_1 \rangle = \langle X_{\varepsilon_8 \pm \varepsilon_i} | i \neq 8 \rangle$. As usual, since $N_G(Y) \geq P_1$ we have $N_G(Y) = P_1$.

Similarly, $G_{xy} \geq P_{18}$, and it is easy to check that $G_{xy} \neq P_1, P_8$. Thus, $G_{xy} = P_{18}$. But $P_{18} \cong \langle X_\alpha | \alpha$ involves $2\alpha_1$ and $1\alpha_8 \rangle = \langle \Delta(x,y) \rangle$. Thus, $\langle \Sigma(x,y) \rangle = Y$.

Now note that $\Sigma(x,y)$ is contained in the group D defined before the digression. Thus, the digression completes the proof.

Remark. Thus, any two distinct commuting points x,y are either contained in a unique line $\langle x,y \rangle \cap \Omega$ or a unique "hyperline" $\Sigma(x,y)$. The uniqueness is further clarified in the next results.

(VII) If x_1 and y_1 are distinct points of $\Sigma(x,y)$, then x_1 and y_1 are collinear in the subspace $\Sigma(x,y)$ if and only if they are collinear t.s. points of the orthogonal space $\langle \Sigma(x,y) \rangle$. If x_1 and y_1 are not collinear, then $\Sigma(x,y) = \Sigma(x_1,y_1)$.

Proof. The first statement follows from the digression and the embedding in D. The second is a consequence of the transitivity of $\Omega^+(14,q)$ on pairs of non-perpendicular points.

(VIII) Given four points x,x',y,y' with x opposite x', y opposite y', and the groups $\langle x,y \rangle$ and $\langle x,y' \rangle$ commuting, there is a unique subspace $D(x,x',y,y')$ of type $D_8(q)$ containing all four of them.

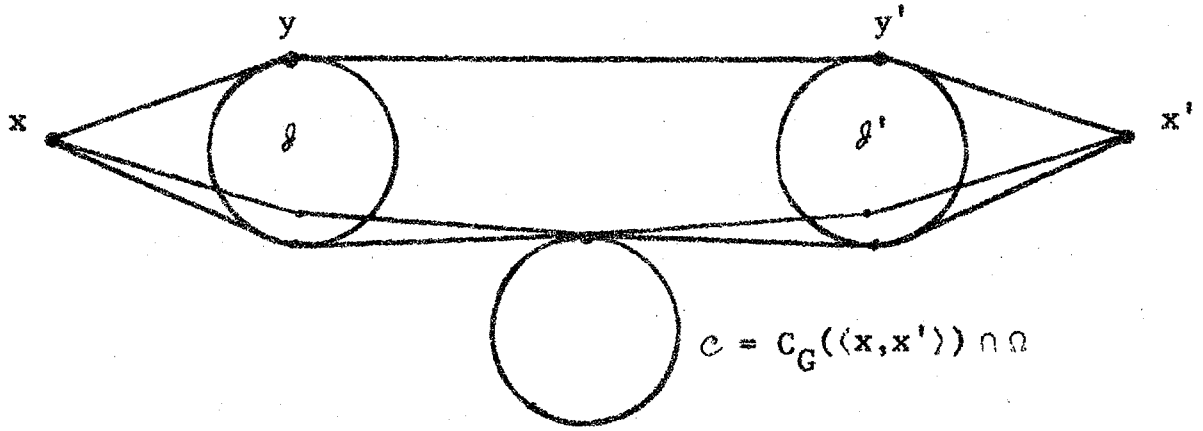
Proof. First we may assume that $x = X_\rho$ and $x' = X_{-\rho}$, and then that $y = X_\sigma$ and $y' = X_{-\sigma}$.

For existence, use the group $D = \langle X_{\pm\alpha_2}, \ldots, X_{\pm\alpha_8}, X_{\pm\rho} \rangle$.

Now assume that $X_{\pm\rho}$ and $X_{\pm\sigma}$ are in a subspace $D_8(q)$. Consider $D_8(q) \cap \Delta(X_\rho, X_\sigma)$, which is contained in $\Delta(X_\rho, X_\sigma)$ and hence in D. But there must be equally many points in $D_8(q)$ collinear with X_ρ and X_σ as there are in D. Hence, $\Delta(X_\rho, X_\sigma) \subset D_8(q)$. It follows similarly that $\langle D_8(q) \rangle \geq \langle \Sigma(X_\rho, X_\sigma), \Sigma(X_{-\rho}, X_{-\sigma}) \rangle = D$, as required.

Remark. The same proof shows that each orthogonal subgroup of G generated by root groups is conjugate to a subgroup of D.

The view from $E_7(q)$.



In this diagram, every line on x produces exactly one point y and one point y', by (IV).

(VIII) Theorem. Suppose that x and x' are opposite points. Let $\mathcal{L}$ and $\mathcal{L}'$ be the set of points y resp. y' as in the diagram; set $\mathcal{C} = C_G(\langle x,x'\rangle) \cap \Omega$.

(i) $\mathcal{L}, \mathcal{L}'$ and $\mathcal{C}$ are subspaces.

(ii) $y \to y'$ is an isomorphism, sending lines of $\mathcal{L}$ to lines of $\mathcal{L}'$.

(iii) Each $\Sigma(x,w)$ contains a unique point w_1 of $\mathcal{C}$, and $\Delta(x,w_1) \subset \mathcal{L}$.

(iv) $\mathcal{C}$ is the set of root groups of a group of type $E_7(q)$.

Proof. We may assume that $x = X_\rho$ and $x' = X_{-\rho}$. Then $G_{xx'} = L_8 H$ with L_8 of type $E_7(q)$. As in (II), L_8 centralizes x and x'.

(iv) This is now clear, since X_σ is in L_8.

(i),(ii) Let y and z be distinct collinear points of $\mathcal{L}$. By the Second Suborbit Lemma, we may assume that $y = X_{\rho-\alpha_8}$ and $z = X_{\rho-\alpha_7-\alpha_8}$. Thus, $\langle y,z\rangle \cap \Omega$ is a line (compare (VI)). But X_{α_7} fixes x,x' and y, and acts transitively on the points $\neq y$ of $\langle y,z\rangle$. (Note that $\langle X_{\alpha_7}, z\rangle$ is isomorphic to a Sylow subgroup of $SL(3,q)$; cf. (3.C).) Since X_{α_7} stabilizes $\mathcal{L}$, it follows that $\langle y,z\rangle \cap \Omega \subset \mathcal{L}$.

Since $y' = X_{-\rho+\alpha_8}$ and $z' = X_{-\rho+\alpha_7+\alpha_8}$ with y' and z' collinear, (IV) completes the proof of (ii).

(iii) By (VI), $N_G(\Sigma(x,w)$ is transitive on the points of $\Sigma(x,w)$. It follows that G_x is transitive on the subspaces $\Sigma(x,w)$ containing x; by (II), so is $G_{xx'}$. We may then assume that $\Sigma(x,w) = \Sigma(X_\rho, X_\sigma)$, and then clearly $X_\sigma \in \Sigma(x,w) \cap \mathcal{C}$. Moreover, generators for $\langle \Delta(X_\rho, X_\sigma) \rangle$ were found in (VI), and all belong to $\mathcal{G}$. Hence, <u>each</u> $\Delta(x,w_1) \subset \mathcal{G}$ for $w_1 \in \mathcal{C}$.

It remains to prove the uniqueness part of (iii). Let $v \in \Sigma(X_\rho, X_\sigma) \cap \mathcal{C}$. As in the suborbit lemmas, we can use L_8 to move the pair (x_σ, v) to a pair (X_σ, X_α). Then also $X_\alpha \in \Sigma(X_\rho, X_\sigma) \cap \mathcal{C}$ since L_8 fixes x and x'. Of the 14 roots β yielding points X_β in $\Sigma(X_\rho, X_\sigma)$, only one of them (namely σ) produces an X_β collinear with X_ρ. Thus, $X_\alpha = X_\sigma$, and hence also $v = X_\sigma$.

<u>Remark</u>. The X_β's inside $\Sigma(X_\rho, X_\sigma)$ form a generalized octahedron (cross-polytope).

The isomorphism $y \to y'$ preserves more than just the collinearity of points, as the next property shows.

(IX) Let $x,x',\mathcal{L},\mathcal{L}',\mathcal{C}$, and $y \to y'$ be as in (VIII). Let $y,z \in \mathcal{L}$, $y \neq z$.

(i) If y and z are collinear, then so are y' and z', and $\langle y,z' \rangle$ is abelian but not a line.

(ii) If $\langle y,z \rangle$ is abelian but not a line, then the same is true of $\langle y',z' \rangle$, and $[y,z'] = [y',z]$ is a point of $\mathcal{C}$.

(iii) If $[y,z] = x$ then $[y',z'] = x'$, and $\langle x,x',y,y',z,z' \rangle \cong SL(3,q)$.

Proof. By (II), (IV) and the Second Suborbit Lemma, we may assume that $y = X_\alpha$ and $z = X_\beta$ with $\alpha = \rho - \alpha_8$ and $\beta = \rho - \alpha_7$, τ $(= {}^{2\ 3\ 4\ 3\ 2\ 1\ 1}_{\quad\ 2})$, or α_8. Then $y' = X_{-\alpha_8}$ and $z' = X_{-\rho+\beta}$.

(i) Here, $\beta = \rho - \alpha_7 - \alpha_8$ and $\alpha \pm (\rho - \beta)$ are not roots.

(ii) Here, $\beta = \tau$ and $\alpha + (-\rho + \beta) = -\beta + (-\rho + \alpha)$ is the root τ^{s_8}.

(iii) Here, $\beta = \alpha_8$, so $\alpha + \beta = \rho$ and $(-\rho + \alpha) + (-\rho + \beta) = -\rho$. Moreover, $\langle x,x',y,y',z,z' \rangle = \langle X_{\pm\alpha_8}, X_{\pm\rho} \rangle$ is isomorphic to $SL(3,q)$, as is seen from the extended Dynkin diagram.

Geometries of type $E_7(q)$, $F_4(q)$, ${}^2E_6(q)$.

For each of these groups, use (long) root groups as points. Then analogues of the two suborbit lemmas can be proved, exactly as before. These geometries arise as subspaces of the $E_8(q)$ geometry. From this fact, or proofs similar to the ones just given, versions of properties (I) - (IX) can be obtained.

REFERENCES

1. E. Artin, Geometric Algebra. Interscience, 1957.

2. M. Aschbacher, A characterization of Chevalley groups over fields of odd order. Annals of Math. 106 (1977) 353-398, 399-468.

3. M. Aschbacher and G. M. Seitz, Involutions in Chevalley groups over fields of even order. Nagoya Math J. 63 (1976), 1-91.

4. E. Bannai, Maximal subgroups of low rank of finite symmetric and alternating groups. J. Fac. Sci. Univ. Tokyo Sect. 1A Math. 18 (1971/2) 475-486.

5. N. Bourbaki, Groupes et Algèbres de Lie VI. Hermann 1968.

6. F. Buekenhout and E. E. Shult, On the foundations of polar geometry. Geom. Ded. 3(1974) 155-170.

7. R. W. Carter, Simple Groups of Lie Type. Wiley 1972.

8. B. N. Cooperstein, The geometry of groups of Lie type. Ph.d. Thesis, U. of Michigan, 1975.

9. B. N. Cooperstein, Minimal degree for a permutation representation of a classical group (to appear in Israel J. Math.).

10. C. W. Curtis, W. M. Kantor and G. M. Seitz, The 2-transitive permutation representations of the finite Chevalley groups. TAMS 218 (1976) 1-57.

11. L. E. Dickson, Linear Groups with an Exposition of the Galois Field Theory. Dover Reprint 1958.

12. L. E. Dickson, On the cyclotomic function, Amer. Math. Monthly 12 (1905), 86-89.

13. J. Dieudonné, La Géométrie des Groupes Classiques. Springer 1955.

14. B. Huppert, Endliche Gruppen, I. Springer 1967.

15. B. Huppert, Geometric Algebra. U. of Illinois at Chicago Circle 1968/69.

16. W. M. Kantor, Rank 3 characterizations of classical geometries. J. Algebra 36(1975) 309-313.

17. W. M. Kantor, Subgroups of classical groups generated by long root elements (to appear in TAMS).

18. W. M. Kantor, Permutation representations of the finite classical groups of small degree or rank (to appear in J. Algebra).

19. W. M. Kantor and R. A. Liebler, The rank 3 permutation representations of the classical groups (submitted).

20. J. McLaughlin, Some groups generated by transvections. Arch. Math. 18 (1967), 364-368.

21. J. McLaughlin, Some subgroups of $SL_n(F_2)$. III. J. Math. 13 (1969) 108-115.

22. W. H. Patton, The minimum index for subgroups in some classical groups: a generalization of a theorem of Galois. Ph.d. Thesis, U. of Illinois at Chicago Circle 1972.

23. D. Perin, On collineation groups of finite projective spaces. Math. Z. 126 (1972) 135-142.

24. F. C. Piper, On elations of finite projective spaces of odd order. JLMS 41 (1966) 641-648.

25. F. C. Piper, On elations of finite projective spaces of even order. JLMS 43 (1968) 456-464.

26. M. A. Ronan, Generalised hexagons. Ph.d. Thesis, U. of Oregon, 1978.

27. G. M. Seitz, Flag-transitive subgroups of Chevalley groups. Ann. of Math. 97 (1973) 27-56.

28. G. M. Seitz, Small rank permutation representations of finite Chevalley groups. J. Alg. 28 (1974) 508-517.

29. E. E. Shult, Groups, polar spaces and related structures. pp. 451-482 in Combinatorics (eds. M. Hall, Jr. and J. H. van Lint), Reidel 1975.

30. J. Tits, Buildings of Spherical Type and Finite BN-pairs. Springer Lecture Notes 386, 1974.

31. A. Wagner, Groups generated by elations. Abh. Hamburg 41 (1974) 199-205.

32. H. Wielandt, Finite Permutation Groups. Academic Press 1964.

33. A. Yanushka, Generalized hexagons of order (t,t). Israel J. Math. 23 (1976) 309-324.