

6 Orthogonal groups

We now turn to the orthogonal groups. These are more difficult, for two related reasons. First, it is not always true that the group of isometries with determinant 1 is equal to its derived group (and simple modulo scalars). Secondly, in characteristic different from 2, there are no transvections, and we have to use a different class of elements.

We let $O(Q)$ denote the isometry group of the non-degenerate quadratic form Q , and $SO(Q)$ the group of isometries with determinant 1. Further, $PO(Q)$ and $PSO(Q)$ are the quotients of these groups by the scalars they contain. We define $\Omega(Q)$ to be the derived subgroup of $O(Q)$, and $P\Omega(Q) = \Omega(Q)/(\Omega(Q) \cap Z)$, where Z consists of the scalar matrices. Sometimes $\Omega(Q) = SO(Q)$, and sometimes it is strictly smaller; but our notation serves for both cases.

In the case where F is finite, we have seen that for even n there is a unique type of non-degenerate quadratic form up to scalar multiplication, while if n is odd there are two types, having germ of dimension 0 or 2 respectively. We write $O^+(n, q)$, $O(n, q)$ and $O^-(n, q)$ for the isometry group of a non-degenerate quadratic form on $\text{GF}(q)^n$ with germ of rank 0, 1, 2 (and n even, odd, even respectively). We use similar notation for SO , $P\Omega$, and so on. Then we write $O^\varepsilon(n, q)$ to mean either $O^+(n, q)$ or $O^-(n, q)$. Note that, unfortunately, this convention (which is standard notation) makes ε the negative of the ε appearing in our general order formula (Theorem 3.17).

Now the order formula for the finite orthogonal groups reads as follows.

$$\begin{aligned}
 |O(2m+1, q)| &= d \prod_{i=1}^m (q^{2i} - 1) q^{2i-1} \\
 &= dq^{m^2} \prod_{i=1}^m (q^{2i} - 1), \\
 |O^+(2m, q)| &= \prod_{i=1}^m (q^i - 1)(q^{i-1} + 1) q^{2i-2} \\
 &= 2q^{m(m-1)} (q^m - 1) \prod_{i=1}^{m-1} (q^{2i} - 1), \\
 |O^-(2m, q)| &= 2(q+1) \prod_{i=1}^{m-1} (q^i - 1)(q^{i+1} + 1) q^{2i}
 \end{aligned}$$

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

where $d = (2, q - 1)$. Note that there is a single difference in sign between the final formulae for $O^\varepsilon(2m, q)$ for $\varepsilon = \pm 1$; we can combine the two and write

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

We have $|\mathrm{SO}(n, q)| = |\mathrm{O}(n, q)|/d$ (except possibly if n is odd and q is even). This is because, with this exclusion, the bilinear form B associated with Q is non-degenerate; and the orthogonal group consists of matrices P satisfying $P^\top AP = A$, where A is the matrix of the bilinear form, so that $\det(P) = \pm 1$. It is easy to show that, for q odd, there are orthogonal transformations with determinant -1 . The excluded case will be dealt with in Section 6.2. We see also that the only scalars in $\mathrm{O}(Q)$ are $\pm I$; and, in characteristic different from 2, we have $-I \in \mathrm{SO}(Q)$ if and only if the rank of the underlying vector space is even. Thus, for q odd, we have

$$|\mathrm{SO}(Q)| = |\mathrm{PO}(Q)| = |\mathrm{O}(Q)|/2,$$

and

$$|\mathrm{PSO}(Q)| = |\mathrm{SO}(Q)|/(n, 2).$$

For q and n even we have $\mathrm{O}(Q) = \mathrm{SO}(Q) = \mathrm{PO}(Q) = \mathrm{PSO}(Q)$.

Exercise 6.1 Let Q be a non-degenerate quadratic form over a field of characteristic different from 2. Show that $\mathrm{O}(Q)$ contains elements with determinant -1 . [Hint: if $Q(v) \neq 0$, then take the transformation which takes v to $-v$ and extend it by the identity on v^\perp (in other words, the reflection in the hyperplane v^\perp).]

6.1 Some small-dimensional cases

We begin by considering some small cases. Let V be a vector space of rank n carrying a quadratic form Q of Witt index r , where $\delta = n - 2r$ is the dimension of the germ of Q . Let $\mathrm{O}(Q)$ denote the isometry group of Q , and $\mathrm{SO}(Q)$ the subgroup of isometries of determinant 1.

Case $n = 1, r = 0$. In this case the quadratic form is a scalar multiple of x^2 . (Replacing q by a scalar multiple does not affect the isometry group.) Then $\mathrm{O}(Q) = \{\pm 1\}$, a group of order 1 or 2 according as the characteristic is or is not 2; and $\mathrm{SO}(Q)$ is the trivial group.

Case $n = 2, r = 1$. The quadratic form is $Q(x_1, x_2) = x_1x_2$, and the isometry group $G = O(Q)$ is

$$\left\{ \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}, \begin{pmatrix} 0 & \lambda \\ \lambda^{-1} & 0 \end{pmatrix} : \lambda \in F^\times \right\},$$

a group with a subgroup H of index 2 isomorphic to F^\times , and such that an element $t \in G \setminus H$ satisfies $t^2 = 1$ and $t^{-1}ht = h^{-1}$ for all $h \in H$. In other words, G is a *generalised dihedral group*. If $F = \text{GF}(q)$, then $O(2, q)$ is a dihedral group of order $2(q-1)$. Note that $H = \text{SO}(Q)$ if and only if the characteristic of F is not 2.

Case $n = 2, r = 0$. In this case, the quadratic form is

$$\alpha x_1^2 + \beta x_1x_2 + \gamma x_2^2,$$

where $q(x) = \alpha x^2 + \beta x + \gamma$ is an irreducible quadratic over F . Let K be a splitting field for q over F , and *assume* that K is a Galois extension (in other words, that q is separable over F : this includes the cases where either the characteristic is not 2 or the field F is finite). Then, up to scalar multiplication, the form Q is equivalent to the K/F norm on the F -vector space K . The orthogonal group is generated by the multiplicative group of elements of norm 1 in K and the Galois automorphism σ of K over F .

In the case $F = \text{GF}(q)$, this group is dihedral of order $2(q+1)$. In the case $F = \mathbb{R}$, the \mathbb{C}/\mathbb{R} norm is

$$z \mapsto z\bar{z} = |z|^2,$$

and so the orthogonal group is generated by multiplication by unit complex numbers and complex conjugation. In geometric terms, it is the group of rotations and reflections about the origin in the Euclidean plane.

Again we see that $\text{SO}(Q)$ has index 2 in $O(F)$ if the characteristic of F is not 2.

Exercise 6.2 Prove that, if K is a Galois extension of F , then the determinant of the F -linear map $x \mapsto \lambda x$ on K is equal to $N_{K/F}(\lambda)$. [Hint: if $\lambda \notin F$, the eigenvalues of this map are λ and λ^σ .]

Case $n = 3, r = 1$. In this case and the next, we describe a group preserving a quadratic form and claim without proof that it is the full orthogonal group. Also, in this case, we assume that the characteristic is not equal to 2.

Let $V = F^2$, and let W be the vector space of all quadratic forms on V (not necessarily non-degenerate). Then $\text{rk}(W) = 3$; a typical element of W is the quadratic form $ux^2 + vxy + wy^2$, where we have represented a typical vector in V as (x, y) . We use the triple (u, v, w) of coefficients to represent this vector of w . Now $\text{GL}(V)$ acts on W by substitution on the variables in the quadratic form. In other words, to the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}(2, F)$$

corresponds the map

$$\begin{aligned} ux^2 + vxy + wy^2 &\mapsto u(ax + cy)^2 + v(ax + cy)(bx + dy) + w(bx + dy)^2 \\ &= (ua^2 + vab + wb^2)x^2 + (2uac + v(ad + bc) + 2wbd)xy \\ &\quad + (uc^2 + vcd + wd^2)y^2, \end{aligned}$$

which is represented by the matrix

$$\rho(A) = \begin{pmatrix} a^2 & 2ac & c^2 \\ ab & ad + bc & cd \\ b^2 & 2bd & d^2 \end{pmatrix} \in \text{GL}(3, F).$$

We observe several things about this representation ρ of $\text{GL}(2, F)$:

- (a) The kernel of the representation is $\{\pm I\}$.
- (b) $\det(\rho(A)) = (\det(A))^3$.
- (c) The quadratic form $Q(u, v, w) = 4uw - v^2$ is multiplied by a factor $\det(A)^2$ by the action of $\rho(A)$.

Hence we find a subgroup of $O(Q)$ which is isomorphic to $\text{SL}^\pm(2, F)/\{\pm I\}$, where $\text{SL}^\pm(2, F)$ is the group of matrices with determinant ± 1 . Moreover, its intersection with $\text{SL}(3, F)$ is $\text{SL}(2, F)/\{\pm 1\}$. In fact, these are the full groups $O(Q)$ and $\text{SO}(Q)$ respectively.

We see that in this case,

$$\text{P}\Omega(Q) \cong \text{PSL}(2, F) \quad \text{if } |F| > 2,$$

and this group is simple if $|F| > 3$.

Case $n = 4, r = 2$. Our strategy is similar. We take the rank 4 vector space over F to be the space $M^{2 \times 2}(F)$, the space of 2×2 matrices over F (where F is any field). The determinant function on V is a quadratic form: $Q(X) = \det(X)$. Clearly X is the sum of two hyperbolic planes (for example, the diagonal and the antidiagonal matrices).

There is an action of the group $GL(2, F) \times GL(2, F)$ on X , by the rule

$$\rho(A, B) : X \mapsto A^{-1}XB.$$

We see that $\rho(A, B)$ preserves Q if and only if $\det(A) = \det(B)$, and $\rho(A, B)$ is the identity if and only if $A = B = \lambda I$ for some scalar λ . So we have a subgroup of $O(Q)$ with the structure

$$((SL(2, F) \times SL(2, F)) \cdot F^\times) / \{(\lambda I, \lambda I) : \lambda \in F^\times\}.$$

Moreover, the map $T : X \mapsto X^\top$ also preserves Q . It can be shown that together these elements generate $O(Q)$.

Exercise 6.3 Show that the above map T has determinant -1 on V , while $\rho(A, B)$ has determinant equal to $\det(A)^{-2} \det(B)^2$. Deduce (from the information given) that $SO(Q)$ has index 2 in $O(Q)$ if and only if the characteristic of F is not 2.

Exercise 6.4 Show that, in the above case, we have

$$P\Omega(Q) \cong PSL(2, F) \times PSL(2, F)$$

if $|F| > 3$.

Exercise 6.5 Use the order formulae for finite orthogonal groups to prove that the groups constructed on vector spaces of ranks 3 and 4 are the full orthogonal groups, as claimed.

6.2 Characteristic 2, odd rank

In the case where the bilinear form is degenerate, we show that the orthogonal group is isomorphic to a symplectic group.

Theorem 6.1 *Let F be a perfect field of characteristic 2. Let Q be a non-degenerate quadratic form in n variables over F , where n is odd. Then $O(Q) \cong Sp(n-1, F)$.*

Proof We know that the bilinear form B is alternating and has a rank 1 radical, spanned by a vector z , say. By multiplying Q by a scalar if necessary, we may assume that $Q(z) = 1$. Let \overline{G} be the group induced on V/Z , where $Z = \langle z \rangle$. Then \overline{G} preserves the symplectic form.

The kernel K of the homomorphism from G to \overline{G} fixes each coset of Z . Since

$$Q(v + az) = Q(v) + a^2,$$

and the map $a \mapsto a^2$ is a bijection of F , each coset of Z contains one vector with each possible value of Q . Thus $K = 1$, and $G \cong \overline{G}$.

Conversely, let \overline{g} be any linear transformation of V/Z which preserves the symplectic form induced by B . The above argument shows that there is a unique permutation g of V lifting the action of \overline{g} and preserving Q . Note that, since \overline{g} induces g on V/Z , it preserves B . We claim that g is linear. First, take any two vectors v, w . Then

$$\begin{aligned} Q(vg + wg) &= Q(vg) + Q(wg) + B(vg, wg) \\ &= Q(v) + Q(w) + B(v, w) \\ &= Q(v + w) \\ &= Q((v + w)g); \end{aligned}$$

and the linearity of \overline{g} shows that $vg + wg$ and $(v + w)g$ belong to the same coset of Z , and so they are equal. A similar argument applies for scalar multiplication. So $\overline{G} = \text{Sp}(n - 1, F)$, and the result is proved. ■

We conclude that, with the hypotheses of the theorem, $O(Q)$ is simple except for $n = 3$ or $n = 5$, $F = \text{GF}(2)$. Hence $O(Q)$ coincides with $\text{P}\Omega(Q)$ with these exceptions.

We conclude by constructing some more 2-transitive groups. Let F be a perfect field of characteristic 2, and B a symplectic form on F^{2m} . Then the set $Q(B)$ of all quadratic forms which polarise to B is a coset of the set of “square-semilinear maps” on V , those satisfying

$$\begin{aligned} L(x + y) &= L(x) + L(y), \\ L(cx) &= c^2L(x) \end{aligned}$$

(these maps are just the quadratic forms which polarise to the zero bilinear form).

In the finite case, where $F = \text{GF}(q)$ (q even), there are thus q^{2m} such quadratic forms, and they fall into two orbits under $\text{Sp}(2m, q)$, corresponding to the two

types of forms. The stabiliser of a form Q is the corresponding orthogonal group $O(Q)$. The number of forms of each type is the index of the corresponding orthogonal group in the symplectic group, which can be calculated to be $q^m(q^m + \varepsilon)/2$ for a form of type ε .

Now specialise further to $F = \text{GF}(2)$. In this case, “square-semilinear” maps are linear. So, given a quadratic form Q polarising to B , we have

$$Q(B) = \{Q + L : L \in V^*\}.$$

Further, each linear form can be written as $x \mapsto B(x, a)$ for some fixed $a \in V$. Thus, there is an $O(Q)$ -invariant bijection between $Q(B)$ and V . By Witt’s Lemma, $O(Q)$ has just three orbits on V , namely

$$\{0\}, \quad \{x \in V : Q(x) = 0, x \neq 0\}, \quad \{x \in V : Q(x) = 1\}.$$

So $O(Q)$ has just three orbits on $Q(B)$, namely

$$\{Q\}, \quad Q^\varepsilon(B) \setminus \{Q\}, \quad Q^{-\varepsilon}(B),$$

where Q has type ε and $Q^\varepsilon(B)$ is the set of all forms of type ε in $Q(B)$.

It follows that $\text{Sp}(2m, 2)$ acts 2-transitively on each of the two sets $Q^\varepsilon(B)$, with cardinalities $2^{m-1}(2^m + \varepsilon)$. The point stabiliser in these actions are $O^\varepsilon(2m, 2)$.

Exercise 6.6 What isomorphisms between symmetric and classical groups are illustrated by the above 2-transitive actions of $\text{Sp}(4, 2)$?

6.3 Transvections and root elements

We first investigate *orthogonal transvections*, those which preserve the non-degenerate quadratic form Q on the F -vector space V .

Proposition 6.2 *There are no orthogonal transvections over a field F whose characteristic is different from 2. If F has characteristic 2, then an orthogonal transvection for a quadratic form Q has the form*

$$x \mapsto x - Q(a)^{-1}B(x, a)a,$$

where $Q(a) \neq 0$ and B is obtained by polarising Q .

Proof Suppose that the transvection $x \mapsto x + (xf)a$ preserves the quadratic form Q , and let B be the associated bilinear form. Then

$$Q(x + (xf)a) = Q(x)$$

for all $x \in V$, whence

$$(xf)^2 Q(a) + (xf)B(x, a) = 0.$$

If $xf \neq 0$, we conclude that $(xf)Q(a) + B(x, a) = 0$. Since this linear equation holds on the complement of a hyperplane, it holds everywhere; that is, $B(x, a) = -(xf)Q(a)$ for all x .

If the characteristic is not 2, then $B(a, a) = 2Q(a)$. Substituting $x = a$ in the above equation, using $af = 0$, we see that $B(a, a) = 0$, so $Q(a) = 0$. But then $B(x, a) = 0$ for all x , contradicting the nondegeneracy of B in this case.

So we may assume that the characteristic is 2. If $B(x, a) = 0$ for all $x \in V$, then $Q(a) \neq 0$ by non-degeneracy. Otherwise, choosing x with $B(x, a) \neq 0$, we see that again $Q(a) \neq 0$. Then

$$xf = -Q(a)^{-1}B(x, a),$$

and the proof is complete. (Incidentally, the fact that f is non-zero now shows that a is not in the radical of B .)

Exercise 6.7 In the characteristic 2 case, replacing a by λa for $a \neq 0$ does not change the orthogonal transvection.

The fact that, if Q is a non-degenerate quadratic form in three variables with Witt rank 1 shows that we can find analogues of transvections acting on three-dimensional sections of V . These are called *root elements*, and they will be used in our simplicity proofs.

A *root element* is a transformation of the form

$$x \mapsto x + aB(x, v)u - aB(x, u)v - a^2Q(v)B(x, u)u$$

where $Q(u) = B(u, v) = 0$. The group of all such transformations for fixed u, v satisfying the above conditions, together with the identity, is called a *root subgroup* $X_{u, v}$.

Exercise 6.8 Prove that the root elements are isometries of Q , that they have determinant 1, and that the root subgroups are abelian. Show further that, if $Q(u) = 0$, then the group

$$X_u = \langle X_{u, v} : v \in u^\perp \rangle$$

is abelian, and is isomorphic to the additive group of $u^\perp / \langle u \rangle$.

Exercise 6.9 Write down the root subgroup $X_{u,v}$ for the quadratic form $Q(x_1, x_2, x_3) = x_1x_3 - x_2^2$ relative to the given basis $\{e_1, e_2, e_3\}$, where $u = e_1$ and $v = e_2$.

Now the details needed to apply Iwasawa's Lemma are similar to, but more complicated than, those that we have seen in the cases of the other classical groups. We summarise the important steps. Let Q be a quadratic form with Witt rank at least 2, and not of Witt rank 2 on a vector space of rank 4 (that is, not equivalent to $x_1x_2 + x_3x_4$). We also exclude the case where Q has Witt index 2 on a rank 5 vector space over $\text{GF}(2)$: in this case $\text{P}\Omega(Q) \cong \text{PSp}(4, 2) \cong S_6$.

- (a) The root subgroups are contained in $\Omega(Q)$, the derived group of $O(Q)$.
- (b) The abelian group X_u is normal in the stabiliser of u .
- (c) $\Omega(Q)$ is generated by the root subgroups.
- (d) $\Omega(Q)$ acts primitively on the set of flat 1-spaces.

Note that the exception of the case of rank 4 and Witt index 2 is really necessary for (d): the group $\Omega(Q)$ fixes the two families of rulings on the hyperbolic quadric shown in Figure 1 on p. 41, and each family is a system of blocks of imprimitivity for this group.

Then from Iwasawa's Lemma we conclude:

Theorem 6.3 *Let Q be a non-degenerate quadratic form with Witt rank at least 2, but not of Witt rank 2 on either a vector space of rank 4 or a vector space of rank 5 over $\text{GF}(2)$. Then $\text{P}\Omega(Q)$ is simple.*

It remains for us to discover the order of $\text{P}\Omega(Q)$ over a finite field. We give the result here, and defer the proof until later. The facts are as follows.

Proposition 6.4 (a) *Let Q have Witt index at least 2, and let F have characteristic different from 2. Then $\text{SO}(Q)/\Omega(Q) \cong F^\times / (F^\times)^2$.*

(b) *Let F be a perfect field of characteristic 2 and let Q have Witt index at least 2; exclude the case of a rank 4 vector space over $\text{GF}(2)$. Then $|\text{SO}(Q) : \Omega(Q)| = 2$.*

The proof of part (a) involves defining a homomorphism from $\text{SO}(Q)$ to $F^\times / (F^\times)^2$ called the *spinor norm*, and showing that it is onto and its kernel is $\Omega(Q)$ except in the excluded case.

In the remaining cases, we work over the finite field $\text{GF}(q)$, and write $O(n, q)$, understanding that if n is even then $O^\varepsilon(n, q)$ is meant.

Proposition 6.5 *Excluding the case q even and n odd:*

$$(a) \quad |\text{SO}(n, q) : \Omega(n, q)| = 2.$$

$$(b) \quad \text{For } q \text{ odd, } -I \in \Omega^\varepsilon(2m, q) \text{ if and only if } q^m \cong \varepsilon \pmod{4}.$$

The last part is proved by calculating the spinor norm of $-I$. Putting this together with the order formula for $\text{SO}(n, q)$ already noted, we obtain the following result:

Theorem 6.6 *For $m \geq 2$, excluding the case $\text{P}\Omega^+(4, 2)$, we have*

$$\begin{aligned} |\text{P}\Omega^\varepsilon(2m, q)| &= \left(q^{m(m-1)} (q^m - \varepsilon) \prod_{i=1}^{m-1} (q^{2i} - 1) \right) / (4, q^m - \varepsilon), \\ |\text{P}\Omega(2m+1, q)| &= \left(q^{m^2} \prod_{i=1}^m (q^{2i} - 1) \right) / (2, q - 1). \end{aligned}$$

Proof For q odd, have already shown that the order of $\text{SO}(n, q)$ is given by the expression in parentheses. We divide by 2 on passing to $\Omega(n, q)$, and another 2 on factoring out the scalars if and only if 4 divides $q^m - \varepsilon$. For q even, $|\text{SO}(n, q)|$ is twice the bracketed expression, and we lose the factor 2 on passing to $\Omega(n, q) = \text{P}\Omega(n, q)$.

Now we note that $|\text{P}\Omega(2m+1, q)| = |\text{P}\text{Sp}(2m, q)|$ for all m . In the case $m = 1$, these groups are isomorphic, since they are both isomorphic to $\text{PSL}(2, q)$. We have also seen that they are isomorphic if q is even. We will see later that they are also isomorphic if $m = 2$. However, they are non-isomorphic for $m \geq 3$ and q odd. This follows from the result of the following exercise.

Exercise 6.10 Let q be odd and $m \geq 2$.

- (a) The group $\text{P}\text{Sp}(2m, q)$ has $\lfloor m/2 \rfloor + 1$ conjugacy classes of elements of order 2.

(b) The group $P\Omega(2m+1, q)$ has m conjugacy classes of elements of order 2.

Hint: if $t \in \text{Sp}(2m, q)$ or $t \in \Omega(2m+1, q) = P\Omega(2m+1, q)$ satisfies $t^2 = 1$, then $V = V^+ \oplus V^-$, where $vt = \lambda v$ for $v \in V^\lambda$; and the subspaces V^+ and V^- are orthogonal. Show that there are m possibilities for the subspaces V^+ and V^- up to isometry; in the symplectic case, replacing t by $-t$ interchanges these two spaces but gives the same element of $\text{PSp}(2m, q)$. In the case $P\text{Sp}(2m, q)$, there is an additional conjugacy class arising from elements $t \in \text{Sp}(2m, q)$ with $t^2 = -1$.

It follows from the Classification of Finite Simple Groups that there are at most two non-isomorphic simple groups of any given order, and the only instances where there are two non-isomorphic groups are

$$P\text{Sp}(2m, q) \not\cong P\Omega(2m+1, q) \text{ for } m \geq 3, q \text{ odd}$$

and

$$PSL(3, 4) \not\cong PSL(4, 2) \cong A_8.$$

The lecture course will not contain detailed proofs of the simplicity of $P\Omega(n, q)$, but at least it is possible to see why $\text{SO}^+(2m, q)$ contains a subgroup of index 2 for q even. Recall from Chapter 3 that, for the quadratic form

$$x_1x_2 + \cdots + x_{2m-1}x_{2m}$$

of Witt index m in $2m$ variables, the flat m -spaces fall into two families \mathcal{F}^+ and \mathcal{F}^- , with the property that the intersection of two flat m -spaces has even codimension in each if they belong to the same family, and odd codimension otherwise. Any element of the orthogonal group must fix or interchange the two families. Now, for q even, $\text{SO}^+(2m, q)$ contains an element which interchanges the two families: for example, the transformation which interchanges the coordinates x_1 and x_2 and fixes all the others. So $\text{SO}^+(2m, q)$ has a subgroup of index 2 fixing the two families, which is $\Omega^+(2m, q)$. (In the case where q is odd, such a transformation has determinant -1 .)