4 Symplectic groups

In this and the next two sections, we begin the study of the groups preserving reflexive sesquilinear forms or quadratic forms. We begin with the symplectic groups, associated with non-degenerate alternating bilinear forms.

4.1 The Pfaffian

The determinant of a skew-symmetric matrix is a square. This can be seen in small cases by direct calculation:

$$\begin{aligned} \det\begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix} &= a_{12}^2, \\ \det\begin{pmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{pmatrix} &= (a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23})^2. \end{aligned}$$

Theorem 4.1 *(a) The determinant of a skew-symmetric matrix of odd size is zero.*

(b) There is a unique polynomial $\mathrm{Pf}(A)$ *in the indeterminates* a_{ij} *for* $1 \le i < j \le 2n$*, having the properties*

(i) if A *is a skew-symmetric* $2n \times 2n$ *matrix with* (i,j) *entry* a_{ij} *for* $1 \le i < j \le 2n$*, then*

$$\det(A) = \mathrm{Pf}(A)^2;$$

(ii) $\mathrm{Pf}(A)$ *contains the term* $a_{12}a_{34}\cdots a_{2n-1\,2n}$ *with coefficient* $+1$.

Proof We begin by observing that, if A is a skew-symmetric matrix, then the form B defined by

$$B(x,y) = xAy^\top$$

is an alternating bilinear form. Moreover, B is non-degenerate if and only if A is non-singular: for $xAy^\top = 0$ for all y if and only if $xA = 0$. We know that there is no non-degenerate alternating bilinear form on a space of odd dimension; so (a) is proved.

We know also that, if A is singular, then $\det(A) = 0$, whereas if A is non-singular, then there exists an invertible matrix P such that

$$PAP^\top = \operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right),$$

so that $\det(A) = \det(P)^{-2}$. Thus, $\det(A)$ is a square in either case.

Now regard a_{ij} as being indeterminates over the field F; that is, let $K = F(a_{ij} : 1 \le i < j \le 2n)$ be the field of fractions of the polynomial ring in $n(2n-1)$ variables over F. If A is the skew-symmetric matrix with entries a_{ij} for $1 \le i < j \le 2n$, then as we have seen, $\det(A)$ is a square in K. It is actually the square of a polynomial. (For the polynomial ring is a unique factorisation domain; if $\det(A) = (f/g)^2$, where f and g are polynomials with no common factor, then $\det(A)g^2 = f^2$, and so f^2 divides $\det(A)$; this implies that g is a unit.) Now $\det(A)$ contains a term

$$a_{12}^2 a_{34}^2 \cdots a_{2n-1\,2n}^2$$

corresponding to the permutation

$$(1\,2)(3\,4)\cdots(2n-1\,2n),$$

and so by choice of sign in the square root we may assume that (ii)(b) holds. Clearly the polynomial $\mathrm{Pf}(A)$ is uniquely determined.

The result for arbitrary skew-symmetric matrices is now obtained by specialisation (that is, substituting values from F for the indeterminates a_{ij}). ∎

Theorem 4.2 *If A is a skew-symmetric matrix and P any invertible matrix, then*

$$\mathrm{Pf}(PAP^\top) = \det(P) \cdot \mathrm{Pf}(A).$$

Proof We have $\det(PAP^\top) = \det(P)^2 \det(A)$, and taking the square root shows that $\mathrm{Pf}(PAP^\top) = \pm \det(P)\,\mathrm{Pf}(A)$; it is enough to justify the positive sign. For this, it suffices to consider the 'standard' skew-symmetric matrix

$$A = \operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right),$$

since all non-singular skew-symmetric matrices are equivalent. In this case, the $(2n-1, 2n)$ entry in $PAP^\top$ contains the term $p_{2n-1\,2n-1}p_{2n\,2n}$, so that $\mathrm{Pf}(PAP^\top)$ contains the diagonal entry of $\det(P)$ with sign $+1$. ∎

Exercise 4.1 A *one-factor* on the set $\{1,2,\ldots,2n\}$ is a partition F of this set into n subsets of size 2. We represent each 2-set$\{i,j\}$ by the ordered pair (i,j) with $i<j$. The *crossing number* $\chi(F)$ of the one-factor F is the number of pairs $\{(i,j),(k,l)\}$ of sets in F for which $i<k<j<l$.

(a) Let $\mathcal{F}_n$ be the set of one-factors on the set $\{1,2,\ldots,2n\}$. What is $|\mathcal{F}_n|$?

(b) Let $A=(a_{ij})$ be a skew-symmetric matrix of order $2n$. Prove that

$$\mathrm{Pf}(A)=\sum_{F\in\mathcal{F}_n}(-1)^{\chi(F)}\prod_{(i,j)\in F}a_{ij}.$$

4.2 The symplectic groups

The *symplectic group* $\mathrm{Sp}(2n,F)$ is the isometry group of a non-degenerate alternating bilinear form on a vector space of rank $2n$ over F. (We have seen that any two such forms are equivalent up to invertible linear transformation of the variables; so we have defined the symplectic group uniquely up to conjugacy in $\mathrm{GL}(2n,F)$.) Alternatively, it consists of the $2n\times 2n$ matrices P satisfying $P^\top AP=A$, where A is a fixed invertible skew-symmetric matrix. If necessary, we can take for definiteness either

$$A=\begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix}$$

or

$$A=\mathrm{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ldots,\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right).$$

The *projective symplectic group* $\mathrm{PSp}(2n,F)$ is the group induced on the set of points of $\mathrm{PG}(2n-1,F)$ by $\mathrm{Sp}(2n,F)$. It is isomorphic to the factor group $\mathrm{Sp}(2n,F)/(Sp(2n,F)\cap Z)$, where Z is the group of non-zero scalar matrices.

Proposition 4.3 *(a)* $\mathrm{Sp}(2n,F)$ *is a subgroup of* $\mathrm{SL}(2n,F)$.

(b) $\mathrm{PSp}(2n,F)\cong\mathrm{Sp}(2n,F)/\{\pm I\}$.

Proof (a) If $P\in\mathrm{Sp}(2n,F)$, then $\mathrm{Pf}(A)=\mathrm{Pf}(PAP^\top)=\det(P)\,\mathrm{Pf}(A)$, so $\det(P)=1$.

(b) If $(cI)A(cI)=A$, then $c^2=1$, so $c=\pm1$. ■

From Theorem 3.17, we have:

Proposition 4.4

$$|\operatorname{Sp}(2n,q)| = \prod_{i=1}^{n}(q^{2i}-1)q^{2i-1} = q^{n^2}\prod_{i=1}^{n}(q^{2i}-1). \quad \blacksquare$$

The next result shows that we get nothing new in the case $2n = 2$.

Proposition 4.5 $\operatorname{Sp}(2,F) \cong \operatorname{SL}(2,F)$ *and* $\operatorname{PSp}(2,F) \cong \operatorname{PSL}(2,F)$.

Proof We show that there is a non-degenerate bilinear form on F^2 preserved by $\operatorname{SL}(2,F)$. The form B is given by

$$B(x,y) = \det\begin{pmatrix} x \\ y \end{pmatrix}$$

for all $x,y \in F^2$, where $\begin{pmatrix} x \\ y \end{pmatrix}$ is the matrix with rows x and y. This is obviously a symplectic form. For any linear map $P : F^2 \to F^2$, we have

$$\begin{pmatrix} xP \\ yP \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} P,$$

whence

$$B(xP,yP) = \det\begin{pmatrix} xP \\ yP \end{pmatrix} = B(x,y)\det(P),$$

and so all elements of $\operatorname{SL}(2,F)$ preserve B, as required.

The second assertion follows on factoring out the group of non-zero scalar matrices of determinant 1, that is, $\{\pm I\}$. $\blacksquare$

In particular, $\operatorname{PSp}(2,F)$ is simple if and only if $|F| > 3$.

There is one further example of a non-simple symplectic group:

Proposition 4.6 $\operatorname{PSp}(4,2) \cong S_6$.

Proof Let $F = \operatorname{GF}(2)$ and $V = F^6$. On V define the "standard inner product"

$$x \cdot y = \sum_{i=1}^{6} x_i y_i$$

(evaluated in F). Let j denote the all-1 vector. Then

$$x \cdot x = x \cdot j$$

for all $x \in X$, so on the rank 5 subspace $j^\perp$, the inner product induces an alternating bilinear form. This form is degenerate — indeed, by definition, its radical contains j — but it induces a non-degenerate symplectic form B on the rank 4 space $j^\perp/\langle j\rangle$. Clearly any permutation of the six coordinates induces an isometry of B. So $S_6 \leq \mathrm{Sp}(4,2) = \mathrm{PSp}(4,2)$. Since

$$|S_6| = 6! = 15 \cdot 8 \cdot 3 \cdot 2 = |\,\mathrm{Sp}(4,2)|,$$

the result is proved. ■

4.3 Generation and simplicity

This subsection follows the pattern used for $\mathrm{PSL}(n,F)$. We show that $\mathrm{Sp}(2n,F)$ is generated by transvections, that it is equal to its derived group, and that $\mathrm{PSp}(2n,F)$ is simple, for $n \geq 2$, with the exception (noted above) of $\mathrm{PSp}(4,2)$.

Let B be a symplectic form. Which transvections preserve B? Consider the transvection $x \mapsto x + (xf)a$, where $a \in V$, $f \in V^*$, and $af = 0$. We have

$$B(x+(xf)a, y+(yf)a) = B(x,y) + (xf)B(a,y) - (yf)B(a,x).$$

So B is preserved if and only if $(xf)B(a,y) = (yf)B(a,x)$ for all $x,y \in V$. We claim that this entails $xf = \lambda B(a,x)$ for all x, for some scalar λ. For we can choose x with $B(a,x) \neq 0$, and define $\lambda = (xf)/B(a,x)$; then the above equation shows that $yf = \lambda B(a,y)$ for all y.

Thus, a *symplectic transvection* (one which preserves the symplectic form) can be written as

$$x \mapsto x + \lambda B(x,a)a$$

for a fixed vector $a \in V$. Note that its centre and axis correspond under the symplectic polarity; that is, its axis is $a^\perp = \{x : B(x,a) = 0\}$.

Lemma 4.7 *For* $r \geq 1$, *the group* $\mathrm{PSp}(2r,F)$ *acts primitively on the points of* $\mathrm{PG}(2r-1,F)$.

Proof For $r = 1$ we know that the action is 2-transitive, and so is certainly primitive. So suppose that $r \geq 2$.

Every point of $\mathrm{PG}(2r-1,F)$ is flat, so by Witt's Lemma, the symplectic group acts transitively. Moreover, any pair of distinct points spans either a flat subspace or a hyperbolic plane. Again, Witt's Lemma shows that the group is transitive on the pairs of each type. (In other words $G = \mathrm{PSp}(2r,F)$ has three orbits on ordered pairs of points, including the diagonal orbit

$$\Delta = \{(p,p) : p \in \mathrm{PG}(2r-1,F)\};$$

we say that $\mathrm{PSp}(2r,F)$ is a *rank* 3 *permutation group* on $\mathrm{PG}(2r-1,F)$.)

Now a non-trivial equivalence relation preserved by G would have to consist of the diagonal and one other orbit. So to finish the proof, we must show:

(a) if $B(x,y) = 0$, then there exists z such that $B(x,z), B(y,z) \neq 0$;

(b) if $B(x,y) \neq 0$, then there exists z such that $B(x,z) = B(y,z) \neq 0$.

This is a simple exercise. ■

Exercise 4.2 Prove (a) and (b) above.

Lemma 4.8 *For $r \geq 1$, the group $\mathrm{Sp}(2r,F)$ is generated by symplectic transvections.*

Proof The proof is by induction by r, the case $r = 1$ having been settled earlier (Theorem 2.6).

First we show that the group H generated by transvections is transitive on the non-zero vectors. Let $u, v \neq 0$. If $B(u,v) \neq 0$, then the symplectic transvection

$$x \mapsto x + \frac{B(x, v-u)}{B(u,v)}(v-u)$$

carries u to v. If $B(u,v) = 0$, choose w such that $B(u,w), B(v,w) \neq 0$ (by (a) of the preceding lemma) and map u to w to v in two steps.

Now it is enough to show that any symplectic transformation g fixing a non-zero vector u is a product of symplectic transvections. By induction, since the stabiliser of u is the symplectic group on $u^\top/\langle u\rangle$, we may assume that g acts trivially on this quotient; but then g is itself a symplectic transvection. ■

Lemma 4.9 *For $r \geq 3$, and for $r = 2$ and $F \neq \mathrm{GF}(2)$, the group $\mathrm{PSp}(2r,F)$ is equal to its derived group.*

Proof If $F \neq \mathrm{GF}(2), \mathrm{GF}(3)$, we know from Lemma 2.8 that any element inducing a transvection on a hyperbolic plane and the identity on the complement is a commutator, so the result follows. The same argument completes the proof provided that we can show that it holds for $\mathrm{PSp}(6,2)$ and $\mathrm{PSp}(4,3)$.

In order to handle these two groups, we first develop some notation which can be more generally applied. For convenience we re-order the rows and columns of the 'standard skew-symmetric matrix' so that it has the form

$$J = \begin{pmatrix} O & I \\ -I & O \end{pmatrix},$$

where O and I are the $r \times r$ zero and identity matrices. (In other words, the ith and $(i+r)$th basis vectors form a hyperbolic pair, for $i = 1, \ldots, r$.) Now a matrix C belongs to the symplectic group if and only if $C^\top JC = J$. In particular, we find that

(a) for all invertible $r \times r$ matrices A, we have

$$\begin{pmatrix} A^{-1} & O \\ O & A^\top \end{pmatrix} \in \mathrm{Sp}(2r, F);$$

(b) for all *symmetric* $r \times r$ matrices B, we have

$$\begin{pmatrix} I & B \\ O & I \end{pmatrix} \in \mathrm{Sp}(2r, F).$$

Now straightforward calculation shows that the commutator of the two matrices in (a) and (b) is equal to

$$\begin{pmatrix} I & B - ABA^\top \\ O & I \end{pmatrix},$$

and it suffices to choose A and B such that A is invertible, B is symmetric, and $B - ABA^\top$ has rank 1.

The following choices work:

(a) $r = 2$, $F = \mathrm{GF}(3)$, $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$;

(b) $r = 3$, $F = \mathrm{GF}(2)$, $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. ■

Theorem 4.10 *The group* $\mathrm{PSp}(2r,F)$ *is simple for* $r \geq 1$*, except for the cases* $(r,F) = (1,\mathrm{GF}(2))$*,* $(1,\mathrm{GF}(3))$*, and* $(2,\mathrm{GF}(2))$*.*

Proof We now have all the ingredients for Iwasawa's Lemma (Theorem 2.7), which immediately yields the conclusion. ■

As we have seen, the exceptions in the theorem are genuine.

Exercise 4.3 Show that $\mathrm{PSp}(4,3)$ is a finite simple group which has no 2-transitive action.

The only positive integers n such that $n(n-1)$ divides $|\mathrm{PSp}(4,3)|$ are $n = 2,3,4,5,6,9,10,16,81$. It suffices to show that the group has no 2-transitive action of any of these degrees. Most are straightforward but $n = 16$ and $n = 81$ require some effort.

(It is known that $\mathrm{PSp}(4,3)$ is the smallest non-abelian finite simple group with this property.)

4.4 A technical result

The result in this section will be needed at one point in our discussion of the unitary groups. It is a method of recognising the groups $\mathrm{PSp}(4,F)$ geometrically.

Consider the polar space associated with $\mathrm{PSp}(4,F)$. Its points are all the points of the projective space $\mathrm{PG}(3,F)$, and its lines are the flat lines (those on which the symplectic form vanishes). We call them F-lines for brevity. Note that the F-lines through a point p of the porojective space form the plane pencil consisting of all the lines through p in the plane $p^\perp$, while dually the F-lines in a plane Π are all those lines of Π containing the point $\Pi^\perp$. Now two points are prthogonal if and only if they line on an F-line.

The geometry of F-lines has the following property:

(a) Given an F-line L and a point p not on L, there is a unique point $q \in L$ such that pq is an F-line.

(The point q is $p^\perp \cap L$.) A geometry with this property (in which two points lie on at most one line) is called a *generalised quadrangle*.

Exercise 4.4 Show that a geometry satisfying the polar space axioms with $r = 2$ is a generalised quadrangle, and conversely.

We wish to recognise, within the geometry, the remaining lines of the projective space. These correspond to hyperbolic planes in the vector space, so we will call them H-lines. Note that the points of a H-line are pairwise non-orthogonal.

We observe that, given any two points p,q not lying on an F-line, the set

$$\{r : pr \text{ and } qr \text{ are F-lines}\}$$

is the set of points of $\{p,q\}^{\perp}$, and hence is the H-line containing p and q. This definition works in any generalized quadrangle, but in this case we have more:

(b) Any two points lie on either a unique F-line or a unique H-line.

(c) The F-lines and H-lines within a set $p^{\perp}$ form a projective plane.

(d) Any three non-collinear points lie in a unique set $p^{\perp}$.

Exercise 4.5 Prove conditions (b)–(d).

Conditions (a)–(d) guarantee that the geometry of F-lines and H-lines is a projective space, hence is isomorphic to $\mathrm{PG}(3,F)$ for some (possibly non-commutative) field F. Then the correspondence $p \leftrightarrow p^{\perp}$ is a polarity of the projective space, such that each point is incident with the corresponding plane. By the Fundamental Theorem of Projective Geometry, this polarity is induced by a symplectic form B on a vector space V of rank 4 over F (which is necessarily commutative).

Hence, again by the FTPG, the automorphism group of the geometry is induced by the group of semilinear transformations of V which preserve the set of pairs $\{(x,y) : B(x,y) = 0\}$. These transformations are composites of linear transformations preserving B up to a scalar factor, and field automorphisms. It follwos that, if $F \neq \mathrm{GF}(2)$, the automorphism group of the geometry has a unique minimal normal subgroup, which is isomorphic to $\mathrm{PSp}(4,F)$.