3 Polarities and forms

3.1 Sesquilinear forms

We saw in Chapter 1 that the projective space $\mathrm{PG}(n-1,F)$ is isomorphic to its dual if and only if the field F is isomorphic to its opposite. More precisely, we have the following. Let σ be an anti-automorphism of F, and V an F-vector space of rank n. A *sesquilinear form* B on V is a function $B : V \times V \to F$ which satisfies the following conditions:

(a) $B(c_1x_1 + c_2x_2, y) = c_1B(x_1,y) + c_2B(x_2,y)$, that is, B is a linear function of its first argument;

(b) $B(x, c_1y_1 + c_2y_2) = B(x,y_1)c_1^\sigma + B(x,y_2)c_2^\sigma$, that is, B is a semilinear function of its second argument, with field anti-automorphism σ.

(The word 'sesquilinear' means 'one-and-a-half'.) If σ is the identity (so that F is commutative), we say that B is a *bilinear form*.

The *left radical* of B is the subspace $\{x \in V : (\forall y \in V)B(x,) = 0\}$, and the *right radical* is the subspace $\{y \in V : (\forall x \in V)B(x,y) = 0\}$.

Exercise 3.1 (a) Prove that the left and right radicals are subspaces.

(b) Show that the left and right radicals have the same rank (if V has finite rank).

(c) Construct a bilinear form on a vector space of infinite rank such that the left radical is zero and the right radical is no-zero.

The sesquilinear form B is called *non-degenerate* if its left and right radicals are zero. (By the preceding exercise, it suffices to assume that one of the radicals is zero.)

A non-degenerate sesquilinear form induces a duality of $\mathrm{PG}(n-1,F)$ (an isomorphism from $\mathrm{PG}(n-1,F)$ to $\mathrm{PG}(n-1,F^\circ)$) as follows: for any $y \in V$, the map $x \mapsto B(x,y)$ is a linear map from V to F, that is, an element of the dual space V^* (which is a left vector space of rank n over F°); if we call this element β_y, then the map $y \mapsto \beta_y$ is a σ-semilinear bijection from V to V^*, and so induces the required duality.

Theorem 3.1 *For $n \geq 3$, any duality of $\mathrm{PG}(n-1,F)$ is induced in this way by a non-degenerate sesquilinear form on $V = F^n$.*

Proof By the Fundamental Theorem of Projective Geometry, a duality is induced by a σ-semilinear bijection ϕ from V to V^*, for some anti-automorphism σ. Set

$$B(x,y) = x(y\phi). \quad \blacksquare$$

We can short-circuit the passage to the dual space, and write the duality as

$$U \mapsto U^\perp = \{x \in V : B(x,y) = 0 \text{ for all } y \in U\}.$$

Obviously, a duality applied twice is a collineation. The most important types of dualities are those whose square is the identity. A *polarity* of $\mathrm{PG}(n,F)$ is a duality $\perp$ which satisfies $U^{\perp\perp} = U$ for all flats U of $\mathrm{PG}(n,F)$.

It will turn out that polarities give rise to a class of geometries (the polar spaces) with properties similar to those of projective spaces, and define groups analogous to the projective groups. If a duality is not a polarity, then any collineation which respects it must commute with its square, which is a collineation; so the group we obtain will lie inside the centraliser of some element of the collineation group. So the "largest" subgroups obtained will be those preserving polarities.

A sesquilinear form B is *reflexive* if $B(x,y) = 0$ implies $B(y,x) = 0$.

Proposition 3.2 *A duality is a polarity if and only if the sesquilinear form defining it is reflexive.*

Proof B is reflexive if and only if $x \in \langle y\rangle^\perp \Rightarrow y \in \langle x\rangle^\perp$. Hence, if B is reflexive, then $U \subseteq U^{\perp\perp}$ for all subspaces U. But by non-degeneracy, $\dim U^{\perp\perp} = \dim V - \dim U^\perp = \dim U$; and so $U = U^{\perp\perp}$ for all U. Conversely, given a polarity $\perp$, if $y \in \langle x\rangle^\perp$, then $x \in \langle x\rangle^{\perp\perp} \subseteq \langle y\rangle^\perp$ (since inclusions are reversed). $\blacksquare$

We now turn to the classification of reflexive forms. For convenience, from now on F will always be assumed to be commutative. (Note that, if the anti-automorphism σ is an automorphism, and in particular if σ is the identity, then F is automatically commutative.)

The form B is said to be σ-*Hermitian* if $B(y,x) = B(x,y)^\sigma$ for all $x,y \in V$. If B is a non-zero σ-Hermitian form, then

(a) for any x, $B(x,x)$ lies in the fixed field of σ;

(b) $\sigma^2 = 1$. For every scalar c is a value of B, say $B(x,y) = c$; then

$$c^{\sigma^2} = B(x,y)^{\sigma^2} = B(y,x)^\sigma = B(x,y) = c.$$

If σ is the identity, such a form (which is bilinear) is called *symmetric*.

A bilinear form b is called *alternating* if $B(x,x) = 0$ for all $x \in V$. This implies that $B(x,y) = -B(y,x)$ for all $x,y \in V$. For

$$0 = B(x+y,x+y) = B(x,x) + B(x,y) + B(y,x) + B(y,y) = B(x,y) + B(y,x).$$

Hence, if the characteristic is 2, then any alternating form is symmetric (but not conversely); but, in characteristic different from 2, only the zero form is both symmetric and alternating.

Clearly, an alternating or Hermitian form is reflexive. Conversely, we have the following:

Theorem 3.3 *A non-degenerate reflexive σ-sesquilinear form is either alternating, or a scalar multiple of a σ-Hermitian form. In the latter case, if σ is the identity, then the scalar can be taken to be* 1.

Proof I will give the proof just for a bilinear form. Thus, it must be proved that a non-degenerate reflexive bilinear form is either symmetric or alternating.

We have

$$B(u,v)B(u,w) - B(u,w)B(u,v) = 0$$

by commutativity; that is, using bilinearity,

$$B(u, B(u,v)w - B(u,w)v) = 0.$$

By reflexivity,

$$B(B(u,v)w - B(u,w)v, u) = 0,$$

whence bilinearity again gives

$$B(u,v)B(w,u) = B(u,w)B(v,u). \tag{1}$$

Call a vector u *good* if $B(u,v) = B(v,u) \neq 0$ for some v. By Equation (1), if u is good, then $B(u,w) = B(w,u)$ for all w. Also, if u is good and $B(u,v) \neq 0$, then v is good. But, given any two non-zero vectors u_1, u_2, there exists v with $B(u_i,v) \neq 0$ for $i = 1,2$. (For there exist v_1, v_2 with $B(u_i,v_i) \neq 0$ for $i = 1,2$, by non-degeneracy; and at least one of v_1, v_2, v_1+v_2 has the required property.) So, if some vector is good, then every non-zero vector is good, and B is symmetric.

But, putting $u = w$ in Equation (1) gives

$$B(u,u)\,(B(u,v) - B(v,u)) = 0$$

for all u,v. So, if u is not good, then $B(u,u) = 0$; and, if no vector is good, then B is alternating. ■

Exercise 3.2 (a) Show that the left and right radicals of a reflexive form are equal.

(b) Assuming Theorem 3.3, prove that the assumption of non-degeneracy in the theorem can be removed.

Exercise 3.3 Let σ be a (non-identity) automorphism of F of order 2. Let E be the subfield $\mathrm{Fix}(\sigma)$.

(a) Prove that F is of degree 2 over E, i.e., a rank 2 E-vector space.

[See any textbook on Galois theory. Alternately, argue as follows: Take $\lambda \in F \setminus E$. Then λ is quadratic over E, so $E(\lambda)$ has degree 2 over E. Now $E(\lambda)$ contains an element ω such that $\omega^\sigma = -\omega$ (if the characteristic is not 2) or $\omega\sigma = \omega + 1$ (if the characteristic is 2). Now, given two such elements, their quotient or difference respectively is fixed by σ, so lies in E.]

(b) Prove that

$$\{\lambda \in F : \lambda\lambda^\sigma = 1\} = \{\varepsilon/\varepsilon^\sigma : \varepsilon \in F\}.$$

[The left-hand set clearly contains the right. For the reverse inclusion, separate into cases according as the characteristic is 2 or not.

If the characteristic is not 2, then we can take $F = E(\omega)$, where $\omega^2 = \alpha \in E$ and $\omega^\sigma = -\omega$. If $\lambda = 1$, then take $\varepsilon = 1$; otherwise, if $\lambda = a + b\omega$, take $\varepsilon = b\alpha + (a-1)\omega$.

If the characteristic is 2, show that we can take $F = E(\omega)$, where $\omega^2 + \omega + \alpha = 0$, $\alpha \in E$, and $\omega^\sigma = \omega + 1$. Again, if $\lambda = 1$, set $\varepsilon = 1$; else, if $\lambda = a + b\omega$, take $\varepsilon = (a+1) + b\omega$.]

Exercise 3.4 Use the result of the preceding exercise to complete the proof of Theorem 3.3 in general.

[If $B(u,u) = 0$ for all u, the form B is alternating and bilinear. If not, suppose that $B(u,u) \neq 0$ and let $B(u,u)^\sigma = \lambda B(u,u)$. Choosing ε as in Exercise 3.3 and re-normalising B, show that we may assume that $\lambda = 1$, and (with this choice) that B is Hermitian.]

3.2 Hermitian and quadratic forms

We now change ground slightly from the last section. On the one hand, we restrict things by excluding some bilinear forms from the discussion; on the other, we

introduce quadratic forms. The loss and gain exactly balance if the characteristic is not 2; but, in characteristic 2, we make a net gain.

Let σ be an automorphism of the commutative field F, of order dividing 2. Let $\mathrm{Fix}(\sigma) = \{\lambda \in F : \lambda^\sigma = \lambda\}$ be the *fixed field* of σ, and $\mathrm{Tr}(\sigma) = \{\lambda + \lambda^\sigma : \lambda \in F\}$ the *trace* of σ. Since σ^2 is the identity, it is clear that $\mathrm{Fix}(\sigma) \supseteq \mathrm{Tr}(\sigma)$. Moreover, if σ is the identity, then $\mathrm{Fix}(\sigma) = F$, and

$$\mathrm{Tr}(\sigma) = \begin{cases} 0 & \text{if } F \text{ has characteristic 2,} \\ F & \text{otherwise.} \end{cases}$$

Let B be a σ-Hermitian form. We observed in the last section that $B(x,x) \in \mathrm{Fix}(\sigma)$ for all $x \in V$. We call the form B *trace-valued* if $B(x,x) \in \mathrm{Tr}(\sigma)$ for all $x \in V$.

Exercise 3.5 Let σ be an automorphism of a commutative field F such that σ^2 is the identity.

(a) Prove that $\mathrm{Fix}(\sigma)$ is a subfield of F.

(b) Prove that $\mathrm{Tr}(\sigma)$ is closed under addition, and under multiplication by elements of $\mathrm{Fix}(\sigma)$.

Proposition 3.4 $\mathrm{Tr}(\sigma) = \mathrm{Fix}(\sigma)$ *unless the characteristic of* F *is* 2 *and* σ *is the identity.*

Proof $E = \mathrm{Fix}(\sigma)$ is a field, and $K = \mathrm{Tr}(\sigma)$ is an E-vector space contained in E (Exercise 3.5). So, if $K \neq E$, then $K = 0$, and σ is the map $x \mapsto -x$. But, since σ is a field automorphism, this implies that the characteristic is 2 and σ is the identity. ■

Thus, in characteristic 2, symmetric bilinear forms which are not alternating are not trace-valued; but this is the only obstruction. We introduce quadratic forms to repair this damage. But, of course, quadratic forms can be defined in any characteristic. However, we note at this point that Theorem 3.3 depends in a crucial way on the commutativity of F; this leaves open the possibility of additional types of polar spaces defined by so-called *pseudoquadratic forms*. We will not pursue this here: see Tits's classification of spherical buildings.

Let V be a vector space over F. A *quadratic form* on V is a function $q : V \to F$ satisfying

(a) $q(\lambda x) = \lambda^2 f(x)$ for all $\lambda \in F$, $x \in V$;

(b) $q(x+y) = q(x) + q(y) + B(x,y)$, where B is bilinear.

Now, if the characteristic of F is not 2, then B is a symmetric bilinear form. Each of q and B determines the other, by

$$\begin{aligned} B(x,y) &= q(x+y) - q(x) - q(y), \\ q(x) &= \tfrac{1}{2}B(x,x), \end{aligned}$$

the latter equation coming from the substitution $x = y$ in (b). So nothing new is obtained.

On the other hand, if the characteristic of F is 2, then B is an alternating bilinear form, and q cannot be recovered from B. Indeed, many different quadratic forms correspond to the same bilinear form. (Note that the quadratic form does give extra structure to the vector space; we'll see that this structure is geometrically similar to that provided by an alternating or Hermitian form.)

We say that the bilinear form B is obtained by *polarisation* of q.

Now let B be a symmetric bilinear form over a field of characteristic 2, which is not alternating. Set $f(x) = B(x,x)$. Then we have

$$\begin{aligned} f(\lambda x) &= \lambda^2 f(x), \\ f(x+y) &= f(x) + f(y), \end{aligned}$$

since $B(x,y) + B(y,x) = 0$. Thus f is "almost" a semilinear form; the map $\lambda \mapsto \lambda^2$ is a homomorphism of the field F with kernel 0, but it may fail to be an automorphism. But in any case, the kernel of f is a subspace of V, and the restriction of B to this subspace is an alternating bilinear form. So again, in the spirit of the vague comment motivating the study of polarities in the last section, the structure provided by the form B is not "primitive". For this reason, we do not consider symmetric bilinear forms in characteristic 2 at all. However, as indicated above, we will consider quadratic forms in characteristic 2.

Now, in characteristic different from 2, we can take either quadratic forms or symmetric bilinear forms, since the structural content is the same. For consistency, we will take quadratic forms in this case too. This leaves us with three "types" of forms to study: alternating bilinear forms; σ-Hermitian forms where σ is not the identity; and quadratic forms.

We have to define the analogue of non-degeneracy for quadratic forms. Of course, we could require that the bilinear form obtained by polarisation is non-

degenerate; but this is too restrictive. We say that a quadratic form q is *non-degenerate* if

$$q(x) = 0 \ \& \ (\forall y \in V) B(x,y) = 0 \quad \Rightarrow \quad x = 0,$$

where B is the associated bilinear form; that is, if the form q is non-zero on every non-zero vector of the radical.

If the characteristic is not 2, then non-degeneracy of the quadratic form and of the bilinear form are equivalent conditions.

Now suppose that the characteristic is 2, and let W be the radical of B. Then B is identically zero on W; so the restriction of q to W satisfies

$$\begin{aligned} q(x+y) &= q(x)+q(y), \\ q(\lambda x) &= \lambda^2 q(x). \end{aligned}$$

As above, f is very nearly semilinear.

The field F is called *perfect* if every element is a square. If F is perfect, then the map $x \mapsto x^2$ is onto, and hence an automorphism of F; so q is indeed semilinear, and its kernel is a hyperplane of W. We conclude:

Theorem 3.5 *Let q be a non-singular quadratic form, which polarises to B, over a field F.*

(a) If the characteristic of F is not 2, then B is non-degenerate.

(b) If F is a perfect field of characteristic 2, then the radical of B has rank at most 1.

Exercise 3.6 Let B be an alternating bilinear form on a vector space V over a field F of characteristic 2. Let $(v_i : i \in I)$ be a basis for V, and $(c_i : i \in I)$ any function from I to F. Show that there is a unique quadratic form q with the properties that $q(v_i) = c_i$ for every $i \in I$, and q polarises to B.

Exercise 3.7 (a) Construct an imperfect field of characteristic 2.

(b) Construct a non-singular quadratic form with the property that the radical of the associated bilinear form has rank greater than 1.

Exercise 3.8 Show that finite fields of characteristic 2 are perfect.

Exercise 3.9 Let B be a σ-Hermitian form on a vector space V over F, where σ is not the identity. Set $f(x) = B(x,x)$. Let $E = \mathrm{Fix}(\sigma)$, and let V' be V regarded as an E-vector space by restricting scalars. Prove that f is a quadratic form on V', which polarises to the bilinear form $\mathrm{Tr}(B)$ defined by $\mathrm{Tr}(B)(x,y) = B(x,y) + B(x,y)^\sigma$. Show further that $\mathrm{Tr}(B)$ is non-degenerate if and only if B is.

3.3 Classification of forms

As explained in the last section, we now consider a vector space V of finite rank equipped with a form of one of the following types: a non-degenerate alternating bilinear form B; a non-degenerate trace-valued σ-Hermitian form B, where σ is not the identity; or a non-singular quadratic form q. In the third case, we let B be the bilinear form obtained by polarising q; then B is alternating or symmetric according as the characteristic is or is not 2, but B may be degenerate. We also let f denote the function q. In the other two cases, we define a function $f : V \to F$ by $f(x) = B(x,x)$ — this is identically zero if b is alternating. See Exercise 3.10 for the Hermitian case.

Such a pair (V,B) or (V,q) will be called a *formed space*.

Exercise 3.10 Let B be a σ-Hermitian form on a vector space V over F, where σ is not the identity. Set $f(x) = B(x,x)$. Let $E = \mathrm{Fix}(\sigma)$, and let V' be V regarded as an E-vector space by restricting scalars. Prove that f is a quadratic form on V', which polarises to the bilinear form $\mathrm{Tr}(B)$ defined by $\mathrm{Tr}(B)(x,y) = B(x,y) + B(x,y)^\sigma$. Show further that $\mathrm{Tr}(b)$ is non-degenerate if and only if B is.

We say that V is *anisotropic* if $f(x) \neq 0$ for all $x \neq 0$. Also, V is a *hyperbolic plane* if it is spanned by vectors v and w with $f(v) = f(w) = 0$ and $B(v,w) = 1$. (The vectors v and w are linearly independent, so V has rank 2.)

Theorem 3.6 *A non-degenerate formed space is the direct sum of a number r of hyperbolic lines and an anisotropic space U. The number r and the isomorphism type of U are invariants of V.*

Proof If V is anisotropic, then there is nothing to prove, since V cannot contain a hyperbolic plane. So suppose that V contains a vector $v \neq 0$ with $f(v) = 0$.

We claim that there is a vector w with $B(v,w) \neq 0$. In the alternating and Hermitian cases, this follows immediately from the non-degeneracy of the form. In the quadratic case, if no such vector exists, then v is in the radical of B; but v is a singular vector, contradicting the non-degeneracy of f.

Multiplying w by a non-zero constant, we may assume that $B(v,w) = 1$.

Now, for any value of λ, we have $B(v, w - \lambda v) = 1$. We wish to choose λ so that $f(w - \lambda v) = 0$; then v and w will span a hyperbolic line. Now we distinguish cases.

(a) If B is alternating, then any value of λ works.

(b) If B is Hermitian, we have

$$\begin{aligned} f(w-\lambda v) &= f(w)-\lambda B(v,w)-\lambda^\sigma B(w,v)+\lambda\lambda^\sigma f(v) \\ &= f(w)-(\lambda+\lambda^\sigma); \end{aligned}$$

and, since B is trace-valued, there exists λ with $\mathrm{Tr}(\lambda) = f(w)$.

(c) Finally, if $f = q$ is quadratic, we have

$$\begin{aligned} f(w-\lambda v) &= f(w)-\lambda B(w,v)+\lambda^2 f(v) \\ &= f(w)-\lambda, \end{aligned}$$

so we choose $\lambda = f(w)$.

Now let W_1 be the hyperbolic line $\langle v, w-\lambda v\rangle$, and let $V_1 = W_1^\perp$, where orthogonality is defined with respect to the form B. It is easily checked that $V = V_1 \oplus W_1$, and the restriction of the form to V_1 is still non-degenerate. Now the existence of the decomposition follows by induction.

The uniqueness of the decomposition will be proved later, as a consequence of Witt's Lemma (Theorem 3.15). ■

The number r of hyperbolic lines is called the *polar rank* of V, and (the isomorphism type of) U is called the *germ* of V.

To complete the classification of forms over a given field, it is necessary to determine all the anisotropic spaces. In general, this is not possible; for example, the study of positive definite quadratic forms over the rational numbers leads quickly into deep number-theoretic waters. I will consider the cases of the real and complex numbers and finite fields.

First, though, the alternating case is trivial:

Proposition 3.7 *The only anisotropic space carrying an alternating bilinear form is the zero space.*

In combination with Theorem 3.6, this shows that a space carrying a non-degenerate alternating bilinear form is a direct sum of hyperbolic planes.

Over the real numbers, Sylvester's theorem asserts that any quadratic form in n variables is equivalent to the form

$$x_1^2+\ldots+x_r^2-x_{r+1}^2-\ldots-x_{r+s}^2,$$

for some r, s with $r+s \le n$. If the form is non-singular, then $r+s = n$. If both r and s are non-zero, there is a non-zero singular vector (with 1 in positions 1 and $r+1$, 0 elsewhere). So we have:

Proposition 3.8 *If V is a real vector space of rank n, then an anisotropic form on V is either positive definite or negative definite; there is a unique form of each type up to invertible linear transformation, one the negative of the other.* ■

The reals have no non-identity automorphisms, so Hermitian forms do not arise.

Over the complex numbers, the following facts are easily shown:

(a) There is a unique non-singular quadratic form (up to equivalence) in n variables for any n. A space carrying such a form is anisotropic if and only if $n \leq 1$.

(b) If σ denotes complex conjugation, the situation for σ-Hermitian forms is the same as for quadratic forms over the reals: anisotropic forms are positive or negative definite, and there is a unique form of each type, one the negative of the other.

For finite fields, the position is as follows.

Theorem 3.9 *(a) An anisotropic quadratic form in n variables over $\mathrm{GF}(q)$ exists if and only if $n \leq 2$. There is a unique form for each n except when $n = 1$ and q is odd, in which case there are two forms, one a non-square multiple of the other.*

(b) Let $q = r^2$ and let σ be the field automorphism $\alpha \mapsto \alpha^r$. Then there is an anisotropic σ-Hermitian form in n variables if and only if $n \leq 1$. The form is unique in each case.

Proof (a) Consider first the case where the characteristic is not 2. The multiplicative group of $\mathrm{GF}(q)$ is cyclic of even order $q-1$; so the squares form a subgroup of index 2, and if η is a fixed non-square, then every non-square has the form $\eta\alpha^2$ for some α. It follows easily that any quadratic form in one variable is equivalent to either x^2 or ηx^2.

Next, consider non-singular forms in two variables. By completing the square, such a form is equivalent to one of x^2+y^2, $x^2+\eta y^2$, $\eta x^2+\eta y^2$.

Suppose first that $q \equiv 1 \pmod 4$. Then -1 is a square, say $-1 = \beta^2$. (In the multiplicative group, -1 has order 2, so lies in the subgroup of even order $\frac{1}{2}(q-1)$ consisting of squares.) Thus $x^2+y^2 = (x+\beta y)(x-\beta y)$, and the first and third forms are not anisotropic. Moreover, any form in 3 or more variables, when

converted to diagonal form, contains one of these two, and so is not anisotropic either.

Now consider the other case, $q \equiv -1 \pmod 4$. Then -1 is a non-square (since the group of squares has odd order), so the second form is $(x+y)(x-y)$, and is not anisotropic. Moreover, the set of squares is not closed under addition (else it would be a subgroup of the additive group, but $\frac{1}{2}(q+1)$ doesn't divide q); so there exist two squares whose sum is a non-square. Multiplying by a suitable square, there exist β, γ with $\beta^2 + \gamma^2 = -1$. Then

$$-(x^2+y^2) = (\beta x + \gamma y)^2 + (\gamma x - \beta y)^2,$$

and the first and third forms are equivalent. Moreover, a form in three variables is certainly not anisotropic unless it is equivalent to $x^2+y^2+z^2$, and this form vanishes at the vector $(\beta, \gamma, 1)$; hence there is no anisotropic form in three or more variables.

The characteristic 2 case is an exercise (see below).

(b) Now consider Hermitian forms. If σ is an automorphism of $\mathrm{GF}(q)$ of order 2, then q is a square, say $q = r^2$, and $\alpha^\sigma = \alpha^r$. We need the fact that every element of $\mathrm{Fix}(\sigma) = \mathrm{GF}(r)$ has the form $\alpha\alpha^\sigma$ (see Exercise 3.3).

In one variable, we have $f(x) = \mu x x^\sigma$ for some non-zero $\mu \in \mathrm{Fix}(\sigma)$; writing $\mu = \alpha\alpha^\sigma$ and replacing x by αx, we can assume that $\mu = 1$.

In two variables, we can similarly take the form to be $xx^\sigma + yy^\sigma$. Now $-1 \in \mathrm{Fix}(\sigma)$, so $-1 = \lambda\lambda^\sigma$; then the form vanishes at $(1, \lambda)$. It follows that there is no anisotropic form in any larger number of variables either. ■

Exercise 3.11 Prove that there is, up to equivalence, a unique non-degenerate alternating bilinear form on a vector space of countably infinite dimension (a direct sum of countably many isotropic planes).

Exercise 3.12 Let F be a finite field of characteristic 2.

(a) Prove that every element of F has a unique square root.

(b) By considering the bilinear form obtained by polarisation, prove that a non-singular form in 2 or 3 variables over F is equivalent to $\alpha x^2 + xy + \beta y^2$ or $\alpha x^2 + xy + \beta y^2 + \gamma z^2$ respectively. Prove that forms of the first shape (with $\alpha, \beta \neq 0$) are all equivalent, while those of the second shape cannot be anisotropic.

3.4 Polar spaces

Polar spaces describe the geometry of vector spaces carrying a reflexive sesquilinear form or a quadratic form in much the same way as projective spaces describe the geometry of vector spaces. We now embark on the study of these geometries; the three preceding sections contain the prerequisite algebra.

First, some terminology. The polar spaces associated with the three types of forms (alternating bilinear, Hermitian, and quadratic) are referred to by the same names as the groups associated with them: *symplectic*, *unitary*, and *orthogonal* respectively. Of what do these spaces consist?

Let V be a vector space carrying a form of one of our three types. Recall that as well as a sesquilinear form b in two variables, we have a form f in one variable — either f is defined by $f(x) = B(x,x)$, or b is obtained by polarising f — and we make use of both forms. A subspace of V on which B vanishes identically is called a *B-flat subspace*, and one on which f vanishes identically is called a *f-flat subspace*. (Note: these terms are not standard; in the literature, such spaces are called *totally isotropic* (t.i.) and *totally singular* (t.s.) respectively.) The unqualified term *flat subspace* will mean a B-flat subspace in the symplectic or unitary case, and a q-flat subspace in the orthogonal case.

The *polar space* associated with a vector space carrying a form is the geometry whose flats are the flat subspaces (in the above sense). Note that, if the form is anisotropic, then the only member of the polar space is the zero subspace. The *polar rank* of a classical polar space is the largest vector space rank of any flat subspace; it is zero if and only if the form is anisotropic. Where there is no confusion, polar rank will be called simply *rank*. (We will soon see that there is no conflict with our earlier definition of rank as the number of hyperbolic planes in the decomposition of the space.) We use the terms *point, line, plane*, etc., just as for projective spaces.

Polar spaces bear the same relation to formed spaces as projective spaces do to vector spaces.

We now proceed to derive some properties of polar spaces. Let Γ be a classical polar space of polar rank r.

(P1) Any flat, together with the flats it contains, is a projective space of dimension at most $r-1$.

(P2) The intersection of any family of flats is a flat.

(P3) If U is a flat of dimension $r-1$ and p a point not in U, then the union of the

planes joining p to points of U is a flat W of dimension $r-1$; and $U \cap W$ is a hyperplane in both U and W.

(P4) There exist two disjoint flats of dimension $r-1$.

(P1) is clear since a subspace of a flat subspace is itself flat. (P2) is also clear. To prove (P3), let $p = \langle y \rangle$. The function $x \mapsto B(x,y)$ on the vector space U is linear; let K be its kernel, a hyperplane in U. Then the line (of the projective space) joining p to a point $q \in U$ is flat if and only if $q \in K$; and the union of all such flat lines is a flat space $W = \langle K, y \rangle$, such that $W \cap U = K$, as required.

Finally, to prove (P4), we use the hyperbolic-anisotropic decomposition again. If $L_1, \ldots, L_r$ are the hyperbolic planes, and x_i, y_i are the distinguished spanning vectors in L_i, then the required flats are $\langle x_1, \ldots, x_r \rangle$ and $\langle y_1, \ldots, y_r \rangle$.

The significance of the geometric properties (P1)–(P4) lies in the major result of Veldkamp and Tits which determines all the geometries of rank at least 3 which satisfy them. All these geometries are polar spaces (as we have defined them) or slight generalisations, together with a couple of exceptions of rank 3. In particular, the following theorem holds:

Theorem 3.10 *A finite geometry satisfying (P1)–(P4) with $r \geq 3$ is a polar space.*

Exercise 3.13 Let $P = \mathrm{PG}(3, F)$ for some (not necessarily commutative) division ring F. Construct a new geometry Γ as follows:

(a) the 'points' of Γ are the lines of P;

(b) the 'lines' of Γ are the plane pencils in P (consisting of all lines lying in a plane Π and containing a point p of Π);

(c) the 'planes' of Γ are of two types: the pencils (consisting of all the lines through a point) and the dual planes (consisting of all the lines in a plane).

Prove that Γ satisfies (P1)–(P4) with $r = 3$.

Prove that, if F is not isomorphic to its opposite, then Γ contains non-isomorphic planes.

(We will see later that, if F is commutative, then Γ is an orthogonal polar space.)

Exercise 3.14 Prove the *Buekenhout–Shult property* of the geometry of points and lines in a polar space: if p is a point not lying on a line L, then p is collinear with one or all points of L.

You should prove this both from the analytic description of polar spaces, and using (P1)–(P4).

In a polar space Γ, given any set S of points, we let $S^\perp$ denote the set of points which are perpendicular to (that is, collinear with) every point of S. Polar spaces have good inductive properties. Let G be a classical polar space. There are two natural ways of producing a "smaller" polar space from G:

(a) Take a point x of G, and consider the quotient space $x^\perp/x$, the space whose points, lines, ... are the lines, planes, ... of G containing x.

(b) Take two non-perpendicular points x and y, and consider $\{x,y\}^\perp$.

In each case, the space constructed is a classical polar space, having the same germ as G but with polar rank one less than that of G. (Note that, in (b), the span of x and y in the vector space is a hyperbolic plane.)

Exercise 3.15 Prove the above assertions.

There are more general versions. For example, if S is a flat of dimension $d-1$, then $S^\perp/S$ is a polar space of rank $r-d$ with the same germ as G. We will see below how this inductive process can be used to obtain information about polar spaces.

We investigate just one type in more detail, the so-called *hyperbolic quadric*, the orthogonal space which is a direct sum of hyperbolic planes (that is, having germ 0). The quadratic form defining this space can be taken to be $x_1x_2 + x_3x_4 + \ldots + x_{2r-1}x_{2r}$.

Proposition 3.11 *The maximal flats of a hyperbolic quadric fall into two classes, with the properties that the intersection of two maximal flats has even codimension in each if and only if they belong to the same class.*

Proof First, note that the result holds when $r = 1$, since then the quadratic form is x_1x_2 and there are just two singular points, $\langle(1,0)\rangle$ and $\langle(0,1)\rangle$. By the inductive principle, it follows that any flat of dimension $r-2$ is contained in exactly two maximal flats.

We take the $(r-1)$-flats and $(r-2)$-flats as the vertices and edges of a graph Γ, that is, we join two $(r-1)$-flats if their intersection is an $(r-2)$-flat. The theorem will follow if we show that Γ is connected and bipartite, and that the distance between two vertices of Γ is the codimension of their intersection. Clearly the

codimension of the intersection increases by at most one with every step in the graph, so it is at most equal to the distance. We prove equality by induction.

Let U be a $(r-1)$-flat and K a $(r-2)$-flat. We claim that the two $(r-1)$-spaces W_1, W_2 containing K have different distances from U. Factoring out the flat subspace $U \cap K$ and using induction, we may assume that $U \cap K = \emptyset$. Then $U \cap K^\perp$ is a point p, which lies in one but not the other of W_1, W_2; say $p \in W_1$. By induction, the distance from U to W_1 is $r-1$; so the distance from U to W_2 is at most r, hence equal to r by the remark in the preceding paragraph.

This establishes the claim about the distance. The fact that Γ is bipartite also follows, since in any non-bipartite graph there exists an edge both of whose vertices have the same distance from some third vertex, and the argument given shows that this doesn't happen in Γ. ■

In particular, the rank 2 hyperbolic quadric consists of two families of lines forming a *grid*, as shown in Figure 1. This is the so-called "ruled quadric", familiar from models such as wastepaper baskets.

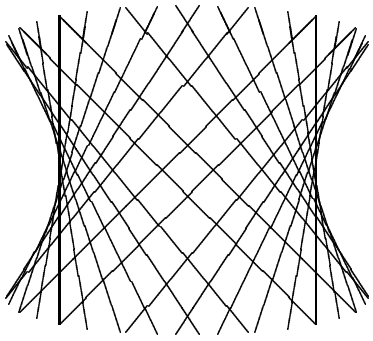

Figure 1: A ruled quadric

Exercise 3.16 Show that Proposition 3.11 can be proved using only properties (P1)–(P4) of polar spaces together with the fact that an $(r-1)$-flat lies in exactly two maximal flats.

3.5 Finite polar spaces

The classification of finite classical polar spaces was achieved by Theorem 3.6. We subdivide these spaces into six families according to their germ, viz., one

symplectic, two unitary, and three orthogonal. (Forms which differ only by a scalar factor obviously define the same polar space.) The following table gives some information about them. In the table, r denotes the polar space rank, and δ the vector space rank of the germ; the rank n of the space is given by $n = 2r + \delta$. The significance of the parameter ε will emerge shortly. This number, depending only on the germ, carries numerical information about all spaces in the family. Note that, in the unitary case, the order of the finite field must be a square.

Type	δ	ε
Symplectic	0	0
Unitary	0	$-\frac{1}{2}$
Unitary	1	$\frac{1}{2}$
Orthogonal	0	-1
Orthogonal	1	0
Orthogonal	2	1

Table 1: Finite polar spaces

Theorem 3.12 *The number of points in a finite polar space of rank 1 is $q^{1+\varepsilon}+1$, where ε is given in Table 1.*

Proof Let V be a vector space carrying a form of rank 1 over GF(q). Then V is the orthogonal direct sum of a hyperbolic line L and an anisotropic germ U of dimension k (say). Let n_k be the number of points.

Suppose that $k > 0$. If p is a point of the polar space, then p lies on the hyperplane $p^\perp$; any other hyperplane containing p is non-degenerate with polar rank 1 and having germ of dimension $k-1$. Consider a parallel class of hyperplanes in the affine space whose hyperplane at infinity is $p^\perp$. Each such hyperplane contains $n_{k-1}-1$ points, and the hyperplane at infinity contains just one, viz., p. So we have

$$n_k - 1 = q(n_{k-1} - 1),$$

from which it follows that $n_k = 1 + (n_0 - 1)q^k$. So it is enough to prove the result for the case $k = 0$, that is, for a hyperbolic line.

In the symplectic case, each of the $q+1$ projective points on a line is isotropic.

Consider the unitary case. We can take the form to be

$$B((x_1,y_1),(x_2,y_2)) = x_1\overline{y_2} + y_1\overline{x_2},$$

where $\overline{x} = x^\sigma = x^r$, $r^2 = q$. So the isotropic points satisfy $x\overline{y} + y\overline{x} = 0$, that is, $\mathrm{Tr}(x\overline{y}) = 0$. How many pairs (x,y) satisfy this? If $y = 0$, then x is arbitrary. If $y \neq 0$, then a fixed multiple of x is in the kernel of the trace map, a set of size $q^{1/2}$ (since Tr is $\mathrm{GF}(q^{1/2})$-linear). So there are

$$q + (q-1)q^{1/2} = 1 + (q-1)(q^{1/2}+1)$$

vectors, i.e., $q^{1/2}+1$ projective points.

Finally, consider the orthogonal case. The quadratic form is equivalent to xy, and has two singular points, $\langle(1,0)\rangle$ and $\langle(1,0)\rangle$. ■

Theorem 3.13 *In a finite polar space of rank r, there are $(q^r-1)(q^{r+\varepsilon}+1)/(q-1)$ points, of which $q^{2r-1+\varepsilon}$ are not perpendicular to a given point.*

Proof We let $F(r)$ be the number of points, and $G(r)$ the number not perpendicular to a given point. (We do not assume that $G(r)$ is constant; this constancy follows from the induction that proves the theorem.) We use the two inductive principles described at the end of the last section.

Claim 1: $G(r) = q^2 G(r-1)$.

Take a point x, and count pairs (y,z), where $y \in x^\perp$, $z \notin x^\perp$, and $z \in y^\perp$. Choosing z first, there are $G(r)$ choices; then $\langle x,z\rangle$ is a hyperbolic line, and y is a point in $\langle x,z\rangle^\perp$, so there are $F(r-1)$ choices for y. On the other hand, choosing y first, the lines through y are the points of the rank $r-1$ polar space $x^\perp/x$, and so there are $F(r-1)$ of them, with q points different from x on each, giving $qF(r-1)$ choices for y; then $\langle x,y\rangle$ and $\langle y,z\rangle$ are non-perpendicular lines in $y^\perp$, i.e., points of $y^\perp/y$, so there are $G(r-1)$ choices for $\langle y,z\rangle$, and so $qG(r-1)$ choices for y. thus

$$G(r)\cdot F(r-1) = qF(r-1)\cdot qG(r-1),$$

from which the result follows.

Since $G(1) = q^{1+\varepsilon}$, it follows immediately that $G(r) = q^{2r-1+\varepsilon}$, as required.

Claim 2: $F(r) = 1 + qF(r-1) + G(r)$.

For this, simply observe (as above) that points perpendicular to x lie on lines of $x^\perp/x$.

Now it is just a matter of calculation that the function $(q^r-1)(q^{r+\varepsilon}+1)/(q-1)$ satisfies the recurrence of Claim 2 and correctly reduces to $q^{1+\varepsilon}+1$ when $r = 1$. ■

Theorem 3.14 *The number of maximal flats in a finite polar space of rank r is*

$$\prod_{i=1}^{r}(1+q^{i+\varepsilon}).$$

Proof Let $H(r)$ be this number. Count pairs (x,U), where U is a maximal flat and $x \in U$. We find that

$$F(r) \cdot H(r-1) = H(r) \cdot (q^r-1)/(q-1),$$

so

$$H(r) = (1+q^{r+\varepsilon})H(r-1).$$

Now the result is immediate. ■

It should now be clear that any reasonable counting question about finite polar spaces can be answered in terms of q, r, ε. We will do this for the associated classical groups at the end of the next section.

3.6 Witt's Lemma

Let V be a formed space, with sesquilinear form B and (if appropriate) quadratic form q. An *isometry* of V is a linear map $g : V \to V$ which satisfies $B(xg,yg) = B(x,y)$ for all $x,y \in V$, and (if appropriate) $q(xg) = q(x)$ for all $x \in V$. (Note that, in the case of a quadratic form, the second condition implies the first.)

The set of all isometries of V forms a group, the *isometry group* of V. This group is our object of study for the next few sections.

More generally, if V and W are formed spaces of the same type, an isometry from V to W is a linear map from V to W satisfying the conditions listed above.

Exercise 3.17 Let V be a (not necessarily non-degenerate) formed space of symplectic or Hermitian type, with radical $V^\perp$. Prove that the natural map from V to $V/V^\perp$ is an isometry.

The purpose of this subsection is to prove *Witt's Lemma*, a transitivity assertion about the isometry group of a formed space.

Theorem 3.15 *Suppose that U_1 and U_2 are subspaces of the formed space V, and $h : U_1 \to U_2$ is an isometry. Then there is an isometry g of V which extends h if and only if $(U_1 \cap V^\perp)h = U_2 \cap V^\perp$.*

In particular, if $V^\perp = 0$, then any isometry between subspaces of V extends to an isometry of V.

Proof Assume that $h : U_1 \to U_2$ is an isometry. Clearly, if h is the restriction of an isometry g of V, then $V^\perp g = V^\perp$, and so

$$(U_1 \cap V^\perp)h = (U_1 \cap V^\perp)g = U_1 g \cap V^\perp g = U_2 \cap V^\perp.$$

We have to prove the converse.

First we show that we may assume that U_1 and U_2 contain $V^\perp$. Suppose not. Choose a subspace W of $V^\perp$ which is a complement to both $U_1 \cap V^\perp$ and $U_2 \cap V^\perp$ (see Exercise 3.18), and extend h to $U_1 \oplus W$ by the identity map on W. This is easily checked to be an isometry to $U_2 \oplus W$.

The proof is by induction on $\mathrm{rk}(U_1/V^\perp)$. If $U_1 = V_\perp = U_2$, then choose any complement W for $V^\perp$ in V and extend h by the identity on W. So the base step of the induction is proved. Assume that the conclusion of Witt's Lemma holds for V', U_1', U_2', h' whenever $\mathrm{rk}(U_1'/(V')^\perp) < \mathrm{rk}(U_1/V^\perp)$.

Let H be a hyperplane of U_1 containing $V^\perp$. Then the restriction f' of f to H has an extension to an isometry g' of V. Now it is enough to show that $h(g')^{-1}$ extends to an isometry; in other words, we may assume that h is the identity on H. Moreover, the conclusion is clear if h is the identity on U_1; so suppose not. Then $\ker(h-1) = H$, and so the image of $h-1$ is a rank 1 subspace P of U_1.

Since h is an isometry, for all $x, y \in U_1$ we have

$$\begin{aligned} B(xh, yh-y) &= B(xh, yh) - B(xh, y) \\ &= B(x,y) - B(xh, y) \\ &= B(x - xh, y). \end{aligned}$$

So, if $y \in H$, then any vector $xh - x$ of P is orthogonal to y; that is, $H \leq P^\perp$.

Now suppose that $P \not\leq U_1^\perp$. Then $U_1 \cap P^\perp = U_2 \cap P^\perp = H$. If W is a complement to H in $P^\perp$, then we can extend h by the identity on W to obtain the required isometry. So we may assume further that $U_1, U_2 \leq P^\perp$. In particular, $P \leq P^\perp$.

Next we show that we may assume that $U_1 = U_2 = P^\perp$. Suppose first that $U_1 \neq U_2$. If $U_i = \langle H, u_i \rangle$ for $i = 1, 2$, let W_0 be a complement for $U_1 + U_2$ in $P^\perp$, and $W = \langle W_0, u_1 + u_2 \rangle$; then h can be extended by the identity on W to an isometry on $P^\perp$. If $U_1 = U_2$, take any complement W to U_1 in $P^\perp$. In either case, the extension is an isometry of $P^\perp$ which acts as the identity on a hyperplane H' of $P^\perp$ containing H. So we may replace U_1, U_2, H by $P^\perp, P^\perp, H'$.

Let $P = \langle x \rangle$ and let $x = uh - u$ for some $u \in U_1$. We have $B(x,x) = 0$. In the orthogonal case, we have

$$q(x) = q(uh - u) = q(uh) + q(u) - B(uh, u) = 2q(u) - B(u,u) = 0.$$

(We have $B(uh,u) = B(u,u)$ because $B(uh-u,u) = 0$.) So P is flat, and there is a hyperbolic plane $\langle u,v\rangle$, with $v \notin P^\perp$. Our job is to extend h to the vector v.

To achieve this, we show first that there is a vector v' such that $\langle uh,v'\rangle^\perp = \langle u,v\rangle^\perp$. This holds because $\langle u,v\rangle^\perp$ is a hyperplane in $\langle uh\rangle^\perp$ not containing $V^\perp$.

Next, we observe that $\langle uh,v'\rangle$ is a hyperbolic plane, so we can choose a vector v'' such that $B(uh,v'') = 1$ and (if relevant) $Q(v'') = 0$.

Finally, we observe that by extending h to map v to v'' we obtain the required isometry of V.

Exercise 3.18 Let U_1 and U_2 be subspaces of a vector space V having the same rank. Show that there is a subspace W of V which is a complement for both U_1 and U_2.

Corollary 3.16 *(a) The ranks of maximal flat subspaces of a formed space are all equal.*

(b) The Witt rank and isometry type of the germ of a non-degenerate formed space are invariants.

Proof (a) Let U_1 and U_2 be maximal flat subspaces. Then both U_1 and U_2 contains $V^\perp$. If $\mathrm{rk}(U_1) < \mathrm{rk}(U_2)$, there is an isometry h from U_1 into U_2. If g is the extension of h to V, then the image of U_2 under g^{-1} is a flat subspace properly containing U_1, contradicting maximality.

(b) The result is clear if V is anisotropic. Otherwise, let U_1 and U_2 be hyperbolic planes. Then U_1 and U_2 are isometric and are disjoint from $V^\perp$. An isometry of V carrying U_1 to U_2 takes $U_1^\perp$ to $U_2^\perp$. Then the result follows by induction. ∎

Theorem 3.17 *Let V_r be a non-degenerate formed space with polar rank r and germ W over $\mathrm{GF}(q)$. Let G_r be the isometry group of V_r. Then*

$$\begin{aligned} |G_r| &= \left(\prod_{i=1}^{r}(q^i-1)(q^{i+\varepsilon}+1)q^{2i-1+\varepsilon}\right)|G_0| \\ &= q^{r(r+\varepsilon)}\left(\prod_{i=1}^{r}(q^i-1)(q^{i+\varepsilon}+1)\right)|G_0|, \end{aligned}$$

where $|G_0|$ is given by the following table:

Type	δ	ε	$\lvert G_0\rvert$
Symplectic	0	0	1
Unitary	0	$-\frac{1}{2}$	1
Unitary	1	$\frac{1}{2}$	$q^{1/2}+1$
Orthogonal	0	-1	1
Orthogonal	1	0	$\begin{cases} 2 & (q \text{ odd}) \\ 1 & (q \text{ even}) \end{cases}$
Orthogonal	2	1	$2(q+1)$

Proof By Theorem 3.13, the number of choices of a vector x spanning a flat subspace is $(q^r-1)(q^{r+\varepsilon}+1)$. Then the number of choices of a vector y spanning a flat subspace and having inner product 1 with x is $q^{2r-1+\varepsilon}$. Then x and y span a hyperbolic plane. Now Witt's Lemma shows that G_r acts transitively on such pairs, and the stabiliser of such a pair is G_{r-1}, by the inductive principle.

In the cases where $\delta = 0$, G_0 is the trivial group on a vector space of rank 0. In the unitary case with $\delta = 1$, G_0 preserves the Hermitian form $xx^{q^{1/2}}$, so consists of multiplication by $(q^{1/2}+1)$st roots of unity. In the orthogonal case with $\delta = 1$, G_0 preserves the quadratic form x^2, and so consists of multiplication by ± 1 only. Finally, consider the orthogonal case with $\delta = 2$. Here we can represent the quadratic form as the norm from $\mathrm{GF}(q^2)$ to $\mathrm{GF}(q)$, that is, $N(x) = x^{q+1}$. The $\mathrm{GF}(q)$-linear maps which preserve this form a dihedral group of order $2(q+1)$: the cyclic group is generated by the $(q+1)$st roots of unity in $\mathrm{GF}(q^2)$, which is inverted by the non-trivial field automorphism over $\mathrm{GF}(q)$ (since, if $x^{q+1} = 1$, then $x^q = x^{-1}$).