6

Polar spaces

Now we begin on our second major theme, polar spaces. This chapter corresponds to the first half of Chapter 1, and gives the algebraic description of polar spaces. The algebraic background required is more elaborate (vector spaces with forms, rather than just vector spaces), accounting for the increased length. The first section, on polarities of projective spaces, provides motivation for the introduction of the (Hermitian and quadratic) forms.

6.1 Dualities and polarities

Recall that the dual V^* of a finite-dimensional (left) vector space V over a skew field F can be regarded as a left vector space of the same dimension over the opposite field F°, and there is thus an inclusion-reversing bijection between the projective spaces $\mathrm{PG}(n,F)$ and $\mathrm{PG}(n,F^\circ)$. If it happens that F and F° are isomorphic, then there exists a *duality* of $\mathrm{PG}(n,F)$, an inclusion-reversing bijection of $\mathrm{PG}(n,F)$.

Conversely, if $\mathrm{PG}(n,F)$ admits a duality (for $n > 1$), then F is isomorphic to F°, as follows from the FTPG (see Section 1.3). We will examine this conclusion and make it more detailed.

So let π be a duality of $\mathrm{PG}(n,F)$, $n > 1$. Composing π with the natural isomorphism from $\mathrm{PG}(n,F)$ to $\mathrm{PG}(n,F^\circ)$, we obtain an inclusion-preserving bijection θ from $\mathrm{PG}(n,F)$ to $\mathrm{PG}(n,F^\circ)$. According to the FTPG, θ is induced by a semilinear transformation T from $V = F^{n+1}$ to its dual space V^*, associated with an isomorphism $\sigma : F \to F^\circ$, which can be regarded as being an anti-automorphism of F:

that is,

$$\begin{aligned}(\mathbf{v}_1+\mathbf{v}_2)T &= \mathbf{v}_1T+\mathbf{v}_2T,\\ (\alpha\mathbf{v})T &= \alpha^\sigma\mathbf{v}T.\end{aligned}$$

Define a function $b : V\times V\to F$ by the rule

$$b(\mathbf{v},\mathbf{w}) = (\mathbf{v})(\mathbf{w}T),$$

that is, the result of applying the element $\mathbf{w}T$ of V^* to $\mathbf{v}$. Then b is a *sesquilinear form*: it is linear as a function of the first argument, and semilinear as a function of the second — this means that

$$b(\mathbf{v},\mathbf{w}_1+\mathbf{w}_2) = b(\mathbf{v},\mathbf{w}_1)+b(\mathbf{v},\mathbf{w}_2)$$

and

$$b(\mathbf{v},\alpha\mathbf{w}) = \alpha^\sigma b(\mathbf{v},\mathbf{w}).$$

(The prefix "sesqui-" means "one-and-a-half".) If we need to emphasise the anti-automorphism σ, we say that b is σ-sesquilinear. If σ is the identity, then the form is *bilinear*.

The form b is also *non-degenerate*, in the sense that

$$(\forall\mathbf{w}\in V)(b(\mathbf{v},\mathbf{w})=0 \quad\Rightarrow\quad \mathbf{v}=0$$

and

$$(\forall\mathbf{v}\in V)(b(\mathbf{v},\mathbf{w})=0 \quad\Rightarrow\quad \mathbf{w}=0.$$

(The second condition asserts that T is one-to-one, so that if $\mathbf{w}\neq 0$ then $\mathbf{w}T$ is a non-zero functional. The first asserts that T is onto: only the zero vector is annihilated by every functional in the dual space.)

So, we have:

Theorem 6.1 *Any duality of* $\mathrm{PG}(n,F)$*, for* $n>1$*, is induced by a non-degenerate* σ*-sesquilinear form on the underlying vector space, where* σ *is an anti-automorphism of* F. ■

Conversely, any non-degenerate sesquilinear form on V induces a duality. We can short-circuit the passage to the dual space, and write the duality as

$$U\mapsto U^\perp = \{\mathbf{v}\in V : b(\mathbf{v},\mathbf{w})=0 \text{ for all } \mathbf{w}\in U\}.$$

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 is a bit difficult to motivate the detailed study of polarities at this stage; but it will turn out that they give rise to a class of geometries (the polar spaces) with properties similar to those of projective spaces. To put it somewhat vaguely, we are trying to add some extra structure to a projective space; if a duality is not a polarity, then its square is a non-identity collineation, and some of the extra structure arises from this collineation. Only in the case of a polarity is the extra structure "primitive".

A sesquilinear form b is *reflexive* if $b(\mathbf{v},\mathbf{w}) = 0$ implies $b(\mathbf{w},\mathbf{v}) = 0$.

Proposition 6.2 *A duality is a polarity if and only if the sesquilinear form defining it is reflexive.*

Proof b is reflexive if and only if

$$\mathbf{v} \in \langle \mathbf{w} \rangle^{\perp} \Rightarrow \mathbf{w} \in \langle \mathbf{v} \rangle^{\perp}.$$

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 $\mathbf{w} \in \langle \mathbf{v} \rangle^{\perp}$, then $\mathbf{v} \in \langle \mathbf{v} \rangle^{\perp\perp} \subseteq \langle \mathbf{w} \rangle^{\perp}$ (since inclusions are reversed). ■

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 $b(\mathbf{w},\mathbf{v}) = b(\mathbf{v},\mathbf{w})^{\sigma}$ for all $\mathbf{v},\mathbf{w} \in V$. This implies that, for any $\mathbf{v}$, $b(\mathbf{v},\mathbf{v})$ lies in the fixed field of σ. If σ is the identity, such a form (which is bilinear) is called *symmetric*.

A bilinear form b is called *alternating* if $b(\mathbf{v},\mathbf{v}) = 0$ for all $\mathbf{v} \in V$. This implies that $b(\mathbf{w},\mathbf{v}) = -b(\mathbf{v},\mathbf{w})$ for all $\mathbf{v},\mathbf{w} \in V$. (Expand $b(\mathbf{v}+\mathbf{w},\mathbf{v}+\mathbf{w}) = 0$, and note that two of the four terms are zero.) 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 6.3 *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.*

I will not give the complete proof of this theorem. The next result shows that $\sigma^2 = 1$, and then the proof of the theorem is given in the case of a bilinear form (that is, when $\sigma = 1$).

Proposition 6.4 *If b is a non-zero reflexive σ-sesquilinear form, then σ^2 is the identity.*

Proof Note first that a form is σ-sesquilinear if and only if it is additive in each variable and satisfies

$$b(\alpha\mathbf{v},\mathbf{w}) = \alpha b(\mathbf{v},\mathbf{w}), \qquad b(\mathbf{v},\beta\mathbf{w}) = b(\mathbf{v},\mathbf{w})\beta^\sigma.$$

Step 1 If b is alternating, then $\sigma = 1$. For we can choose $\mathbf{v}$ and $\mathbf{w}$ with $b(\mathbf{v},\mathbf{w}) = -b(\mathbf{w},\mathbf{v}) = 1$. Then for any $\alpha \in F$, we have

$$\begin{aligned} \alpha &= \alpha b(\mathbf{v},\mathbf{w}) \\ &= b(\alpha\mathbf{v},\mathbf{w}) \\ &= -b(\mathbf{w},\alpha\mathbf{v}) \\ &= -b(\mathbf{w},\mathbf{v})\alpha^\sigma \\ &= \alpha^\sigma. \end{aligned}$$

(Note that this step does not require non-degeneracy, merely that b is not identically zero.)

So we can assume that there exists $\mathbf{v}$ with $b(\mathbf{v},\mathbf{v}) \neq 0$. Multiplying b by a non-zero scalar (this does not affect the hypotheses), we may assume that $b(\mathbf{v},\mathbf{v}) = 1$.

Step 2 Assume for a contradiction that $\sigma^2 \neq 1$. For any vector $\mathbf{w}$, if $b(\mathbf{w},\mathbf{v}) \neq 0$, then we can replace $\mathbf{w}$ by its product with a non-zero scalar to assume $b(\mathbf{w},\mathbf{v}) = 1$. Then $b(\mathbf{w}-\mathbf{v},\mathbf{v}) = 0$, and so $b(\mathbf{v},\mathbf{w}-\mathbf{v}) = 0$, whence $b(\mathbf{v},\mathbf{w}) = 1$. We *claim* that $b(\mathbf{w},\mathbf{w}) = 1$.

Proof Suppose that $\alpha = b(\mathbf{w},\mathbf{w}) \neq 1$. Note first that $b(\mathbf{w}-\alpha\mathbf{v},\mathbf{v}) = 0$, and so $b(\mathbf{w},\mathbf{w}-\alpha\mathbf{v}) = 0$, whence $\alpha = \alpha^\sigma$. Take any element $\lambda \in F$ with $\lambda \neq 1$, and choose $\mu \in F$ such that $\mu^\sigma = (1-\lambda)^{-1}(\alpha-\lambda)$. Since $\alpha \neq 1$, we have $\mu \neq 1$; and

$$\mu^\sigma - \lambda\mu^\sigma = \alpha - \lambda.$$

This implies, first, that $\lambda = (\alpha-\mu^\sigma)(1-\mu^\sigma)^{-1}$, and second that

$$b(\mathbf{w}-\lambda\mathbf{v},\mathbf{w}-\mu\mathbf{v}) = \alpha-\lambda-\mu^\sigma+\lambda\mu^\sigma = 0.$$

Hence $b(\mathbf{w}-\mu\mathbf{v},\mathbf{w}-\lambda\mathbf{v}) = 0$, and we obtain

$$\alpha-\mu-\lambda^\sigma+\mu\lambda^\sigma = 0.$$

Applying σ to this equation and using the fact that $\alpha^\sigma = \alpha$, we obtain

$$\alpha-\mu^\sigma-\lambda^{\sigma^2}+\lambda^{\sigma^2}\mu^\sigma = 0,$$

whence

$$\lambda^{\sigma^2} = (\alpha-\mu^\sigma)(1-\mu^\sigma)^{-1} = \lambda.$$

But λ was an arbitrary element different from 1. Since clearly $1^\sigma = 1$, we have $\sigma^2 = 1$, contrary to assumption.

Step 3 Let $W = \mathbf{v}^\perp$. Then $V = \langle\mathbf{v}\rangle \oplus W$, and $\mathrm{rk}(W) \geq 1$. For any $\mathbf{x} \in W$, we have $b(\mathbf{v},\mathbf{v}) = b(\mathbf{v}+\mathbf{x},\mathbf{v}) = 1$, and so by Step 2, we have $b(\mathbf{v}+\mathbf{x},\mathbf{v}+\mathbf{x}) = 1$. Thus $b(\mathbf{x},\mathbf{x}) = -2$. Putting $x = 0$, we see that F must have characteristic 2, and that $b|W$ is alternating. But then Step 1 shows that $b|W$ is identically zero, whence W is contained in the radical of b, contrary to the assumed non-degeneracy.

Proof of Theorem 6.3 We have

$$b(\mathbf{u},\mathbf{v})b(\mathbf{u},\mathbf{w}) - b(\mathbf{u},\mathbf{w})b(\mathbf{u},\mathbf{v}) = 0$$

by commutativity; that is, using bilinearity,

$$b(\mathbf{u}, b(\mathbf{u},\mathbf{v})\mathbf{w} - b(\mathbf{u},\mathbf{w})\mathbf{v}) = 0.$$

By reflexivity,

$$b(b(\mathbf{u},\mathbf{v})\mathbf{w} - b(\mathbf{u},\mathbf{w})\mathbf{v}, \mathbf{u}) = 0,$$

whence bilinearity again gives

$$b(\mathbf{u},\mathbf{v})b(\mathbf{w},\mathbf{u}) = b(\mathbf{u},\mathbf{w})b(\mathbf{v},\mathbf{u}). \tag{6.1}$$

Call a vector $\mathbf{u}$ *good* if $b(\mathbf{u},\mathbf{v}) = b(\mathbf{v},\mathbf{u}) \neq 0$ for some $\mathbf{v}$. By (6.1), if $\mathbf{u}$ is good, then $b(\mathbf{u},\mathbf{w}) = b(\mathbf{w},\mathbf{u})$ for all $\mathbf{w}$. Also, if $\mathbf{u}$ is good and $b(\mathbf{u},\mathbf{v}) \neq 0$, then $\mathbf{v}$ is good. But, given any two non-zero vectors $\mathbf{u}_1,\mathbf{u}_2$, there exists $\mathbf{v}$ with $b(\mathbf{u}_i,\mathbf{v}) \neq 0$ for $i = 1,2$. (For there exist $\mathbf{v}_1,\mathbf{v}_2$ with $b(\mathbf{u}_i,\mathbf{v}_i) \neq 0$ for $i = 1,2$, by non-degeneracy; and at least one of $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_1+\mathbf{v}_2$ has the required property.) So, if some vector is good, then every non-zero vector is good, and b is symmetric.

But, putting $\mathbf{u} = \mathbf{w}$ in (6.1) gives

$$b(\mathbf{u},\mathbf{u})(b(\mathbf{u},\mathbf{v}) - b(\mathbf{v},\mathbf{u})) = 0$$

for all $\mathbf{u},\mathbf{v}$. So, if $\mathbf{u}$ is not good, then $b(\mathbf{u},\mathbf{u}) = 0$; and, if no vector is good, then b is alternating. ∎

In the next few sections, we develop this theme further.

Exercises

1. Let b be a sesquilinear form on V. Define the *left* and *right radicals* of b to be the subsets

$$\{\mathbf{v} \in V : (\forall \mathbf{w} \in V) b(\mathbf{v},\mathbf{w}) = 0\}$$

and

$$\{\mathbf{v} \in V : (\forall \mathbf{w} \in V) b(\mathbf{w},\mathbf{v}) = 0\}$$

respectively. Prove that the left and right radicals are subspaces of the same rank (if V has finite rank).

(Note: If the left and right radicals are equal, this subspace is called the *radical* of b. This holds if b is reflexive.)

2. Give an example of a bilinear form on an infinite-rank vector space whose left radical is zero and whose right radical is non-zero.

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 $\varepsilon = (a+1)+b\omega$.]

4. Use the result of Exercise 3 to complete the proof of Theorem 6.3 in general.

[If $b(\mathbf{u},\mathbf{u}) = 0$ for all $\mathbf{u}$, the form b is alternating and *bilinear.* If not, suppose that $b(\mathbf{u},\mathbf{u}) \neq 0$ and let $b(\mathbf{u},\mathbf{u})^\sigma = \lambda b(\mathbf{u},\mathbf{u})$. Choosing ε as in Exercise 2 and renormalising b, show that we may assume that $\lambda = 1$, and (with this choice) that b is Hermitian.]

6.2 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 $b(\mathbf{v},\mathbf{v}) \in \mathrm{Fix}(\sigma)$ for all $\mathbf{v} \in V$. We call the form b *trace-valued* if $b(\mathbf{v},\mathbf{v}) \in \mathrm{Tr}(\sigma)$ for all $\mathbf{v} \in V$.

Proposition 6.5 *We have* $\mathrm{Tr}(\sigma) = \mathrm{Fix}(\sigma)$ *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 1). 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 Proposition 6.5 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*. These will be discussed briefly in Section 7.6.

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

- $f(\lambda\mathbf{v}) = \lambda^2 f(\mathbf{v})$ for all $\lambda \in F$, $\mathbf{v} \in V$;
- $f(\mathbf{v}+\mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w}) + b(\mathbf{v},\mathbf{w})$, where b is bilinear.

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

$$b(\mathbf{v},\mathbf{w}) = f(\mathbf{v}+\mathbf{w}) - f(\mathbf{v}) - f(\mathbf{w})$$

and

$$f(\mathbf{v}) = \tfrac{1}{2}b(\mathbf{v},\mathbf{v}),$$

the latter equation coming from the substitution $\mathbf{v} = \mathbf{w}$ in the second defining condition. So nothing new is obtained.

On the other hand, if the characteristic of F is 2, then b is an alternating bilinear form, and f 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 is obtained by *polarisation* of f.

Now let b be a symmetric bilinear form over a field of characteristic 2, which is not alternating. Set $f(\mathbf{v}) = b(\mathbf{v},\mathbf{v})$. Then we have

$$f(\lambda\mathbf{v}) = \lambda^2 f(\mathbf{v})$$

and

$$f(\mathbf{v}+\mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w}),$$

since $b(\mathbf{v},\mathbf{w})+b(\mathbf{w},\mathbf{v})=0$. 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 f is *non-singular* if

$$(f(\mathbf{v})=0\ \&\ (\forall\mathbf{w}\in V)b(\mathbf{v},\mathbf{w})=0)\quad\Rightarrow\quad\mathbf{v}=0$$

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

If the characteristic is not 2, then non-singularity is equivalent to non-degeneracy of the bilinear form.

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

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

As above, f is very nearly semilinear. The field F is called *perfect* if every element is a square. In this case, f is indeed semilinear, and its kernel is a hyperplane of W. We conclude:

Theorem 6.6 *Let f 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.

Exercises

1. 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 $\mathrm{Fix}(\sigma)$.

2. Let b be an alternating bilinear form on a vector space V over a field F of characteristic 2. Let $(\mathbf{v}_i : i \in I)$ be a basis for V, and q any function from I to F. Show that there is a unique quadratic form with the properties that $f(\mathbf{v}_i) = q(i)$ for every $i \in I$, and f polarises to b.

3. (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.

4. Show that finite fields of characteristic 2 are perfect. (Hint: the multiplicative group is cyclic of odd order.)

6.3 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 σ-Hermitian form b, where σ is not the identity; or a non-singular quadratic form f. In the third case, we let b be the bilinear form obtained by polarising f; then b is alternating or symmetric according as the characteristic is or is not 2, but b may be degenerate. In the other two cases, we define a function $f : V \to F$ defined by $f(\mathbf{v}) = b(\mathbf{v},\mathbf{v})$ — this is identically zero if b is alternating. See Exercise 1 for the Hermitian case.

We say that V is *anisotropic* if $f(\mathbf{v}) \neq 0$ for all $\mathbf{v} \neq 0$. Also, V is a *hyperbolic line* if it is spanned by vectors $\mathbf{v}$ and $\mathbf{w}$ with $f(\mathbf{v}) = f(\mathbf{w}) = 0$ and $b(\mathbf{v},\mathbf{w}) = 1$. (The vectors $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, so V has rank 2; so, projectively, it is a "line".)

Theorem 6.7 *A space carrying a form of one of the above types 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. (V cannot contain a hyperbolic line.) So suppose that V contains a vector $\mathbf{v} \neq 0$ with $f(\mathbf{v}) = 0$.

We claim that there is a vector $\mathbf{w}$ with $b(\mathbf{v},\mathbf{w}) \neq 0$. 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 $\mathbf{v}$ is in the radical of b; but $\mathbf{v}$ is a singular vector, contradicting the non-singularity of f.

Multiplying $\mathbf{w}$ by a non-zero constant, we may assume that $b(\mathbf{v},\mathbf{w}) = 1$.

Now, for any value of λ, we have $b(\mathbf{v},\mathbf{w}-\lambda\mathbf{v}) = 1$. We wish to choose λ so that $f(\mathbf{w}-\lambda\mathbf{v}) = 0$; then $\mathbf{v}$ and $\mathbf{w}$ will span a hyperbolic line. Now we distinguish cases. If b is alternating, then any value of λ works. If b is Hermitian, we have

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

and, since b is trace-valued, there exists λ with $\mathrm{Tr}(\lambda) = f(\mathbf{w})$. Finally, if f is quadratic, we have

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

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

Now let W_1 be the hyperbolic line $\langle \mathbf{v}, \mathbf{w}-\lambda\mathbf{v}\rangle$, 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 or non-singular, as appropriate. Now the existence of the decomposition follows by induction.

I will omit the proof of uniqueness. ■

The number r of hyperbolic lines is called the *polar rank* or *Witt index* of V. I do not know of a commonly accepted term for U; I will call it the *germ* of V, for reasons which will become clear shortly.

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 6.8 *The only anisotropic space carrying an alternating bilinear form is the zero space.* ■

In combination with Theorem 6.7, this shows that a space carrying a non-degenerate alternating bilinear form is a direct sum of hyperbolic lines.

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 $r+s\leq n$. 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 6.9 *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 6.10 *(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 Exercise 3).

(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 1 of Section 6.2).

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. ■

Exercises

1. Let b be a σ-Hermitian form on a vector space V over F, where σ is not the identity. Set $f(\mathbf{v}) = b(\mathbf{v}, \mathbf{v})$. 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 $\mathrm{Tr}(b)(\mathbf{v}, \mathbf{w}) = b(\mathbf{v}, \mathbf{w}) + b(\mathbf{v}, \mathbf{w})^\sigma$. Show further that $\mathrm{Tr}(b)$ is non-degenerate if and only if b is.

2. 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 lines).

3. 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.

6.4 Classical 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 $f(\mathbf{v}) = b(\mathbf{v}, \mathbf{v})$, 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 *totally isotropic subspace* (or *t.i. subspace*), while a subspace on which f vanishes identically is called a *totally singular subspace* (or *t.s. subspace*). Every t.s. subspace is t.i., but the converse is false. In the case of alternating forms, every subspace is t.s.! I frequently use the expression *t.i. or t.s. subspace*, to mean a t.i. subspace (in the symplectic or unitary case) or a t.s. subspace (in the orthogonal case).

The *classical polar space* (or simply the *polar space*) associated with a vector space carrying a form is the geometry whose flats are the t.i. or t.s. subspaces (in the above sense). (Concerning the terminology: the term “polar space” is normally reserved for a geometry satisfying the axioms of Tits, which we will meet shortly. But every classical polar space is a polar space, so the terminology here should cause no confusion.) 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 t.i. or t.s. 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 polar rank as the number of hyperbolic lines in the decomposition of the space.) We use the terms *point*, *line*, *plane*, etc., just as for projective spaces.

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

First, we identify the two definitions of polar space rank. We use the expression for V as the direct sum of r hyperbolic lines and an anisotropic subspace given by Theorem 6.7. Any t.i. or t.s. subspace meets each hyperbolic line in at most a point, and meets the anisotropic germ in the zero space; so its rank is at most r. But the span of r t.i. or t.s. points, one chosen from each hyperbolic line, is a t.i. or t.s. subspace of rank r.

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

This is clear since a subspace of a t.i. or t.s. subspace is itself t.i. or t.s. The next property is also clear.

(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 lines 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.

Proof Let $p = \langle \mathbf{w} \rangle$. The function $\mathbf{v} \mapsto b(\mathbf{v}, \mathbf{w})$ 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 t.i. or t.s. if and only if $q \in K$; and the union of all such t.i. or t.s. lines is a t.i. or t.s. space $W = \langle K, \mathbf{v} \rangle$, such that $W \cap U = K$, as required.

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

Proof Use the hyperbolic-anisotropic decomposition again. If $L_1, \ldots, L_r$ are the hyperbolic lines, and $\mathbf{v}_i, \mathbf{w}_i$ are the distinguished spanning vectors in L_i, then the required flats are $\langle \mathbf{v}_1, \ldots, \mathbf{v}_r \rangle$ and $\langle \mathbf{w}_1, \ldots, \mathbf{w}_r \rangle$.

Next, we specialise to the case $r = 2$. (A polar space of rank 1 is just an unstructured collection of points.) A polar space of rank 2 consists of points and lines, and has the following properties. (The first two are immediate consequences of (P2) and (P3) respectively.)

(Q1) Two points lie on at most one line.

(Q2) If L is a line, and p a point not on L, then there is a unique point of L collinear with p.

(Q3) No point is collinear with all others.

For, by (P4), there exist disjoint lines; and, given any point p, at least one of these lines does not contain p, and p fails to be collinear with some point of this line.

A geometry satisfying (Q1), (Q2) and (Q3) is called a *generalised quadrangle*. Such geometries play much the same rôle in the theory of polar spaces as projective planes do in the theory of projective spaces. We will return to them later.

Note that (Q1) holds in a polar space of arbitrary rank.

Another property of polar spaces, which is proved by almost the same argument as (P3), is the following extension of (Q2):

(BS) If L is a line, and p a point not on L, then p is collinear with one or all points of L.

In a polar space G, for 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. It follows from (BS) that, for any set S, the set $S^\perp$ is a (linear) subspace of G (that is, if two points of $S^\perp$ are collinear, then the line joining them lies wholly in $S^\perp$). Moreover, for any point x, $x^\perp$ is a hyperplane of G (that is, a subspace which meets every line).

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 line.) 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 and in the next section 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* or *hyperbolic orthogonal space*, the orthogonal space which is a direct sum of hyperbolic lines (that is, having germ 0). The quadratic form defining this space can be taken to be $x_1x_2 + x_3x_4 + \ldots + x_{2r-1}x_{2r}$.

Theorem 6.11 *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 t.s. 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 Fig. 6.1. This is the so-called "ruled quadric", familiar from models such as wastepaper baskets.

Exercises

1. Prove (BS).

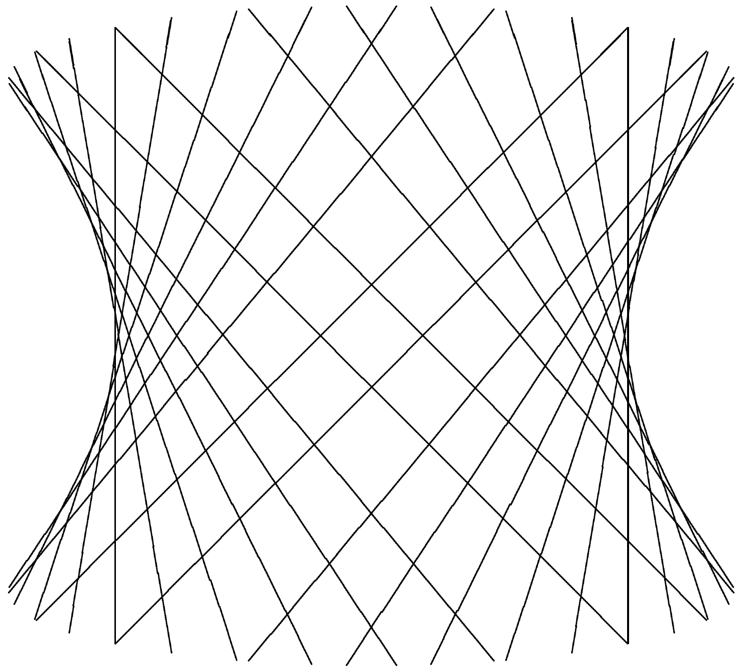

Figure 6.1: A grid

2. Prove the assertions above about $x^{\perp}/x$ and $\{x,y\}^{\perp}$.

3. Show that Theorem 6.11 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.

6.5 Finite polar spaces

The classification of finite classical polar spaces was achieved by Theorem 6.7. 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, n the vector space rank. 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	n	ε
Symplectic	$2r$	0
Unitary	$2r$	$-\frac{1}{2}$
Unitary	$2r+1$	$\frac{1}{2}$
Orthogonal	$2r$	-1
Orthogonal	$2r+1$	0
Orthogonal	$2r+2$	1

Table 6.1: Finite classical polar spaces

Theorem 6.12 *The number of points in a finite polar space of rank 1 is $q^{1+\varepsilon}+1$, where ε is given in Table 6.1.*

Proof Let V be a vector space carrying a form of rank 1 over $\mathrm{GF}(q)$. 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 $\bar{x}=x^\sigma=x^r$, $r^2=q$. So the isotropic points satisfy $x\bar{y}+y\bar{x}=0$, that is, $\mathrm{Tr}(x\bar{y})=0$. 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 6.13 *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.

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

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.

Step 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 Step 2 and correctly reduces to $q^{1+\varepsilon}+1$ when $r=1$. ∎

Theorem 6.14 *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.