An Introduction to Finite Geometry

Simeon Ball and Zsuzsa Weiner

5 September 2011

Contents

Preface

These notes are an updated version of notes that were compiled during part of a graduate course in combinatorics that I gave at the Universitat Politècnica de Catalunya in October 2003. Zsuzsa has corrected and clarified a number of the proofs and we have tried to eliminate typos and the like but inevitably the subtler ones remain unnoticed by us. The notes are intended to be used as course notes and should provide material for about 12 hours of lectures.

Chapter 1 introduces the geometries $PG(n, q)$, the axioms of a projective and affine planes and its highlight is a proof of the Bruck-Ryser-Chowla theorem.

Chapter 2 looks at arcs and maximum distance separable codes. Its highlight is a proof of Segre's theorem on ovals.

Chapter 3 introduces polar geometries, includes the classification of sesquilinear forms, the Birkhoff-von Neumann theorem, and the classification of (the classical) polar spaces.

Chapter 4 contains basic results derived from the axioms of a generalised quadrangle and inversive planes. Its highlight is the construction of the Tits ovoid.

There are many topics in finite geometry that have not been touched upon in these notes, diagram geometries, blocking sets, partial geometries for example. It should not be inferred that the topics covered in these notes are more interesting or relevant than those not covered.

I have taken material from a number of sources all of which will provide further reading and references. The title should be self-explanatory as to which part of the course it relates to. *Combinatorics: topics, techniques, algorithms* by P. J. Cameron [3] is a textbook for undergraduates and the lecture notes *Projective and polar spaces* [4] are for graduates and are available on-line. *The geometry of the classical groups* by D. E. Taylor [17] is out of print, but well worth trying to get hold of. The books by J. W. P. Hirschfeld [9], [8], and together with J. A. Thas [10], provide a welter of information enough to feed any appetite. *Finite generalised quadrangles* by S. E. Payne and J. A. Thas [14] is out of print but there are rumours that a LaTeXversion may be available soon. The classic reference *Finite geometries* by P. Dembowski [7] was reprinted in 1997 but contains no proofs. *Projective planes* by D. R. Hughes and F. Piper [11] is out of print but a very interesting read and should be in most libraries. *The algebraic theory of spinors and Clifford algebras* by C. Chevalley [5] should provide interesting further reading to Chapter 3.

April 2007.

I have changed the proof of Segre's theorem (Theorem 2.2.1) so that it demonstrates the ideas used in [1] to prove the MDS conjecture over prime fields.

Simeon Ball, Barcelona, September 2011.

Chapter 1

Projective geometries

1.1 Finite fields

A *field* is a set K with two operations, usually called addition and multiplication, with the property that K is an additive group with identity 0 and $K \setminus \{0\}$ is a multiplicative group.

Theorem 1.1.1 (Galois). *A finite field has q elements, where q is the power of a prime. The field of order q is unique.*

We denote the finite field of order q as $GF(q)$, although it is also denoted $\mathbb{F}_q$ by many. We will need the following properties and definitions relating to finite fields. The details of the following facts can be found in Lidl and Niederreiter [13].

(i) For all $x \in GF(q)$ we have $x^q = x$.

(ii) Let p be a prime. The field $GF(p)$ consists of the set Z/pZ where addition and multiplication are defined modulo p.

(iii) The finite field $GF(p^h)$ can be constructed in the following way. Let $f \in GF(p)[x]$ be a polynomial of degree h, irreducible over $GF(p)$. The quotient ring $GF(p)[x]/(f(x))$ has p^h elements and with the multiplication and addition defined as in this quotient ring, it is the field $GF(p^h)$.

(iv) The prime p is called the *characteristic* of the field.

(v) An element ϵ is called *primitive* if $\{\epsilon^i \mid i = 0, 1, \ldots, q-2\} = GF(q) \setminus \{0\}$. The multiplicative group $GF(q) \setminus \{0\}$ is usually denoted $GF(q)^*$ and is cyclic.

(vi) $GF(p^r)$ is a subfield of $GF(p^h)$ if and only if r divides h.

(vii) $GF(p^h)$ is a vector space of rank h over $GF(p)$.

1.2 Projective spaces

Let $V(n+1,q)$ be a vector space of rank $n+1$ over $GF(q)$. The projective space $PG(n,q)$ is the geometry whose points, lines, planes, $\ldots$, hyperplanes are the subspaces of $V(n+1,q)$ of rank 1, 2, 3, $\ldots$, n. The dimension of a subspace of $PG(n,q)$ is one less than the rank of a subspace of $V(n+1,q)$.

The incidence structure in Figure 1.1 is what we get if we put $n=q=2$. It has a group of 168 automorphisms, isomorphic to the 3×3 non-singular matrices whose elements come from $GF(2)$.

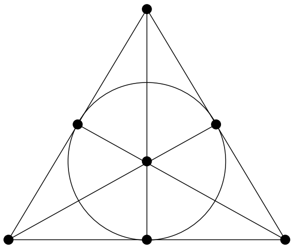

Figure 1.1: $PG(2,2)$ $\ldots$ The Fano plane

As in linear algebra $\langle (x_0,x_1,\ldots,x_n),(y_0,y_1,\ldots,y_n),\ldots,(z_0,z_1,\ldots,z_n)\rangle$ is the space spanned by the vectors $\mathbf{x}=(x_0,x_1,\ldots,x_n),\mathbf{y}=(y_0,y_1,\ldots,y_n)$ and $\mathbf{z}=(z_0,z_1,\ldots,z_n)$.

A *hyperplane* is a subspace of co-dimension 1. If H a hyperplane and l is a line not contained in H then $H\cap l$ is a point.

The geometry $PG(2,q)$ has the property that every two lines are incident in a (unique) point. The rank of the vector space $V(3,q)$ is 3 and the lines U and V are subspaces of rank 2. Hence the rank of $U\cap V$ is 1, so $U\cap V$ is a point.

Proposition 1.2.1. *The number of subspaces of rank k in $V(n,q)$ is*

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \frac{(q^n-1)(q^n-q)\ldots(q^n-q^{k-1})}{(q^k-1)(q^k-q)\ldots(q^k-q^{k-1})}.$$

Proof. The number of k-tuples of linearly independent vectors in a vector space of rank n is

$$(q^n-1)(q^n-q)\ldots(q^n-q^{k-1}).$$

The number of subspaces of rank k is the number of k-tuples of linearly independent vectors in $V(n,q)$ divided by the number of k-tuples of linearly independent vectors in $V(k,q)$. □

The next proposition follows similarly.

Proposition 1.2.2. *The number of subspaces of rank k through a given subspace of rank $d \leq k$ in $V(n,q)$ is* $\begin{bmatrix} n-d \\ k-d \end{bmatrix}_q$.

1.3 Desargues' theorem

We say that the triangles ABC and $A'B'C'$ of $PG(n,q)$ are *in perspective* if the lines AA', BB' and CC' are concurrent.

Theorem 1.3.1. *Assume that in $PG(n,q)$, ABC and $A'B'C'$ are two triangles in perspective. Let AB be the intersection point of the lines $\langle A,B\rangle$ and $\langle A',B'\rangle$ and define the points AC and BC similarly (see Figure 1.2). Then the points AB, AC and BC are collinear.*

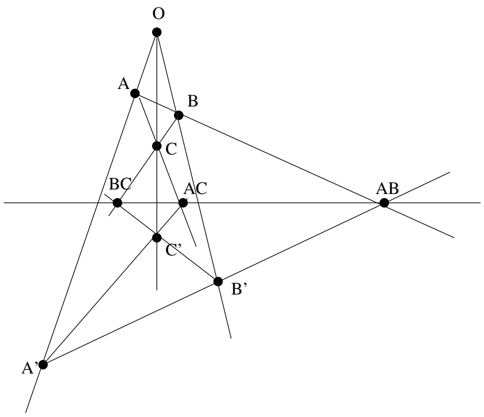

Figure 1.2: Desargues' configuration

Proof. The points AB, AC and AB lie in the planes $\langle A,B,C\rangle$ and $\langle A',B',C'\rangle$. If Figure 1.2 is not contained in a plane then these planes are distinct planes both contained in the 3-space $\langle O,A,B,C\rangle$, and so their intersection is a line.

If Figure 1.2 is contained in a plane π then for every hyperplane H containing the line $\langle AB,BC\rangle$, we can choose two points D and D' (on a line through O), not in π, such that AD (the intersection point of the lines $\langle A,D\rangle$ and $\langle A',D'\rangle$) is in H. To do so, first we choose the point AD, then we find D and D'. Let ℓ be a line through O, not in π and let f be the intersection line of the plane $\langle\langle A,A'\rangle,\ell\rangle$ and H. For the point AD, choose any point on $f \setminus (\langle A',A'\rangle \cup \ell)$. Define D to be the intersection point of the lines $\langle A,AD\rangle$ and ℓ, and D' to be the intersection point of the lines $\langle A',AD\rangle$ and ℓ.

Similarly to AD, define the points BD and CD. The argument of the previous paragraph implies that AD, AB and BD are collinear and, since AB and AD are in H, we have that BD is in H. Now again from the previous paragraph BD, BC and CD are collinear and hence CD is in H. And finally since AD, AC and CD are collinear, we have that AC is in H. Now AC is in every hyperplane containing the line $\langle AB, BC\rangle$ so it must be incident with the line $\langle AB, BC\rangle$. □

1.4 Projective planes

In this section we give axioms of an incidence structure that mimics those properties that $PG(2, q)$ has. This is a common occurrence in the study of incidence structures. We will often take a naturally occurring object and define a more general object with properties it possesses.

A *projective plane* is an incidence structure of points and lines with the following properties.

(PP1) Every two points are incident with a unique line.

(PP2) Every two lines are incident with a unique point.

(PP3) There are four points, no three collinear.

Note that the axioms (PP1)-(PP3) are self-dual. Hence the dual of a projective plane is also a projective plane. So if we prove a theorem for points in a projective plane then the dual result holds automatically for lines.

We have already seen that the geometry $PG(2, q)$ is an incidence structure satisfying these properties. It is called the *Desarguesian projective plane* because of the following theorem, a partial proof of which can be found in [4].

Theorem 1.4.1. *If π is a projective plane with the property that for every pair of triangles ABC and $A'B'C'$ in perspective the points AB, AC and BC, defined as in Figure 1.2, are collinear then π is $PG(2, q)$ for some q.*

Proposition 1.4.2. *Every point in a projective plane is incident with a constant $n + 1$ lines. Dually, every line is incident with $n + 1$ points.*

Proof. Let P be a point not incident with a line l. By (PP1) and (PP2) the number of points incident with l is equal to the number of lines incident with P. By (PP3) there is a point $Q \neq P$, that is not incident with l. The number of lines incident with Q is equal to the number of points incident with l which is equal to the number of lines incident with P. The points P and Q were chosen arbitrarily so every point is incident with a constant number of lines. □

The *order* of a projective plane is one less then the number of points incident with a line.

Proposition 1.4.3. *A projective plane of order n has n^2+n+1 points and n^2+n+1 lines.*

Proof. Let P be a point of a projective plane. There are $n+1$ lines incident with P and each is incident with n other points. Hence the number of points in a projective plane of order n is $n(n+1)+1$. The number of lines follows from the dual argument. □

A projective line over $GF(q)$ has $\begin{bmatrix} 2 \\ 1 \end{bmatrix}_q = (q^2-1)/(q-1) = q+1$ points. Hence the projective plane $PG(2,q)$ has order q and so there are examples of projective planes of order n for every prime power n. There are many projective planes known that are not isomorphic to $PG(2,q)$, however all known examples have prime power order. This brings us to the first, and probably the most famous, of the many unsolved problems in finite geometry.

Conjecture 1.4.4. *The order of a projective plane is the power of a prime.*

Finally, we note that similarly to defining projective planes, one could define projective higher dimensional spaces (mimicking $PG(n,q)$), but it turns out that every such "abstract" projective space is isomorphic to $PG(n,q)$ for some n.

1.5 The Bruck-Ryser-Chowla theorem

The aim of this section is to prove the following theorem.

Theorem 1.5.1. *If there is a projective plane of order n and $n = 1$ or 2 mod 4 then n is the sum of two squares.*

Before we go ahead and prove this theorem let us look at the possible small orders. The numbers 1 and 2 mod 4 are $2,5,6,9,10,13,14,\ldots$. Now $2 = 1^2+1^2$, $5 = 1^2+2^2$, $9 = 0^2+3^2$, $13 = 2^2+3^2$ and there are projective planes of these orders as we have seen. However the theorem implies that there is no projective plane of order 6, nor 14. We shall see that the non-existence of projective planes of order 6 also follows from Euler's proof that there are no two mutually orthogonal latin squares of order 6.

Lam, Thiel and Swiercz [12] concluded, with the aid of a computer, that is there no projective plane of order 10. The smallest possible counter-example to Conjecture 1.4.4 is therefore a plane of order 12.

The proof detailed below is from [3]; the odd explanation has been added.

We will need some lemmas from number theory.

LEMMA 1.5.2. *We have the following identities.*

$$(a_1^2+a_2^2)(x_1^2+x_2^2) = (a_1x_1-a_2x_2)^2+(a_1x_2+a_2x_1)^2.$$

and

$$(a_1^2+a_2^2+a_3^2+a_4^2)(x_1^2+x_2^2+x_3^2+x_4^2) = y_1^2+y_2^2+y_3^2+y_4^2$$

where

$$y_1 = a_1x_1-a_2x_2-a_3x_3-a_4x_4,$$

$$y_2 = a_1x_2+a_2x_1+a_3x_4-a_4x_3,$$

$$y_3 = a_1x_3+a_3x_1+a_4x_2-a_2x_4,$$

$$y_4 = a_1x_4+a_4x_1+a_2x_3-a_3x_2.$$

Proof. The identities hold by direct calculation. □

LEMMA 1.5.3. *Let p be a prime. If there are two integers such that*

$$x_1^2+x_2^2 = 0 \pmod p$$

then p is the sum of two squares.

Proof. Choose r minimal such that there exist x_i satisfying

$$x_1^2+x_2^2 = rp.$$

Assume $r > 1$, otherwise we are finished, and we shall obtain a contradiction. Put $u_1 = x_1 \bmod r$ and $u_2 = -x_2 \bmod r$ such that $|u_i| \le r/2$. Now

$$u_1^2+u_2^2 = x_1^2+x_2^2 = 0 \pmod r$$

so $u_1^2+u_2^2 = rs$ for some $s < r$. Moreover

$$(u_1^2+u_2^2)(x_1^2+x_2^2) = r^2ps = (u_1x_1-u_2x_2)^2+(u_1x_2+u_2x_1)^2,$$

and

$$u_1x_1-u_2x_2 = x_1^2+x_2^2 = 0 \pmod r$$

and

$$u_1x_2+u_2x_1 = x_1x_2-x_1x_2 = 0 \pmod r).$$

Hence

$$ps = ((u_1x_1-u_2x_2)/r)^2+((u_1x_2+u_2x_1)/r)^2,$$

which, since $s < r$, contradicts the minimality of r. □

Lemma 1.5.4. *Let p be a prime. If there are four integers such that*

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 0 \pmod{p}$$

then p is the sum of four squares.

Proof. The proof is as in Lemma 1.5.3. Choose r minimal such that there exist x_i satisfying

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = rp.$$

Assume $r > 1$, put $u_1 = x_1 \bmod r$, $u_2 = -x_2 \bmod r$, $u_3 = -x_3 \bmod r$ and $u_4 = -x_4 \bmod r$ such that $|u_i| \leq r/2$ and use the sum of four squares identity from Lemma 1.5.2 to get a contradiction. □

Lemma 1.5.5. *Every number is the sum of four squares.*

Proof. By Lemma 1.5.2 it suffices to prove the supposition for prime numbers. Now $2 = 1^2 + 1^2 + 0^2 + 0^2$ so let us assume that p is an odd prime.

If -1 is a quadratic residue modulo p (i.e. there exists an x such that $x^2 = m \bmod p$) then there exists an x such that

$$x^2 + 1 = 0 \pmod{p}$$

and Lemma 1.5.3 implies that p is the sum of two squares.

If not then let m be the smallest quadratic non-residue. Since -1 and m are quadratic non-residues $-m$ is a quadratic residue and since m is minimal $m - 1$ is also a quadratic residue. Hence there exist x and y such that

$$x^2 = -m \pmod{p}$$

and

$$y^2 = m - 1 \pmod{p}.$$

Now $x^2 + y^2 + 1 = 0 \bmod p$ and by Lemma 1.5.4 p is the sum of four squares. □

Lemma 1.5.6. *If $nx^2 = w^2 + y^2$ has a solution in integers then n is the sum of two squares.*

Proof. Assume first that n can be written as $n = p_1 p_2 \dots p_t$ where p_i are distinct primes. Now

$$w^2 + y^2 = 0 \pmod{p_i}$$

and so Lemma 1.5.3 implies p_i is the sum of two squares. But then it follows from Lemma 1.5.2 that n is the sum of two squares.

If not then $n = m^2 r$ where r is the product of distinct primes. Now

$$r(mx)^2 = w^2 + y^2,$$

and we have just seen that this implies we can write $r = a^2 + b^2$ for some a and b. Thus $n = (ma)^2 + (mb)^2$. □

We are now ready to prove Theorem 1.5.1.

Proof. (of Theorem 1.5.1.) Let $\{P_i \mid i = 1, \ldots, N\}$ be the points of a projective plane of order n and let $\{l_i \mid i = 1, \ldots, N\}$ be the lines, $N = n^2 + n + 1$. Let $A = (a_{ij})$ be the matrix defined by

$$a_{ij} = \begin{cases} 1 & \text{if } P_i \in l_j, \\ 0 & \text{if } P_i \notin l_j. \end{cases}$$

The axioms (PP1) and (PP2) imply that

$$A^T A = J + nI,$$

where A^T is the transpose of the matrix A and J is the all one matrix. Let $\mathbf{z} = A\mathbf{x}$ where $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ and the x_i's are indeterminates. Then we have that

$$\mathbf{z}^T\mathbf{z} = \mathbf{x}^T A^T A\mathbf{x} = \mathbf{x}^T J\mathbf{x} + n\mathbf{x}^T\mathbf{x},$$

and so

$$z_1^2 + z_2^2 + \ldots + z_N^2 = w^2 + n(x_1^2 + x_2^2 + \ldots + x_N^2),$$

where $w = x_1 + x_2 + \ldots + x_N$. Now add nx_{N+1}^2 to both sides of this equation.

$$z_1^2 + z_2^2 + \ldots + z_N^2 + nx_{N+1}^2 = w^2 + n(x_1^2 + x_2^2 + \ldots + x_N^2 + x_{N+1}^2). \tag{1.1}$$

Note that since $n = 1$ or $2 \bmod 4$, $N+1 = 0 \bmod 4$. By Lemma 1.5.5, n is the sum of four squares and so by Lemma 1.5.2 there exist y_j that are linear combinations of the $\{x_i \mid i = 1, \ldots, N\}$ such that

$$z_1^2 + z_2^2 + \ldots + z_N^2 + nx_{N+1}^2 = w^2 + y_1^2 + y_2^2 + \ldots + y_N^2 + y_{N+1}^2. \tag{1.2}$$

Note that, by definition, the z_j are also linear combinations of the $\{x_i \mid i = 1, \ldots, N\}$. Now at least one of the z_j's and at least one of the y_j's must be linear combinations of the x_j's that include x_1 (since $\mathbf{z} = A\mathbf{x}$ and there are $n+1$ 1's in the first row of A) so let us assume without loss of generality that z_1 and y_1 do. If we put $z_1 = y_1$ then we can solve for x_1 unless the coefficient of x_1 in y_1 and z_1 is the same. If this is the case put $z_1 = -y_1$ and solve for x_1. Now when we substitute this solution into (1.2) the z_1^2 term will cancel with the y_1^2 term.

We can repeat this process with x_2 and continue reducing the number of indeterminates, unless at some stage when we make a substitution one (or more) of the x_j's, say x_l, no longer appears in any of the remaining z_j's or y_j's. Let x_l be one of the four $x_k, x_{k+1}, x_{k+2}, x_{k+3}$ for which by Lemma 1.5.2 $(y_k\ y_{k+1}\ y_{k+2}\ y_{k+3}) = (x_k\ x_{k+1}\ x_{k+2}\ x_{k+3})B$, for some matrix B determined by n. If, after substituting x_j for $j < l$ with linear combinations of x_j's with $j \geq l$, there are no $y_j's$ in which x_l occurs, then putting all $x_j = 0$ for $j > l$ we have that $(y_k\ y_{k+1}\ y_{k+2}\ y_{k+3}) = 0$ and so B is singular. This is not the case since its determinant is $(a_1^2+a_2^2+a_3^3+a_4^2)^2 = n^4$.

Therefore we can continue reducing the number of indeterminates in (1.2) until we have

$$nx_{N+1}^2 = w^2 + y_{N+1}^2,$$

where w and y_{N+1} are rational number multiples of x_{N+1}. Now choose x_{N+1} so that this equation has integer solutions. Lemma 1.5.6 implies that n is the sum of two squares. □

1.6 Affine spaces

The affine space $AG(n, q)$ is the geometry whose points, lines, planes, . . . , hyperplanes, are the cosets of the subspaces of $V(n, q)$ of rank 0, 1, 2, . . . , $n - 1$. The dimension of a subspace of $AG(n, q)$ is the rank of a subspace of $V(n, q)$.

The incidence structure Figure 1.3 is what we get if we put $n = 2$ and $q = 3$.

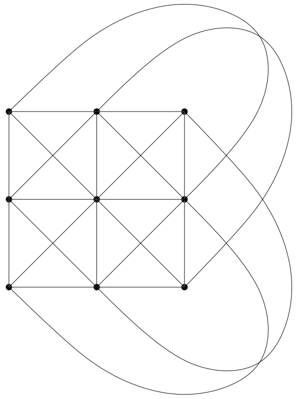

Figure 1.3: $AG(2, 3)$

1.7 Affine planes

An *affine plane* is an incidence structure of points and lines with the following properties.

(AP1) Every two points are incident with a unique line.

(AP2) Given a point P and a line l such that $P \notin l$ then there exists a unique line m such that $P \in m$ and $m \cap l = \emptyset$.

(PP3) There are three points that are not collinear.

It is a simple matter to check that $AG(2, q)$ is an example of an affine plane.

If m and l are lines of an affine plane such that $m \cap l = \emptyset$ then we say that m and l are parallel. If m and l are parallel and l and r are parallel then m and r are parallel, for if not then there is a point $P \in m \cap r$ such that $P \notin l$ which contradicts (AP2). So parallelism is an equivalence relation.

Let $\mathcal{P}$ be the set of points of an affine plane, let $\mathcal{L}$ be the set of lines and let $\mathcal{E}$ be the set of equivalence classes of parallel lines. Each line l belongs to an equivalence class, say $E \in \mathcal{E}$. Define a new line $l^+ = l \cup \{E\}$. The incidence structure whose points are $\mathcal{P} \cup \mathcal{E}$ and whose lines are $\{l^+ \mid l \in \mathcal{L}\}$ and the line at infinity $l_\infty = \{E \mid E \in \mathcal{E}\}$ is a projective plane.

On the other hand if we delete a line and all the points incident with that line from a projective plane then the remaining structure is an affine plane. The deleted line is often called the "line at infinity". It is interesting to note that deleting different lines from a projective plane can yield non isomorphic affine planes.

If we complete $AG(2, 3)$ to a projective plane we get $PG(2, 3)$, as seen in Figure 1.4.

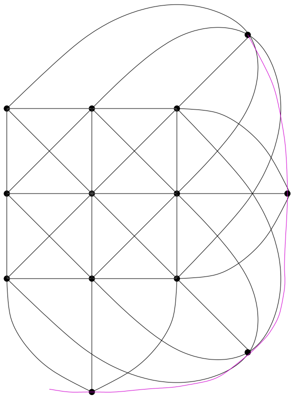

Figure 1.4: $PG(2, 3)$

Proposition 1.7.1. *In an affine plane every line is incident with a constant n points and every point is incident with $n + 1$ lines.*

Proof. Complete the affine plane to projective plane. Then the proof follows immediately from Lemma 1.4.2. □

We define the integer n to be the order of an affine plane.

Proposition 1.7.2. *An affine plane of order n has n^2 points and $n^2 + n$ lines.*

Proof. This follows immediately from Lemma 1.4.3. □

1.8 Mutually orthogonal latin squares

A *latin square of order* n is an $n \times n$ matrix with entries from the set $\{1, 2, \ldots, n\}$ with the property that every element of $\{1, 2, \ldots, n\}$ appears exactly once in each row and column.

A pair of latin squares $A = (a_{ij})$ and $B = (b_{ij})$ are called *orthogonal* if for all (k, l) there exist a unique (i, j) such that $a_{ij} = k$ and $b_{ij} = l$.

E.g.: the matrices

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

are orthogonal latin squares.

Let $N(n)$ denote the maximum number of mutually orthogonal latin squares of order n.

Proposition 1.8.1. *There are at most $n-1$ mutually orthogonal latin squares of order n, so $N(n) \leq n-1$.*

Proof. Note that after permuting the set $\{1, 2, \ldots, n\}$, in each latin square individually, the $n-1$ mutually orthogonal latin squares will still be mutually orthogonal. Hence we can suppose that the 1 is the entry $(1, 1)$ in each of the mutually orthogonal latin squares. In each latin square the $(n-1)^2$ entries that are not in the first row or the first column contain $n-1$ 1's and 1 does not occur in the same cell in two different matrices since the entry $(1, 1)$ is 1 in all the latin squares, by assumption. Hence $N(n)(n-1) \leq (n-1)^2$. □

Given a set of $n-1$ mutually orthogonal latin squares $A_1, A_2 \ldots A_n$, we can construct an affine plane of order n in the following way. Let the set $\{(i, j) \mid i, j = 1, 2, \ldots, n\}$ be the points, the sets

$$\{(x, j) \mid x = 1, 2, \ldots, n\} \text{ where } j = 1, 2, \ldots, n$$

be n "horizontal" lines, the sets

$$\{(j, x) \mid x = 1, 2, \ldots, n\} \text{ where } j = 1, 2, \ldots, n$$

be n "vertical" lines and for each A_m we define for each $k = 1, 2, \ldots, n$ a line

$$\{(i, j) \mid (A_m)_{ij} = k\}.$$

On the other hand, given an affine plane of order n, we can construct $n-1$ mutually orthogonal latin squares by fixing two parallel classes as the horizontal lines and the vertical lines, coordinatising the points with respect to the horizontal and vertical lines and following the above construction in reverse.

Since we know that there are affine planes of order n whenever n is the power of a prime, in these cases we can attain the bound in Proposition 1.8.1. However, if n is not the power of prime, $f(n)$ would be an indication of how close we can get to constructing an affine (and hence a projective) plane. As we have seen as a consequence of Theorem 1.5.1 there is no affine plane of order 6. In fact there are not even 2 orthogonal latin squares of order 6, which was conjectured and partially proved by Euler. There are however two orthogonal latin squares of order 10; it is not known if there are three mutually orthogonal latin squares of order 10.

Given mutually orthogonal latin squares of order r and s we can construct mutually orthogonal latin squares of order rs in the following way. Consider a latin square of order r whose entries come from an r-element set G. Define a multiplication $*$ on G by the rule $g_i * g_j = g_k$ if g_k is the (i,j) entry in the latin square. Then two latin squares $(G,*)$ and $(G,\circ)$ are orthogonal if for all g_k, $g_l \in G$ there exists a unique pair (i,j) such that $g_i * g_j = g_k$ and $g_i \circ g_j = g_l$. If $(H,*)$ and $(H,\circ)$ are orthogonal latin squares of order s then the latin squares $(G\times H,*)$ and $(G\times H,\circ)$ are orthogonal, where $(G\times H,*)$ is defined by the rule

$$(g_1,h_1)*(g_2,h_2)=(g_1*g_2,h_1*h_2).$$

Proposition 1.8.2. *If $n = p_1^{a_1}\dots p_r^{a_r}$ where p_i are distinct primes and $a_i > 0$ then $N(n) \geq q-1$ where q is the smallest $p_i^{a_i}$.*

Proof. We can construct $p_i^{a_i}-1$ mutually orthogonal latin squares of order $p_i^{a_i}$ from $AG(2,p_i^{a_i})$. □

Corollary 1.8.3. *If $n \neq 2$ modulo 4 then there exist at least 2 orthogonal latin squares of order n.*

1.9 The groups $GL(n,q)$ and $PSL(n,q)$.

The general linear group $GL(n,q)$ is the group of non-singular linear transformations of $V(n,q)$. It is isomorphic to the multiplicative group of $(n\times n)$ non-singular matrices whose entries come from $GF(q)$. The order of $GL(n,q)$ is

$$(q^n-1)(q^n-q)\dots(q^n-q^{n-1}).$$

The group of automorphisms $Aut(GF(q))$ of $GF(q)$, where $q=p^h$ and p is prime, is generated by the map

$$x \mapsto x^p$$

and has order h.

We say that two objects $\mathcal{S}$ and $\mathcal{S}'$ in $PG(n-1,q)$ are (projectively) equivalent if there exists an $A \in GL(n,q)$ and a $\sigma \in Aut(GF(q))$ such that

$$\mathcal{S}' = \{\langle A\mathbf{x}^\sigma\rangle \mid \mathbf{x}\in\mathcal{S}\}.$$

The group $GL(n, q)$ acts 2-transitively on the points of $PG(n-1, q)$, however it is not 3-transitive (a set of three collinear points is not equivalent to three non-collinear points).

As the next theorem shows, the above transformations are the only collineations of $PG(n, q)$. We omit the proof.

THEOREM 1.9.1. *Any collineation of $PG(n, q)$, $n \geq 2$, can be represented by*

$$\mathbf{x} \mapsto \mathbf{x}^{\sigma} A,$$

where $\sigma \in Aut(GF(q))$ and $A \in GL(n, q)$.

It is not difficult to show that there exists a unique linear transformation that maps a given ordered $(n+2)$-tuple of points in general position (no $(n+1)$ of them are on a hyperplane) to another given ordered $(n+2)$-tuple of points in general position. This means that if there are given two coordinate systems in $PG(n, q)$ (with their base points), then we can switch from one to the other by a linear transformation.

The set

$$\{A \in GL(n, q) \mid \det(A) = 1\}$$

is a subgroup of $GL(n, q)$ and is denoted $SL(n, q)$. The group $PSL(n, q)$ is the group of collineations induced by $SL(n, q)$ on $PG(n-1, q)$ and is isomorphic to

$$SL(n, q)/(SL(n, q) \cap Z(n, q))$$

where $Z(n, q) = \{\lambda I \mid \lambda \in GF(q)^*\}$.

The group $PSL(n, q)$ is a simple group (it has no non-trivial normal subgroups) unless $(n, q) = (2, 2)$ or $(n, q) = (2, 3)$.

1.10 Exercises

1. Check that the polynomials $x^3 + x + 1$ and $x^3 + x^2 + 1$ are both irreducible over $GF(2)$. Write out the multiplication table for the elements in the quotient rings $GF(2)[x]/(x^3 + x + 1)$ and $GF(2)[x]/(x^3 + x^2 + 1)$. What is the isomorphism between these tables?

2. Prove Desargues' Theorem using coordinates.

3. *A spread of* $PG(3, q)$ is a set of lines with the property that every point of $PG(3, q)$ is incident with a unique line of the spread. Prove that a spread of $PG(3, q)$ is a set of $q^2 + 1$ lines.

Let V be a vector space of rank 2 over $GF(q^2)$. Using the fact that $GF(q^2)$ can be viewed as a vector space of rank 2 over $GF(q)$ construct a spread of $PG(3, q)$.

4. Let q be the order of a projective plane π. A (k, r)-*arc* $\mathcal{K}$ of π is a set of k points with the property that every line of π is incident with at most $r \leq q$ points of $\mathcal{K}$. Prove that

$$k \leq rq - q + r$$

and that if r does not divide q then

$$k \leq rq - q + r - 1.$$

An $(rq - q + r, r)$-arc is called a *maximal arc* of *degree* r.

Chapter 2

Arcs and maximum distance separable codes

2.1 Ovals

A *conic* is a set of points of $PG(2,q)$ that are zeros of a non-degenerate homogeneous quadratic form (in 3 variables), for example, $f(\mathbf{x}) = x^2 - yz$. All conics in $PG(2,q)$ are equivalent, as we shall see in Section 3.8, so we can deduce the properties of a conic from the conic

$$\mathcal{C} = \{\langle (x,y,z) \rangle \mid x^2 = yz\}.$$

There is just one point of $\mathcal{C}$ that has $z = 0$ and that is $\langle (0,1,0) \rangle$. If we put $z = 1$ we see that there are q other points in $\mathcal{C}$, $\{\langle (t, t^2, 1) \rangle \mid t \in GF(q)\}$. Every line is incident with at most two points of $\mathcal{C}$. Let $P \in \mathcal{C}$. Of the $q+1$ lines that are incident with P there are q incident with one other point of $\mathcal{C}$ and one that is incident with 1 point of $\mathcal{C}$.

An *oval* $\mathcal{O}$ is a set of $q+1$ points in a projective plane of order q, with the property that every line is incident with at most two points of $\mathcal{O}$. It is immediate that through each point P of $\mathcal{O}$, there is exactly one line whose intersection with $\mathcal{O}$ is just the point P. Such a line is called a *tangent*, hence there are $q+1$ tangents to $\mathcal{O}$. Lines meeting $\mathcal{O}$ in two points are called *secants*.

A conic is an example of an oval in $PG(2,q)$. The following proposition shows that when q is even and large enough the conics are not the only examples.

Proposition 2.1.1. *The set*

$$\mathcal{A} = \{\langle (1, t, t^{2^i}) \rangle \mid t \in GF(2^h)\} \cup \{\langle (0,0,1) \rangle\},$$

is an oval in $PG(2, 2^h)$ *if and only if* $(i,h) = 1$.

Proof. Every line incident with $\langle(0,0,1)\rangle$ of the form $y = ax$ is incident with one other point of $\mathcal{A}$, namely $\langle(1,a,a^{2^i})\rangle$. The line $x = 0$ is clearly a tangent of $\mathcal{A}$ at $\{\langle(0,0,1)\rangle\}$.

The other lines are of the form $z = ax + by$, so let us consider their intersection with $\mathcal{A}$. The line $z = ax + by$ is incident with $\langle(1,t,t^{2^i})\rangle$ whenever $t^{2^i} = bt + a$. If $u^{2^i} = bu + a$ and $v^{2^i} = bv + a$ then $u^{2^i} - v^{2^i} = b(u - v)$. Now $u^{2^i} - v^{2^i} = (u - v)^{2^i}$ since all the binomial coefficients in the expansion of $(u - v)^{2^i}$ are even and hence zero in the field $GF(2^h)$. Thus $(u - v)^{2^i} = (u - v)b$ and so if $u \neq v$ then $b = (u - v)^{2^i - 1}$. The hypothesis $(i,h) = 1$ implies that $(2^i - 1, 2^h - 1) = 1$ and so there are integers m and n such that $m(2^i - 1) + n(q - 1) = 1$. Hence $b^m = u - v$ and so there are at most two points of $\mathcal{A}$ incident with the line $z = ax + by$. □

The set $\mathcal{A}$ is called a *translation oval.*

Let us do some simple counting arguments relating to ovals.

Proposition 2.1.2. *Let $\mathcal{O}$ be an oval in a projective plane or order q, q odd. Every point that is not a point of $\mathcal{O}$ is incident with zero or two tangents to $\mathcal{O}$.*

Proof. Let x_i be the number of points that are incident with i tangents to $\mathcal{O}$. We count in two ways the number of pairs (P,l) where l is a tangent to $\mathcal{O}$ and $P \in l \setminus \mathcal{O}$,

$$\sum ix_i = q(q+1).$$

Now count triples (P,l,m) where l and m are distinct tangents to $\mathcal{O}$ and $P \in l \cap m$,

$$\sum i(i-1)x_i = q(q+1).$$

Hence

$$\sum i(i-2)x_i = 0.$$

Now $q + 1$ is even, so the number of points in $\mathcal{O}$ and the number of lines incident with a point are both even. So counting the points of $\mathcal{O}$ on lines incident with a point $P \notin \mathcal{O}$, we have that P is incident with an even number of tangents. Therefore $x_i = 0$ if i is odd and every term in the sum in non-negative. We conclude that $x_i = 0$ unless $i = 0$ or 2. □

Proposition 2.1.3. *Let $\mathcal{O}$ be an oval in a projective plane of order q, q even. Every point that is not a point of $\mathcal{O}$ is incident with one or $q + 1$ tangents to $\mathcal{O}$.*

Proof. This is Exercise 8, the solution of which can be found in Chapter 5.1. □

The above proposition implies that the tangents are incident with a common point. This common point is called the *nucleus* of the oval. (See Exercise 8.) The set of $q + 2$ points that we get if we add the nucleus to the oval is called a *hyperoval*; it is a maximal arc of degree 2. When q is odd no such nucleus exists.

2.2 Segre's theorem

This section contains a proof of Segre's theorem, which may be more complicated than other proofs of this theorem, see [15] for the original proof, see [3] for an alternative proof, but which has the advantage of demonstrating the ideas used to prove Theorem 2.5.2 in [1].

THEOREM 2.2.1. *An oval in $PG(2,q)$, q odd, is a conic.*

Proof. Let $\langle x\rangle$, $\langle y\rangle$ and $\langle z\rangle$ be three points of and oval $\mathcal{O}$ of $PG(2,q)$. With respect to the basis $\{x,y,z\}$ let the tangents at these points be $\alpha_{21}X_2+\alpha_{31}X_3=0$, $\alpha_{12}X_1+\alpha_{32}X_3=0$ and $\alpha_{13}X_1+\alpha_{23}X_2=0$ respectively.

Let $\langle s\rangle\in\mathcal{O}\setminus\{\langle x\rangle,\langle y\rangle,\langle z\rangle\}$. The line joining $\langle z\rangle$ and $\langle s\rangle$ is $s_2X_1-s_1X_2=0$, where $s=(s_1,s_2,s_3)$ are the coordinates of s with respect to the basis $\{x,y,z\}$.

Since $\mathcal{O}$ is an oval the set

$$\{\frac{s_2}{s_1}\mid\langle s\rangle\in\mathcal{O}\setminus\{\langle x\rangle,\langle y\rangle,\langle z\rangle\}\}\cup\{-\frac{\alpha_{13}}{\alpha_{23}}\}$$

contains every non-zero element of $\mathbb{F}_q$. Thus,

$$-\frac{\alpha_{13}}{\alpha_{23}}\prod_{\langle s\rangle\in\mathcal{O}\setminus\{\langle x\rangle,\langle y\rangle,\langle z\rangle\}}\frac{s_2}{s_1}=-1.$$

Define $T_x(X)=\alpha_{21}X_2+\alpha_{31}X_3$, $T_y(X)=\alpha_{12}X_1+\alpha_{32}X_3$ and $T_z(X)=\alpha_{13}X_1+\alpha_{23}X_2$. Since $T_z(x)=\alpha_{13}$ and $T_z(y)=\alpha_{23}$ we have $T_z(x)\prod s_2=T_z(y)\prod s_1$. Similarly, $T_x(y)\prod s_3=T_x(z)\prod s_2$ and $T_y(z)\prod s_1=T_y(x)\prod s_3$ and so

$$T_x(y)T_y(z)T_z(x)=T_x(z)T_y(x)T_z(y). \tag{2.1}$$

Let $\langle u\rangle$, $\langle v\rangle$ and $\langle w\rangle$ be three points of and oval $\mathcal{O}\setminus\langle x\rangle$. By interpolation

$$T_x(X)=T_x(u)\frac{\det(X,v,x)}{\det(u,v,x)}+T_x(v)\frac{\det(X,u,x)}{\det(v,u,x)}$$

since both sides are polynomials of degree 1 and agree at 2 points. Putting $X=w$ and rearranging gives

$$T_x(w)\det(u,v,x)+T_x(v)\det(w,u,x)+T_x(u)\det(v,w,x)=0 \tag{2.2}$$

Permuting the roles of x,u,v,w we also have that

$$T_u(w)\det(x,v,u)+T_u(v)\det(w,x,u)+T_u(x)\det(v,w,u)=0,$$

$$T_v(w)\det(x,u,v)+T_v(u)\det(w,x,v)+T_v(x)\det(u,w,v)=0,$$

$$T_w(u)\det(x,v,w)+T_w(v)\det(u,x,w)+T_w(x)\det(v,u,w)=0.$$

Now, by (2.2) and (2.1) we have

$$T_w(x)\det(u,v,x)+T_v(x)\frac{T_w(v)}{T_v(w)}\det(w,u,x)+T_u(x)\frac{T_w(u)}{T_u(w)}\det(v,w,x)=0,$$

and so

$$\det(u,v,x)(T_w(u)\det(x,v,w)+T_w(v)\det(u,x,w))+$$
$$\frac{T_w(v)}{T_v(w)}\det(w,u,x)(T_v(w)\det(x,u,v)+T_v(u)\det(w,x,v))-$$
$$\frac{T_w(u)}{T_u(w)}\det(v,w,x)(T_u(w)\det(x,v,u)+T_u(v)\det(w,x,u))=0,$$

and rearranging (using (2.1) for the third coefficient) gives

$$2T_w(u)\det(u,v,x)\det(x,v,w)+2T_w(v)\det(u,v,x)\det(u,x,w)+$$
$$2T_v(u)\frac{T_w(v)}{T_v(w)}\det(w,u,x)\det(w,x,v))=0.$$

Now with respect to the basis $\{u,v,w\}$, we see that an arbitrary point $\langle x\rangle$ of $\mathcal{O}$ satisfies the equation of a conic, namely

$$2T_w(u)x_3x_1+2T_w(v)x_3x_2+2T_v(u)\frac{T_w(v)}{T_v(w)}x_2x_1=0.$$

□

2.3 Maximum distance separable codes

A *code of length* n is a set of n-tuples (called *codewords*) of a set (called the *alphabet*). The *distance* between two codewords is the number of coordinates in which they differ, i.e. if $\mathbf{x}=(x_i)$ and $\mathbf{y}=(y_i)$ then the distance between them is

$$d(\mathbf{x},\mathbf{y})=|\{\ i\mid x_i\neq y_i\}|.$$

The *minimum distance* of a code $\mathcal{C}$ is the minimum value of $d(\mathbf{x},\mathbf{y})$ where $\mathbf{x}$ and $\mathbf{y}$ are any two distinct codewords of $\mathcal{C}$. A code with minimum distance at least $2e+1$ can correct up to e errors. So if we receive a codeword that has been distorted in at most e entries, then we can correctly deduce which codeword was sent. We say that the code is an e-error correcting code.

Proposition 2.3.1. *Suppose that the alphabet of a code of length n has size a. If the minimum distance is d then the code has at most a^{n-d+1} codewords.*

Proof. Take any $n-d+1$ coordinates. There are a^{n-d+1} ways of filling these coordinates with entries from the alphabet. Since there are more than a^{n-d+1} codewords, then there are two codewords that have the same entries in these $n-d+1$ coordinates. But then the minimum distance is at most $d-1$. □

A code of length n, with an alphabet of size a and minimum distance d that has a^{n-d+1} codewords is called a *maximum distance separable (MDS) code.*

A *linear code* $\mathcal{C}$ has alphabet $GF(q)$ and codewords that form a subspace of $V(n,q)$. If the rank of the subspace is k then we say that $\mathcal{C}$ is an $[n,k,d]$-code over $GF(q)$. The number of codewords in $\mathcal{C}$ is q^k and so by Proposition 2.3.1 we have

$$k \leq n-d+1,$$

which is called the *Singleton bound.*

A $[n, n-d+1, d]$-code over $GF(q)$ is an MDS code.

The *dual code* $\mathcal{C}^{\perp}$ of a linear code $\mathcal{C}$ is the set

$$\{\mathbf{y} \in \mathcal{C} \mid \mathbf{x} \cdot \mathbf{y} = 0 \text{ for all } \mathbf{x} \in \mathcal{C}\},$$

where $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n$. The dual code is clearly a linear code and it has rank $n-k$.

PROPOSITION 2.3.2. *The dual code of a linear MDS* $[n, n-d+1, d]$*-code is a linear MDS* $[n, d-1, n-d+2]$*-code.*

Proof. This is Exercise 6, the solution of which can be found in Chapter 5.1. □

The *weight* $wt(\mathbf{x})$ of a codeword $\mathbf{x}$ is the number of non-zero coordinates of $\mathbf{x}$. If $\mathbf{x}$ and $\mathbf{y}$ are two codewords of a linear code $\mathcal{C}$ then $\mathbf{x}-\mathbf{y} \in \mathcal{C}$, so the minimum weight of a codeword is equal to the minimum distance.

Let $\mathbf{c_1}, \mathbf{c_2}, \ldots, \mathbf{c_k}$ be a basis for a linear $[n,k,d]$-code over $GF(q)$. For all $\alpha \in GF(q)^k$, the codeword

$$\sum_{i=1}^{k} \alpha_i \mathbf{c_i},$$

has at most $n-d$ zero entries. Define vectors $\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n} \in V(k,q)$ by the rule

$$(\mathbf{a_j})_i = (\mathbf{c_i})_j,$$

and let $\mathcal{A} = \{\mathbf{a_j} \mid j = 1, 2, \ldots, n\}$. Consider the hyperplane of $V(k,q)$

$$\sum_{i=1}^{k} \alpha_i x_i = 0.$$

The number of vectors of $\mathcal{A}$ that are incident with this hyperplane is equal to the number of j such that

$$\sum_{i=1}^{k} \alpha_i (\mathbf{a_j})_i = 0,$$

which is equal to the number of j such that

$$\sum_{i=1}^{k} \alpha_i (\mathbf{c_i})_j = 0,$$

which is equal to the number of j such that

$$\left(\sum_{i=1}^{k} \alpha_i \mathbf{c_i} \right)_j = 0,$$

which is equal to the number of zero entries in $\sum_{i=1}^{k} \alpha_i \mathbf{c_i}$, which we observed was at most $n-d$.

Therefore $\mathcal{A}$ is a set (possibly multiset) of n vectors of $V(k,q)$ with the property that each hyperplane contains at most $n-d$ vectors of $\mathcal{A}$. In projective terms, $\mathcal{A}$ is a set of n points of $PG(k-1,q)$ with the property that each hyperplane is incident with at most $n-d$ points of $\mathcal{A}$ and the next section is devoted to the corresponding point sets. If the code is an MDS code then $k-1 = n-d$.

2.4 Arcs

An *arc* $\mathcal{A}$ is a set of at least $r+1$ points in $PG(r,q)$ with the property that every hyperplane is incident with at most r points of $\mathcal{A}$.

Therefore, from a linear MDS $[n, n-d+1, d]$-code we can construct an arc of $PG(n-d,q)$ with n points and from an arc with n points in $PG(r,q)$ we can construct a linear MDS $[n, r+1, n-r]$-code. In terms of the code, if we fix the minimum distance we would like to find a code with length n as large as possible to be able to send as much information as possible. If we fix the length n then we would like to find a code with distance d as large as possible to be able to correct the maximum number of errors. However, because we would like to fix the dimension of the projective space r, we fix $n-d$ and then look for an arc of size n, with n as large as possible.

Proposition 2.4.1. *The set of points of* $PG(r,q)$

$$\mathcal{A} = \{\langle(1, t, t^2, \ldots, t^r)\rangle \mid t \in GF(q)\} \cup \{\langle(0, 0, \ldots, 0, 1)\rangle\},$$

is an arc with $q+1$ *points.*

Proof. Consider the intersection of a hyperplane

$$\sum_{i=0}^{r} \alpha_i x_i = 0,$$

with the set $\mathcal{A}$. If $\alpha_r \neq 0$ then the hyperplane is incident with a point of $\mathcal{A}$ for each t satisfying

$$\sum_{i=0}^{r} \alpha_i t^i = 0,$$

which, since it is an equation of degree r, has at most r solutions. If $\alpha_r = 0$ then the hyperplane is incident with the point $\langle(0, 0, \ldots, 0, 1)\rangle$ of $\mathcal{A}$ and additional points of $\mathcal{A}$ for each t satisfying

$$\sum_{i=0}^{r-1} \alpha_i t^i = 0,$$

which, since it is an equation of degree $r-1$, has at most $r-1$ solutions. □

Hence for any prime power q and any d we can construct a linear MDS $[q+1, q+2-d, d]$-code over $GF(q)$. When $r = 2$, we have already seen examples of arcs with $q+1$ points, the ovals. Moreover when q is even, we also have examples of hyperovals, which are arcs with $q+2$ points. Hyperovals give linear MDS $[q+2, 3, q]$-codes over $GF(q)$ and their dual codes are linear MDS $[q+2, q-1, q]$-codes, by Proposition 2.3.2.

The following proposition deals with the (less interesting) case when r is large compared to q. Note that we can always construct an arc with $r+2$ points by taking

$$\mathcal{A} = \{\langle(1, 0, \ldots, 0)\rangle, \langle(0, 1, 0, \ldots, 0)\rangle, \ldots, \langle(0, \ldots, 0, 1)\rangle\} \cup \{\langle(1, 1, \ldots, 1)\rangle\}.$$

Proposition 2.4.2. *If $r \geq q-1$ then an arc in $PG(r, q)$ has at most $r+2$ points.*

Proof. This is Exercise 10, the solution of which can be found in Chapter 5.1. □

2.5 The main conjecture and a question of Segre

Now we are ready to mention what is called the main conjecture for MDS codes. Here we quote it in terms of arcs.

Conjecture 2.5.1. *Let $\mathcal{A}$ be an arc of $PG(r, q)$ with $r \leq q-1$.*

(i) If q is even and $r = 2$ or $q-2$ then $|\mathcal{A}| \leq q+2$.

(ii) If q is odd or $3 \leq r \leq q-3$ then $|\mathcal{A}| \leq q+1$.

Chao and Kaneta have verified by computer that the conjecture holds for $q \leq 27$. In a generalisation of the proof of Segre's theorem, Theorem 2.2.1, Conjecture 2.5.1 was shown to hold for q prime. In [1], the following is proven.

THEOREM 2.5.2. *An arc $\mathcal{A}$ in $PG(r,p)$, p prime and $r \leq p-1$, satisfies $|\mathcal{A}| \leq p+1$, with equality if and only if $\mathcal{A}$ is equivalent to the example in Proposition 2.4.1.*

As we have seen we have constructions for the bound in the conjecture. The following proposition provides us with a weaker upper bound.

PROPOSITION 2.5.3. *If $\mathcal{A}$ is an arc of $PG(r,q)$ then $|\mathcal{A}| \leq q+r$.*

Proof. Take any subset S of $\mathcal{A}$ of size $r-1$. If the points of S do not span a subspace of dimension $r-2$ then we can find a hyperplane that contains $r+1$ points of $\mathcal{A}$, contradicting the definition of an arc. The space of dimension $r-2$ spanned by the points of S is contained in $q+1$ hyperplanes, each of which is incident with at most one point of $\mathcal{A} \setminus S$. Hence

$$|\mathcal{A}| \leq q+1+r-1.$$

□

In relation to the main conjecture Segre, in 1955, asked the following question:

For what values of r and q do there exist arcs of size $q+1$ in $PG(r,q)$ that are not equivalent to the example in Proposition 2.4.1?

He knew that when $r=2$ or $r=q-2$ and q is even there were other examples. Indeed he proved that when q is non-square and even then the set

$$\mathcal{A} = \{\langle(1,t,t^6)\rangle \mid t \in GF(q)\} \cup \{\langle(0,0,1)\rangle\},$$

is an arc with $q+1$ points (so it is an oval). It is projectively inequivalent to Example 2.4.1 for $q \geq 32$.

When q is odd or when $3 \leq r \leq q-3$ there is only one non-classical arc (not equivalent to Example 2.4.1) of size $q+1$ known. It was discovered by Glynn in 1986 and is the following.

PROPOSITION 2.5.4. *The set of points of $PG(4,9)$*

$$\mathcal{A} = \{\langle(1,t,t^2+\eta t^6,t^3,t^4)\rangle \mid t \in GF(9)\} \cup \{\langle(0,0,0,0,1)\rangle\},$$

where η is a fixed element satisfying $\eta^4 = -1$, is an arc of size 10.

Proof. Let $|\,1\ t_i\ t_i^2\ t_i^3\ t_i^4\,|$ represent the determinant whose i-th row is

$$(\,1,\ t_i,\ t_i^2,\ t_i^3,\ t_i^4\,).$$

If $\mathcal{A}$ is not an arc then it contains 5 dependent points. Hence either there exist distinct t_i such that

$$|\,1\ t_i\ t_i^2+\eta t_i^6\ t_i^3\ t_i^4\,| = 0,$$

or one of these 5 points is $\langle(0,0,0,0,1)\rangle$ and there exist distinct t_i such that the determinant

$$|\,1\ t_i\ t_i^2+\eta t_i^6\ t_i^3\,| = 0.$$

In the first case we have that

$$|\,1\ t_i\ t_i^2\ t_i^3\ t_i^4\,| = -\eta|\,1\ t_i\ t_i^6\ t_i^3\ t_i^4\,|. \tag{2.3}$$

Now cubing both sides and first applying $t^9 = t$, then using 2.3 we have

$$\begin{aligned}|\,1\ t_i^3\ t_i^6\ t_i\ t_i^4\,| &= -\eta^3|\,1\ t_i^3\ t_i^2\ t_i\ t_i^4\,| = -\eta^3|\,1\ t_i\ t_i^2\ t_i^3\ t_i^4\,| \\ &= \eta^4|\,1\ t_i\ t_i^6\ t_i^3\ t_i^4\,| = \eta^4|\,1\ t_i^3\ t_i^6\ t_i\ t_i^4\,|,\end{aligned}$$

and so $\eta^4 = 1$. Similarly $|\,1\ t_i\ t_i^2+\eta t_i^6\ t_i^3\,| = 0$ implies $\eta^4 = 1$. □

The final two theorems that we include here without a proof relate to Segre's question and the main conjecture. The first is a result proved by Segre [16], from which Theorem 2.2.1 follows as a corollary.

THEOREM 2.5.5. *Let $\mathcal{A}$ be an arc of $PG(2,q)$ and let $\mathcal{A}^*$ be the set of lines dual to the points of $\mathcal{A}$. Let*

$$k = \begin{cases} q+2-|\mathcal{A}| & \text{when } q \text{ is even,} \\ \\ 2(q+2-|\mathcal{A}|) & \text{when } q \text{ is odd.} \end{cases}$$

Then there is a polynomial f homogeneous in three variables of degree k whose zeros include the set $\mathcal{Z}$ of points that lie on exactly one line of $\mathcal{A}^*$.

Moreover, when q is odd, for each point $P \in \mathcal{Z}$, if l_P is line of $\mathcal{A}^*$ incident with P then f mod l_P has a zero of degree 2 at P.

The generalisation of this theorem to higher dimensions is the following theorem by Blokhuis, Bruen and Thas [2].

THEOREM 2.5.6. *Let $\mathcal{A}$ be an arc of $PG(r,q)$ and let $\mathcal{A}^*$ be the set of hyperplanes dual to the points of $\mathcal{A}$. Let*

$$k = \begin{cases} q+r-|\mathcal{A}| & \text{when } q \text{ is even,} \\ \\ 2(q+r-|\mathcal{A}|) & \text{when } q \text{ is odd.} \end{cases}$$

Then there is a polynomial f homogeneous in $r+1$ variables of degree k whose zeros include the set $\mathcal{Z}$ of points that lie on exactly $r-1$ hyperplanes of $\mathcal{A}^*$.

Moreover, when q is odd, for each point $P \in \mathcal{Z}$, if l_P is a line that is the intersection of the $r-1$ hyperplanes of $\mathcal{A}^*$ incident with P then f mod l_P has a zero of degree 2 at P.

2.6 Exercises

5. Prove that any five points, no three collinear, are contained in a unique conic of $PG(2,q)$. Deduce that the number of conics is

$$(q^2+q+1)q^2(q-1).$$

6. Prove that the dual of a linear MDS code is a linear MDS code.

7. Prove that the number of polynomials of degree at most $q-1$ in $GF(q)[X]$ is the same as the number of functions from $GF(q)$ to $GF(q)$. Conclude that every function from $GF(q)$ to $GF(q)$ can be represented by a polynomial of degree at most $q-1$ in $GF(q)[X]$.

8. Prove that the tangents to an oval $\mathcal{O}$ in $PG(2,q)$, q even, are incident with a common point N.

The point N is called the *nucleus* of the oval $\mathcal{O}$. The union of an oval with its nucleus is called a *hyperoval.*

9. Without loss of generality assume that a hyperoval $\mathcal{H}$ of $PG(2,q)$, q even, contains the point $\langle(0,0,1)\rangle$ and the point $\langle(0,1,0)\rangle$.

1. By considering the lines incident with $\langle(0,0,1)\rangle$, show that there exists a function f such that

$$\mathcal{H} = \{\langle(0,0,1)\rangle\} \cup \{\langle(0,1,0)\rangle\} \cup \{\langle(1,x,f(x))\rangle \mid x \in GF(q)\}.$$

2. By considering the lines incident with $\langle(0,1,0)\rangle$, show that the map $x \mapsto f(x)$ is a permutation of $GF(q)$.

3. By considering the lines incident with a fixed point $\langle(1,s,f(s))\rangle$ prove that for all $s \in GF(q)$

$$x \mapsto (f(x+s)+f(s))/x$$

is a permutation of $GF(q)$.

4. Conclude that if a function f satisfies the properties in (b) and (c) then the set $\mathcal{H}$ in (a) is a hyperoval.

10. Prove that an arc in $PG(r,q)$, when $r \geq q-1$, has at most $r+2$ points.

Chapter 3

Polar geometries

3.1 Dualities and polarities

A *duality* π of $V(n,q)$ is a map from the subspaces of $V(n,q)$ to the subspaces of $V(n,q)$ that reverses inclusion. In projective terms π takes points to hyperplanes, lines to spaces of co-dimension 2,.... If $P \in l$ then $\pi(l) \subset \pi(P)$, etc...

A *σ-sesquilinear form* on $V(n,q) = V$ is a map

$$\beta : V \times V \mapsto GF(q),$$

such that

$$\beta(\mathbf{u}+\mathbf{w},\mathbf{v}) = \beta(\mathbf{u},\mathbf{v}) + \beta(\mathbf{w},\mathbf{v}),$$
$$\beta(\mathbf{u},\mathbf{w}+\mathbf{v}) = \beta(\mathbf{u},\mathbf{w}) + \beta(\mathbf{u},\mathbf{v}),$$
$$\beta(a\mathbf{u},b\mathbf{v}) = ab^{\sigma}\beta(\mathbf{u},\mathbf{v}),$$

where σ is an automorphism of $GF(q)$. If $\sigma = 1$ then β is called *bilinear*.

A form is *degenerate* if there exists a $\mathbf{w} \neq \mathbf{0}$ such that $\beta(\mathbf{u},\mathbf{w}) = 0$ for all $\mathbf{u} \in V$ or $\beta(\mathbf{w},\mathbf{u}) = 0$ for all $\mathbf{u} \in V$.

Given a non-degenerate σ-sesquilinear form β, we can construct a duality

$$\pi : X \mapsto \{\mathbf{u} \mid \beta(\mathbf{u},\mathbf{v}) = 0, \text{ for all } \mathbf{v} \in X\},$$

where $X \leq V$. Note that for all $\mathbf{u} \in V$ the hyperplane $\pi(\mathbf{u})$ is the kernel of the linear map

$$\Phi_{\mathbf{u}} : \mathbf{v} \mapsto \beta(\mathbf{u},\mathbf{v})^{\sigma^{-1}}.$$

The σ^{-1} in the exponent is necessary to convert the σ-linear map to a linear map.

The converse is also true; a duality is induced by a non-degenerate σ-sesquilinear form.

A σ-sesquilinear form β is called *reflexive* if $\beta(\mathbf{u},\mathbf{v})=0$ implies $\beta(\mathbf{v},\mathbf{u})=0$.
A duality π induces a map π^* where

$$\pi^*(\mathbf{u})=\{\mathbf{v}\in V \mid \mathbf{u}\in\pi(\mathbf{v})\}.$$

If π is a duality constructed from a σ-sesquilinear form β then

$$\pi^*: X\mapsto\{\mathbf{u}\mid\beta(\mathbf{v},\mathbf{u})=0, \text{ for all } \mathbf{v}\in X\}.$$

It is called a *polarity* if $\pi\pi^*$ is the identity on $PG(n-1,q)$.
For any vector $\mathbf{u}$, we define $\mathbf{u}$ *perp* that is:

$$\langle\mathbf{u}\rangle^\perp := \{\mathbf{v}\in V\mid\beta(\mathbf{u},\mathbf{v})=0\}.$$

Similarly for subspaces,

$$U^\perp := \{\mathbf{v}\in V\mid\beta(\mathbf{u},\mathbf{v})=0 \text{ for all } \mathbf{u}\in U\}.$$

PROPOSITION 3.1.1. *If π is a duality constructed from a σ-sesquilinear form β then π is a polarity if and only if β is reflexive.*

Proof. Let π be a polarity. Then $\mathbf{u}\in\langle\mathbf{v}\rangle^\perp$ implies $\mathbf{v}\in\langle\mathbf{u}\rangle^\perp$ and so $\beta(\mathbf{u},\mathbf{v})=0$ implies $\beta(\mathbf{v},\mathbf{u})=0$. Conversely, if β is reflexive then $X\subseteq X^{\perp\perp}$ and by non-degeneracy $\dim X^{\perp\perp}=n-\dim X^\perp=n-(n-\dim X)=\dim X$, hence $X=X^{\perp\perp}$. □

3.2 The classification of forms

In this section we will prove the Birkhoff and von Neumann theorem.

THEOREM 3.2.1. *Let β be a non-degenerate σ-sesquilinear reflexive form on $V=V(n,q)$. Up to a scalar factor β is one of the following types.*

(i) Alternating: $\beta(\mathbf{u},\mathbf{u})=0$ for all $\mathbf{u}\in V$.

(ii) Symmetric: $\beta(\mathbf{u},\mathbf{v})=\beta(\mathbf{v},\mathbf{u})$ for all $\mathbf{u},\mathbf{v}\in V$.

(iii) Hermitian: $\beta(\mathbf{u},\mathbf{v})=\beta(\mathbf{v},\mathbf{u})^\sigma$ for all $\mathbf{u},\mathbf{v}\in V$ where $\sigma^2=1$, $\sigma\neq1$.

Note that (iii) implies that $\beta(\mathbf{w},\mathbf{w})\in GF(\sqrt{q})$, $\mathbf{w}\in V$. Furthermore, any scalar multiple of β will define the same polarity as β on the subspaces of $PG(n-1,q)$.

Proof. For all $\mathbf{u}\in V$ define linear maps $\Psi_\mathbf{u}: V\to GF(q)$ by $\Psi_\mathbf{u}:\mathbf{v}\mapsto\beta(\mathbf{v},\mathbf{u})$ and $\Phi_\mathbf{u}: V\to GF(q)$ by $\Phi_\mathbf{u}:\mathbf{v}\mapsto\beta(\mathbf{u},\mathbf{v})^{\sigma^{-1}}$.

The duality π constructed from β takes $\mathbf{u}$ to ker$\Phi_{\mathbf{u}}$ and π^* takes ker$\Psi_{\mathbf{u}}$ to $\mathbf{u}$. The linear maps are determined by their kernels up to a scalar factor and so $\pi\pi^*$ is the identity map if and only if there is a $\lambda \in GF(q)$ such that

$$\beta(\mathbf{u},\mathbf{v})^{\sigma^{-1}} = \lambda\beta(\mathbf{v},\mathbf{u}), \tag{3.1}$$

for all $\mathbf{u},\mathbf{v} \in V$.

Either $\beta(\mathbf{u},\mathbf{u}) = 0$ for all $\mathbf{u} \in V$ or there exists a $\mathbf{w} \in V$ such that $\beta(\mathbf{w},\mathbf{w}) = a \neq 0$. If we replace β by $a^{-1}\beta$ and put $\mathbf{u} = \mathbf{v} = \mathbf{w}$ in (3.1) then $\lambda = 1$. By non-degeneracy there is a vector $\mathbf{v} \in V \setminus \langle\mathbf{w}\rangle^{\perp}$. The form β takes every value in $GF(q)$ since for all $\nu \in GF(q)$ we have $\beta(\nu\mathbf{v},\mathbf{w}) = \nu\beta(\mathbf{v},\mathbf{w})$. Hence for all $\alpha \in GF(q)$

$$\alpha = \beta(\mathbf{u},\mathbf{v}) = \beta(\mathbf{v},\mathbf{u})^{\sigma} = \beta(\mathbf{u},\mathbf{v})^{\sigma^2} = \alpha^{\sigma^2}$$

and therefore $\sigma^2 = 1$. □

3.3 An application to graphs of fixed degree and diameter

A polarity π of $PG(2,q)$ takes points to lines and has the property that for any points P and Q

$$P \in \pi(Q) \text{ if and only if } Q \in \pi(P).$$

We construct a graph G whose vertices are the points of $PG(2,q)$ and there is an edge between P and Q ($P \sim Q$) if and only if $P \in \pi(Q)$. There are $q+1$ points on a line in $PG(2,q)$ so the (maximum) degree of the graph is $q+1$. We do not allow loops, so if $P \in \pi(P)$ then the degree of the vertex P will be q.

Proposition 3.3.1. *The length of the shortest path between two vertices is at most 2.*

Proof. Let P and R be two non-adjacent vertices. The neighbours of P are the points incident with $\pi(P)$. In $PG(2,q)$ $\pi(P)$ and $\pi(R)$ are incident in a point Q and so there is a path of length two $P \sim Q \sim R$ between P and R. □

In a graph G the *diameter* refers the the maximum length of the shortest path between two distinct vertices. Therefore we have constructed a graph G of degree $q+1$ and diameter 2.

Proposition 3.3.2. *A graph G of degree Δ and diameter d has at most*

$$1 + \Delta((\Delta-1)^d - 1)/(\Delta - 2)$$

vertices.

Proof. Let v be any vertex of G. There are at most Δ of vertices at distance 1 from v, at most $\Delta(\Delta-1)$ of vertices at distance 2 from v, at most $\Delta(\Delta-1)^2$ of vertices at distance 3 from v, etc..... Since there are no vertices of G at distance more than d from v the number of vertices of G is at most

$$1+\Delta+\Delta(\Delta-1)+\ldots+\Delta(\Delta-1)^{d-1}.$$

□

When $d=2$ this bound is Δ^2+1, whereas the graph we constructed from the polarity has $\Delta^2-\Delta+1$ vertices.

3.4 Quadratic forms

A *quadratic form* on $V(n,q)=V$ is a function $Q:V(n,q)\rightarrow GF(q)$ with the properties

(QF1) $Q(a\mathbf{v})=a^2Q(\mathbf{v})$ for all $a\in GF(q)$ and $\mathbf{v}\in V$;

(QF2) $\beta(\mathbf{u},\mathbf{v})=Q(\mathbf{u}+\mathbf{v})-Q(\mathbf{u})-Q(\mathbf{v})$ is a bilinear form.

We say that β is the *polar form* of Q. Note that β is symmetric.

A quadratic form is *singular* if there exists a $\mathbf{u}\neq\mathbf{0}$ such that $Q(\mathbf{u})=\beta(\mathbf{u},\mathbf{v})=0$ for all $\mathbf{v}\in V$.

We want to show that it is enough to consider alternating, hermitian and quadratic forms, since the symmetric is included in these forms.

If q is odd then $Q(\mathbf{u})=\beta(\mathbf{u},\mathbf{u})/2$ and we can recover the quadratic form from the symmetric bilinear form β. However, when q is even $\beta(\mathbf{u},\mathbf{u})=Q(2\mathbf{u})-2Q(\mathbf{u})=0$ for all $\mathbf{u}\in V$ and the bilinear form is alternating.

Note that an alternating form is a symmetric form when q is even since for all $\mathbf{u},\mathbf{v}\in V$ we have $\beta(\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v})=0$, which when we expand implies $\beta(\mathbf{u},\mathbf{v})=-\beta(\mathbf{v},\mathbf{u})=\beta(\mathbf{v},\mathbf{u})$.

If β is symmetric and q is even then

$$\beta(\mathbf{u}+\lambda\mathbf{v},\mathbf{u}+\lambda\mathbf{v})=\beta(\mathbf{u},\mathbf{u})+\lambda^2\beta(\mathbf{v},\mathbf{v}).$$

If $\beta(\mathbf{v},\mathbf{v})\neq 0$ then when $\lambda=(\beta(\mathbf{u},\mathbf{u})/\beta(\mathbf{v},\mathbf{v}))^{1/2}$, $\beta(\mathbf{u}+\lambda\mathbf{v},\mathbf{u}+\lambda\mathbf{v})=0$. Thus every subspace of rank two contains a vector $\mathbf{w}$ for which $\beta(\mathbf{w},\mathbf{w})=0$. Moreover, if $\beta(\mathbf{u},\mathbf{u})=0$ and $\beta(\mathbf{v},\mathbf{v})=0$ then every vector $\mathbf{w}$ in their span satisfies $\beta(\mathbf{w},\mathbf{w})=0$. So we can conclude that if β is symmetric but not alternating, and q is even then V contains a hyperplane H on which β is alternating.

We will from now on consider the geometries formed by the subspaces on which these forms are zero. Thus we restrict our attention to alternating, hermitian and quadratic forms, the symmetric forms being included in the quadratic forms for q odd and for q even, restricting to a hyperplane if necessary, the symmetric forms are dealt with by the alternating forms.

3.5 Polar spaces

Let β be a non-degenerate σ-sesquilinear form on $V(n,q) = V$. A vector $\mathbf{u}$ is called *isotropic* if $\beta(\mathbf{u},\mathbf{u}) = 0$.

A subspace U is called *totally isotropic* if $\beta(\mathbf{u},\mathbf{v}) = 0$ for all $\mathbf{u},\mathbf{v} \in U$.

A pair $(\mathbf{u},\mathbf{v})$ of isotropic vectors is called a *hyperbolic pair* if $\beta(\mathbf{u},\mathbf{v}) = 1$. The line $\langle \mathbf{u},\mathbf{v}\rangle$ is called a hyperbolic line. This is well defined since $\beta(\mathbf{u},\mathbf{v}) = 1$ implies that $\beta(\mathbf{v},\mathbf{u}) = 1$.

Furthermore, note also that $\beta(\mathbf{u},\mathbf{v}) = 0$ means that $\beta(\mathbf{v},\mathbf{u}) = 0$.

A subspace U is called *anisotropic* if $\beta(\mathbf{u},\mathbf{u}) \neq 0$ for all $\mathbf{u} \in U$.

The *polar space* associated to a σ-sesquilinear form on $V(n,q) = V$ is the geometry whose points, lines, planes, …, are the totally isotropic subspaces of $V(n,q)$ of rank 1, 2, 3, ….

Let Q be a non-singular quadratic form on $V(n,q) = V$.

A vector $\mathbf{u}$ is called *singular* if $Q(\mathbf{u}) = 0$.

A subspace U is called *totally singular* if $Q(\mathbf{u}) = 0$ for all $\mathbf{u} \in U$.

A pair $(\mathbf{u},\mathbf{v})$ of singular vectors is called a *hyperbolic pair* if $\beta(\mathbf{u},\mathbf{v}) = 1$. The line $\langle \mathbf{u},\mathbf{v}\rangle$ is called a hyperbolic line.

A subspace U is called *anisotropic* if $Q(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in U$.

The *polar space* associated to a quadratic form on $V(n,q) = V$ is the geometry whose points, lines, planes, …, are the totally singular subspaces of $V(n,q)$ of rank 1, 2, 3, ….

If the form is alternating the corresponding polar space is called *symplectic.*

If the form is hermitian the corresponding polar space is called *unitary.*

If the form is quadratic the corresponding polar space is called *orthogonal.*

In all cases the vector space V is called the *ambient space* of the polar space.

Lemma 3.5.1. *Suppose that L is a subspace of rank 2 of $V(n,q)$ that contains an isotropic vector $\mathbf{u}$ with respect to an alternating or hermitian form β. Either β restricted to L is degenerate or there is a vector $\mathbf{v}$ such that $(\mathbf{u},\mathbf{v})$ is a hyperbolic pair and $L = \langle \mathbf{u},\mathbf{v}\rangle$.*

Proof. If β restricted to L is not degenerate then there is a vector $\mathbf{w}$ such that $\beta(\mathbf{u},\mathbf{w}) = a \neq 0$.

If β is alternating then take $\mathbf{v} = a^{-1}\mathbf{w}$.

If β is hermitian then choose d such that $d^\sigma \neq -d$ and let

$$c = (d + d^\sigma)^{-1} d^\sigma \beta(\mathbf{w}, \mathbf{w}).$$

Then $c^\sigma = (d + d^\sigma)^{-1} d\beta(\mathbf{w}, \mathbf{w})$ and $c + c^\sigma = \beta(\mathbf{w}, \mathbf{w})$. Put $a := \beta(\mathbf{w}, \mathbf{w})$ and take $\mathbf{v} = -a^{-1-\sigma} c\mathbf{u} + a^{-\sigma}\mathbf{w}$. Check that

$$\beta(\mathbf{u}, \mathbf{v}) = \beta(\mathbf{u}, a^{-\sigma}\mathbf{w}) = a^{-1}\beta(\mathbf{u}, \mathbf{w}) = 1$$

and

$$\begin{aligned}\beta(\mathbf{v}, \mathbf{v}) = \beta(-a^{-1-\sigma}c\mathbf{u}, -a^{-1-\sigma}c\mathbf{u}) + \beta(-a^{-1-\sigma}c\mathbf{u}, -a^{-\sigma}\mathbf{w}) + \\ \beta(a^{-\sigma}\mathbf{w}, -a^{-1-\sigma}c\mathbf{u}) + \beta(a^{-\sigma}\mathbf{w}, a^{-\sigma}\mathbf{w}) \\ = -a^{-2-\sigma}ca - a^{-2\sigma-1}a^\sigma c^\sigma + a^{-\sigma-1}(c + c^\sigma) = 0.\end{aligned}$$

□

Lemma 3.5.2. *Suppose that L is a subspace of rank 2 of $V(n, q)$ that contains a singular vector $\mathbf{u}$ with respect to a quadratic form Q. Either Q restricted to L is singular or there is a vector $\mathbf{v}$ such that $(\mathbf{u}, \mathbf{v})$ is a hyperbolic pair and $L = \langle \mathbf{u}, \mathbf{v} \rangle$.*

Proof. Let β be the polar form of Q. If Q restricted to L is not singular then there is a vector $\mathbf{w}$ such that $\beta(\mathbf{u}, \mathbf{w}) = a \neq 0$. Take $\mathbf{v} = -a^{-2}Q(\mathbf{w})\mathbf{u} + a^{-1}\mathbf{w}$. Check that $\beta(\mathbf{u}, \mathbf{v}) = \beta(\mathbf{u}, a^{-1}\mathbf{w}) = aa^{-1} = 1$ and

$$\begin{aligned}Q(\mathbf{v}) = \beta(-a^{-2}Q(\mathbf{w})\mathbf{u}, a^{-1}\mathbf{w}) + Q(-a^{-2}Q(\mathbf{w})\mathbf{u}) + Q(a^{-1}\mathbf{w}) \\ = -a^{-3}Q(\mathbf{w})a + a^{-2}Q(\mathbf{w}) = 0,\end{aligned}$$

as required. □

Recall that for any subspace $U \leq V$ we define

$$U^\perp = \{\mathbf{v} \in V \mid \beta(\mathbf{u}, \mathbf{v}) = 0 \text{ for all } \mathbf{u} \in U\}.$$

Theorem 3.5.3. *Let β be a non-degenerate alternating or hermitian form on $V(n, q)$. Let W be a maximal totally isotropic subspace and let r be the rank of W. There exists a basis*

$$\{\mathbf{e_i} \mid i = 1, 2, \ldots r\} \cup \{\mathbf{f_i} \mid i = 1, 2, \ldots r\}$$

for a subspace $X \leq V = V(n, q)$ such that $W = \langle e_1, e_2, \ldots, e_r \rangle$ and

$$V = X \oplus U,$$

where $(\mathbf{e_i}, \mathbf{f_i})$ is a hyperbolic pair and U is anisotropic.

Proof. If $r = 0$ then there is nothing to prove. If not, there is an isotropic vector $\mathbf{e_1} \in W$ and since the form is non-degenerate there exists a $\mathbf{v} \in V$ such that $\beta(\mathbf{e_1}, \mathbf{v}) \neq 0$ and β restricted to $\langle \mathbf{e_1}, \mathbf{v} \rangle$ is non-degenerate. By Lemma 3.5.1 there is a hyperbolic pair $(\mathbf{e_1}, \mathbf{f_1})$ such that

$$V = \langle \mathbf{e_1}, \mathbf{f_1} \rangle \oplus V_1,$$

where $V_1 = \langle \mathbf{e_1}, \mathbf{f_1} \rangle^{\perp}$. Now let $W_1 = W \cap V_1$ and check that the rank of W_1 is $rk(W) + rk(V_1) - rk(W + V_1) = r + n - 2 - (n-1) = r - 1$.

If β restricted to V_1 is degenerate then there exists a $0 \neq \mathbf{u} \in V_1$ such that $\beta(\mathbf{u}, \mathbf{w}) = 0$ for all $\mathbf{w} \in V_1$. However $\beta(\mathbf{u}, \mathbf{v}) = 0$ for all $\mathbf{v} \in \langle \mathbf{e_1}, \mathbf{f_1} \rangle$ so β would be degenerate. Now apply the same argument again with the vector space V_1 and the totally isotropic subspace W_1. □

Note that r is invariant given V and β. If there exists a maximal totally isotropic subspace W' of rank $s < r$ and $\mathbf{e'_1}, \ldots, \mathbf{e'_s}$ is basis for W' then there is an endomorphism α that takes $\mathbf{e_i}$ to $\mathbf{e'_i}$ for $i = 1, \ldots, s$ and which extends to a change of basis endomorphism on the whole space V but which takes $\mathbf{e_r}$ to a vector which is perpendicular to all vectors in W'. By the previous theorem there is no such vector.

The *rank of a polar space* is defined to be the r in the previous theorem. Note that $r \leq n/2$.

THEOREM 3.5.4. *Let Q be a non-singular quadratic form on $V(n, q) = V$. Let W be a maximal totally isotropic subspace of rank r. Then there exists a basis*

$$\{\mathbf{e_i} \mid i = 1, 2, \ldots, r\} \cup \{\mathbf{f_i} \mid i = 1, 2, \ldots r\}$$

for a subspace $X \leq V$ such that $W = \langle \mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_r} \rangle$ and

$$V = X \oplus U,$$

where $(\mathbf{e_i}, \mathbf{f_i})$ is a hyperbolic pair and U is anisotropic.

Proof. If $r = 0$ then there is nothing to prove.

If not, there is a singular vector $\mathbf{e_1} \in W$ and since the form is non-singular there exists a $\mathbf{v} \in V$ such that $\beta(\mathbf{e_1}, \mathbf{v}) \neq 0$, where β is the polar form of Q. Now Q restricted to $\langle \mathbf{e_1}, \mathbf{v} \rangle$ is non-singular so by lemma 3.5.2 there is a hyperbolic pair $(\mathbf{e_1}, \mathbf{f_1})$ such that

$$V = \langle \mathbf{e_1}, \mathbf{f_1} \rangle \oplus V_1,$$

where $V_1 = \langle \mathbf{e_1}, \mathbf{f_1} \rangle^{\perp}$. Let $W_1 = W \cap V_1$ and check the rank of W_1 is $rk(W) + rk(V_1) - rk(W + V_1) = r + n - 2 - (n-1) = r - 1$.

If Q restricted to V_1 is singular then there exists a $0 \neq \mathbf{u} \in V_1$ such that $Q(\mathbf{u}) = 0$ and $\beta(\mathbf{u}, \mathbf{w}) = 0$ for all $\mathbf{w} \in V_1$. However $\beta(\mathbf{u}, \mathbf{v}) = 0$ for all $\mathbf{v} \in \langle \mathbf{e_1}, \mathbf{f_1} \rangle$ so Q would be singular. Now apply the same argument again with the vector space V_1 and the totally singular subspace W_1. □

Again one can show that r is invariant.

Corollary 3.5.5. *Let ρ be a polar space of rank r with anisotropic space U. If $(\mathbf{u},\mathbf{v})$ is a hyperbolic pair then $\langle \mathbf{u},\mathbf{v}\rangle^{\perp}$ is a polar space of rank $r-1$ with anisotropic space U.*

Proof. Extend $(\mathbf{u},\mathbf{v}) = (\mathbf{e_1},\mathbf{f_1})$ to a basis of hyperbolic lines as in Theorem 3.5.3 or Theorem 3.5.4 where appropriate. The totally isotropic (singular) subspaces of β restricted to $\langle \mathbf{u},\mathbf{v}\rangle^{\perp} = \langle \mathbf{e_2},\mathbf{f_2},\dots,\mathbf{e_r},\mathbf{f_r}\rangle \oplus U$ is a polar space of rank $r-1$ with anisotropic space U. □

Corollary 3.5.6. *Let ρ be a polar space of rank r with anisotropic space U. If $\mathbf{u}$ is an isotropic (singular) vector then $\langle \mathbf{u}\rangle^{\perp}/\mathbf{u}$ is a polar space of rank $r-1$ with anisotropic space U.*

Proof. Extend $\mathbf{u} = \mathbf{e_1}$ to a basis of hyperbolic lines as in Theorem 3.5.3 or Theorem 3.5.4 where appropriate. The totally isotropic subspaces of β restricted to $\langle \mathbf{u}\rangle^{\perp}/\mathbf{u} = \langle \mathbf{e_2},\mathbf{f_2},\dots,\mathbf{e_r},\mathbf{f_r}\rangle \oplus U$ is a polar space of rank $r-1$ with anisotropic space U. □

We will use the following obsevation several times.

Corollary 3.5.7. *Let ρ be a polar space of rank r with anisotropic space U. For any vector $\langle \mathbf{x},\,\mathbf{x}\rangle^{\perp} \cap \rho$ is a cone with vertex x and base a polar space of rank $r-1$ with anisotropic space U.*

3.6 Symplectic spaces

Let β be a non-degenerate alternating form on $V(n,q) = V$. An anisotropic subspace has rank 0 and so according to Theorem 3.5.3 there is a basis $\{\mathbf{e_i} \mid i = 1,2,\dots r\} \cup \{\mathbf{f_i} \mid i = 1,2,\dots r\}$ of V such that $(\mathbf{e_i},\mathbf{f_i})$ is a hyperbolic pair. Note that $n = 2r$.

Choose $\mathbf{e_1} = (1,0,0,\dots,0)$ and $\mathbf{f_1} = (0,1,0,\dots,0)$. Consider β restricted to $\langle \mathbf{e_1},\mathbf{f_1}\rangle$. It follows from $\beta(\mathbf{e_1},\mathbf{e_1}) = 0$, $\beta(\mathbf{f_1},\mathbf{f_1}) = 0$ and $\beta(\mathbf{e_1},\mathbf{f_1}) = 1$ that β in matrix form

$$\beta(\mathbf{u},\mathbf{v}) = \mathbf{u}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\mathbf{v} = u_1v_2 - v_1u_2.$$

By induction, we can choose a basis for V such that

$$\beta(\mathbf{u},\mathbf{v}) = u_1v_2 - v_1u_2 + u_3v_4 - v_3u_4 + \ldots + u_{n-1}v_n - v_{n-1}u_n.$$

Note that this means that all non-degenerate alternating forms on V are equivalent, i.e. if β_1 and β_2 are non-degenerate alternating forms then there exists an $f \in GL(n,q)$ such that

$$\beta_1(f(\mathbf{u}), f(\mathbf{v})) = \beta_2(\mathbf{u},\mathbf{v})$$

for all $\mathbf{u}, \mathbf{v} \in V$.

The symplectic polar space of rank r formed from the totally isotropic subspaces of a non-degenerate alternating form on $V(2r, q)$ is denoted $W(2r-1, q)$. The elements $f \in GL(2r, q)$ with the property that

$$\beta(f(\mathbf{u}), f(\mathbf{v})) = \beta(\mathbf{u}, \mathbf{v})$$

for all $\mathbf{u}, \mathbf{v} \in V$ form a group under composition. This group is called the *symplectic group* and is denoted $Sp(2r, q)$. The *projective symplectic group*

$$PSp(2r, q) = Sp(2r, q)/\{\pm I\}$$

is a simple group unless $(r, q) = (1, 2)$ or $(r, q) = (1, 3)$.

The symplectic polar space of rank 2 over $GF(2)$ is drawn in Figure 3.1.

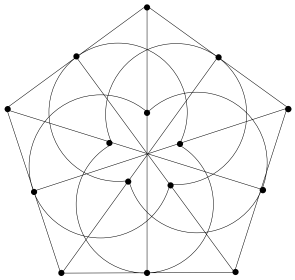

Figure 3.1: $W(3, 2)$

3.7 Unitary spaces

The following lemma restricts the dimension of an anisotropic subspace of a vector space equipped with a non-degenerate hermitian form. In this section $x^\sigma = x^{\sqrt{q}}$.

Lemma 3.7.1. *Let β be a non-degenerate hermitian form on $V(n, q) = V$. If the rank of V is at least 2 then V has an isotropic vector.*

Proof. Suppose that $\mathbf{v}$ is not an isotropic vector and put $b = \beta(\mathbf{v}, \mathbf{v})$. For a vector $\mathbf{u} \in \langle \mathbf{v} \rangle^\perp$,

$$\beta(\mathbf{u} + a\mathbf{v}, \mathbf{u} + a\mathbf{v}) = \beta(\mathbf{u}, \mathbf{u}) + a^{\sigma+1}\beta(\mathbf{v}, \mathbf{v}).$$

Now $-\beta(\mathbf{u}, \mathbf{u})/b \in GF(\sqrt{q})$ and the map $x \mapsto x^{\sigma+1}$ is a surjective map from $GF(q)$ to $GF(\sqrt{q})$ so we can choose a such that $a^{\sigma+1} = -\beta(\mathbf{u}, \mathbf{u})/b$. □

Let $(\mathbf{e_1}, \mathbf{f_1})$ be a hyperbolic pair.

Choose $\mathbf{e_1} = (1, 0, 0, \dots, 0)$ and $\mathbf{f_1} = (0, 1, 0, \dots, 0)$. Consider β restricted to $\langle \mathbf{e_1}, \mathbf{f_1} \rangle$. It follows from $\beta(\mathbf{e_1}, \mathbf{e_1}) = 0$, $\beta(\mathbf{f_1}, \mathbf{f_1}) = 0$ and $\beta(\mathbf{e_1}, \mathbf{f_1}) = 1$ that β in matrix form

$$\beta(\mathbf{u}, \mathbf{v}) = \mathbf{u} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{v}^\sigma = u_1 v_2^\sigma + u_2 v_1^\sigma.$$

According to Lemma 3.7.1 an anisotropic subspace has rank 0 or 1. Therefore according to Theorem 3.5.3 we can extend $(\mathbf{e_1}, \mathbf{f_1})$ to a basis for $V(n, q)$ such that

$$\beta(\mathbf{u}, \mathbf{v}) = u_1 v_2^\sigma + u_2 v_1^\sigma + u_3 v_4^\sigma + u_4 v_3^\sigma + \dots + u_{n-1} v_n^\sigma + u_n v_{n-1}^\sigma$$

if $n = 2r$ and

$$\beta(\mathbf{u}, \mathbf{v}) = u_1 v_2^\sigma + u_2 v_1^\sigma + u_3 v_4^\sigma + u_4 v_3^\sigma + \dots + u_{n-2} v_{n-1}^\sigma + u_{n-1} v_{n-2}^\sigma + u_n v_n^\sigma$$

if $n = 2r + 1$.

The unitary polar spaces of rank r formed from the totally isotropic subspaces of a non-degenerate hermitian form on $V(n, q)$ are denoted $H(2r-1, q)$ and $H(2r, q)$ according to whether $n = 2r$ or $n = 2r + 1$. The elements $f \in GL(n, q)$ with the property that

$$\beta(f(\mathbf{u}), f(\mathbf{v})) = \beta(\mathbf{u}, \mathbf{v})$$

for all $\mathbf{u}, \mathbf{v} \in V$ form a group under composition. This group is called the *unitary group* and is denoted $U(n, q)$. The *special unitary group* $SU(n, q)$ is isomorphic to the group $\{f \in U(n, q) \mid \det f = 1\}$. The *projective special unitary group*

$$PSU(n, q) = SU(n, q)/\{aI \mid a^{\sigma+1} = 1 \text{ and } a^n = 1\}$$

is a simple group unless $(n, q) = (2, 2)$, $(n, q) = (3, 2)$ or $(n, q) = (2, 3)$.

3.8 Orthogonal Spaces

The following lemma restricts the dimension of an anisotropic subspace of a vector space carrying a non-singular quadratic form.

Lemma 3.8.1. *Let Q be a non-singular quadratic form on $V(n, q) = V$. If the rank of V is at least* 3 *then V has an isotropic vector.*

Proof. Let $\mathbf{u}$ be any non-zero vector. Since the rank of V is at least 3, the rank of $\langle \mathbf{u} \rangle^\perp$ is at least 2 and we can choose a $\mathbf{v} \in \langle \mathbf{u} \rangle^\perp \setminus \langle \mathbf{u} \rangle$, $\mathbf{v} \neq \mathbf{0}$. Let β be the polar form of Q. For all $\lambda, \nu \in GF(q)$ we have that $\beta(\lambda \mathbf{u}, \nu \mathbf{v}) = 0$ and so by (QF1) and (QF2)

$$Q(\lambda \mathbf{u} + \nu \mathbf{v}) = \lambda^2 Q(\mathbf{u}) + \nu^2 Q(\mathbf{v}).$$

If q is even every element of $GF(q)$ is a square and we can choose λ and ν such that $\lambda^2 Q(\mathbf{u}) + \nu^2 Q(\mathbf{v}) = 0$.

If q is odd then, since the rank of V is at least 3, we can take a $\mathbf{w} \in \langle \mathbf{u}, \mathbf{v} \rangle^{\perp}$, $\mathbf{w} \neq \mathbf{0}$. The sets $\{\lambda^2 Q(\mathbf{u}) \mid \lambda \in GF(q)\}$ and $\{-\nu^2 Q(\mathbf{v}) - Q(\mathbf{w}) \mid \nu \in GF(q)\}$ both contain $(q+1)/2$ elements so there exists a λ and a ν such that

$$\lambda^2 Q(\mathbf{u}) = -\nu^2 Q(\mathbf{v}) - Q(\mathbf{w})$$

and with this choice $Q(\lambda \mathbf{u} + \nu \mathbf{v} + \mathbf{w}) = 0$. □

Let $(\mathbf{e_1}, \mathbf{f_1})$ be a hyperbolic pair.

Choose $\mathbf{e_1} = (1, 0, 0, \ldots, 0)$ and $\mathbf{f_1} = (0, 1, 0, \ldots, 0)$. Consider Q restricted to $\langle \mathbf{e_1}, \mathbf{f_1} \rangle$. It follows from $Q(\mathbf{e_1}) = 0$, $Q(\mathbf{f_1}) = 0$ and $\beta(\mathbf{e_1}, \mathbf{f_1}) = 1$ that Q in matrix form is

$$Q(\mathbf{u}) = \mathbf{u} \begin{pmatrix} 0 & c \\ 1-c & 0 \end{pmatrix} \mathbf{u} = u_1 u_2.$$

According to Lemma 3.8.1 an anisotropic subspace U has rank 0, 1 or 2. Therefore according to Theorem 3.5.4 we can extend $(\mathbf{e_1}, \mathbf{f_1})$ to a basis for $V(n, q)$.

If the rank of U is 0 then $n = 2r$ and the basis can be chosen such that

$$Q(\mathbf{u}) = u_1 u_2 + u_3 u_4 + \ldots + u_{n-1} u_n;$$

this is the *hyperbolic* case.

If the rank of U is 1 then $n = 2r + 1$ and the basis can be chosen such that

$$Q(\mathbf{u}) = u_1 u_2 + u_3 u_4 + \ldots + u_{n-2} u_{n-1} + a u_n^2;$$

this is the *parabolic* case. By applying the change of basis $u_i \mapsto a u_i$, for i even, and scaling we can assume that $a = 1$.

If the rank of U is 2 then $n = 2r + 2$ and the basis can be chosen such that

$$Q(\mathbf{u}) = u_1 u_2 + u_3 u_4 + \ldots + u_{n-3} u_{n-2} + f(u_{n-1}, u_n),$$

where f is a homogeneous quadratic polynomial irreducible over $GF(q)$ (note that Q restricted to $U = \langle e_{n-1}, e_n \rangle$ is $f(u_{n-1}, u_n)$, which must be irreducible since U is anisotropic); this is the *elliptic* case.

Note that this means that all conics are equivalent. Recall from Section 2.1 that a conic is the set of singular points of a non-singular quadratic form Q in a vector space of rank 3. The above implies that we can find a linear transformation $g \in GL(3, q)$ such that $Q(g(\mathbf{x})) = xy + z^2$.

The orthogonal polar spaces of rank r formed from the totally singular subspaces of a non-singular quadratic form on $V(n, q)$ are denoted $Q^+(2r - 1, q)$, $Q(2r, q)$ and $Q^-(2r + 1, q)$ according to whether the rank of the anisotropic space is 0, 1 or 2. They are called the hyperbolic space, the parabolic space and the elliptic space, respectively. The elements $f \in GL(n, q)$ with the property that

$$Q(f(\mathbf{u})) = Q(\mathbf{u})$$

for all $\mathbf{u} \in V$ form a group under composition. This group is called the *orthogonal group* and is denoted $O^+(2r-1,q)$, $O(2r,q)$ and $O^-(2r+1,q)$ accordingly. The derived subgroup of a group G is $\{g^{-1}h^{-1}gh \mid g,h \in G\}$ and in the case of the orthogonal groups we denote these as $\Omega^*(n,q)$. The group $P\Omega^*(n,q) = \Omega^*(n,q)/\{aI|aI \in \Omega^*(n,q)\}$ is a simple group for $n \geq 3$ unless $(n,r) = (4,2)$ or $(n,q) = (3,3)$.

3.9 Counting in polar spaces

The previous sections show that there are six polar spaces of a given rank r. In Table 3.9 we assign to each a label ϵ whose use will be clear from the following theorems.

Name	Polar Space	ϵ
Symplectic	$W(2r-1,q)$	0
Unitary	$U(2r-1,q)$	$-\frac{1}{2}$
Unitary	$U(2r,q)$	$\frac{1}{2}$
Hyperbolic	$Q^+(2r-1,q)$	-1
Parabolic	$Q(2r,q)$	0
Elliptic	$Q^-(2r+1,q)$	1

Table 3.1: The polar spaces of rank r

Theorem 3.9.1. *The number of points in a polar space of rank* 1 *is* $q^{1+\epsilon}+1$ *where* ϵ *is defined as in table 3.9.*

Proof. By Theorem 3.5.3 and Theorem 3.5.4, the ambient space of a polar space of rank 1 is the direct sum

$$L \oplus U$$

where L is a hyperbolic line and U is an anisotropic subspace. Let n_k be the number of points in a polar space ρ of rank 1 whose anisotropic space U has rank k. If x is a point of ρ then $x = x^\perp \cap \rho$ and the other hyperplanes that are incident with x meet ρ in a polar space of rank 1 with anisotropic space of rank $k-1$. Hence

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

and so $n_k = q^k(n_0 - 1) + 1$.

To determine n_0 we must take each type of polar space of rank r separately.

Symplectic: $V = L$ is of rank 2 and every vector is isotropic. Therefore there are $q+1$ points in the polar space and $\epsilon = 0$.

Unitary: By Section 3.7, we can choose a basis for V such that

$$\beta(\mathbf{u},\mathbf{v}) = u_1 v_2^\sigma + u_2 v_1^\sigma.$$

The point $\mathbf{u} = \langle (u_1,u_2)\rangle$ of the polar space satisfies

$$u_1 u_2^\sigma + u_2 u_1^\sigma = 0.$$

If $u_1 = 0$ then there is just one solution $\langle(0,1)\rangle$ and if not we can put $u_1 = 1$ and there are $\sqrt{q}$ points corresponding to the solutions of $u_2^\sigma + u_2 = 0$. Hence $n_0 = \sqrt{q}+1$ and $\epsilon = -\frac{1}{2}$ and $n_1 = q\sqrt{q}+1$, so $\epsilon = \frac{1}{2}$ in this case.

Orthogonal: By Section 3.8, we can choose a basis for V such that

$$Q(\mathbf{u}) = u_1 u_2.$$

The point $\mathbf{u} = \langle (u_1,u_2)\rangle$ of the polar space satisfies

$$u_1 u_2 = 0$$

and they are the points $\langle(1,0)\rangle$ and $\langle(0,1)\rangle$. Hence $n_0 = 2$ and $\epsilon = -1$, $n_1 = q+1$ and $\epsilon = 0$ in this case, and $n_2 = q^2+1$ and $\epsilon = 1$ in this case. □

Recall that the points of a polar space of rank r are the isotropic subspaces of rank one.

THEOREM 3.9.2. *The number of points in a polar space of rank r is*

$$(q^r-1)(q^{r+\epsilon}+1)/(q-1),$$

$q^{2r-1+\epsilon}$ *of which are not collinear to a given point.*

Proof. Let $F(r)$ be the number of points in a polar space of rank r and let $G(r)$ be the number of points not collinear with a given point. The proof goes by induction. We do not assume that $G(r)$ is independent of the given point, it will also follow by induction. Fix a point x of the polar space. Let us count the pairs (y,z) where $z \notin x^\perp$ and $y \in x^\perp \cap z^\perp$. By Corollary 3.5.5 and Corollary 3.5.6 $\langle x,z\rangle^\perp$ and $\langle x\rangle^\perp/\langle x\rangle$ are polar spaces of rank $r-1$. Choosing z first there are $G(r)F(r-1)$ pairs and choosing y first there are $qF(r-1)qG(r-1)$ pairs. Hence $G(r) = q^2 G(r-1)$. Theorem 3.9.1 says that $G(1) = q^{\epsilon+1}$ and so $G(r) = q^{2r-1+\epsilon}$. Now counting points of the polar space collinear and not collinear with x gives

$$F(r) = 1 + qF(r-1) + G(r) = q^{2r-1+\epsilon} + 1 + qF(r-1).$$

One can check that $F(r) = (q^r-1)(q^{r+\epsilon}+1)/(q-1)$ satisfies this recurrence and that $F(1) = 1 + q^{\epsilon+1}$, in accordance with Theorem 3.9.1. □

The maximal totally isotropic subspaces of a polar space of rank r have rank r and projective dimension $r-1$. The following theorem tells us how many there are.

THEOREM 3.9.3. *The number of subspaces of dimension $r-1$ in a polar space of rank r is*

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

Proof. Let $H(r)$ be the number of (totally isotropic) subspaces of dimension $r-1$ in a polar space of rank r and count pairs (x, U) where x is a point of U, a subspace of dimension $r-1$. Choosing x first, the number of pairs is $F(r)H(r-1)$, since $U \subset U^{\perp} \subset x^{\perp}$, and choosing U first, the number of pairs is $H(r)(q^r-1)/(q-1)$. Theorem 3.9.2 implies $H(r) = (q^{r+\epsilon}+1)H(r-1)$ and Theorem 3.9.1 implies $H(1) = q^{\epsilon+1}+1$. □

3.10 Exercises

11. Let Q be a non-singular quadratic form on $V(3,q)$ and let β be the symmetric bilinear form defined by $\beta(\mathbf{x},\mathbf{y}) = Q(\mathbf{x}+\mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})$. Let G be the graph whose vertices are the points of $PG(2,q)$ and where $\langle\mathbf{x}\rangle$ is joined to $\langle\mathbf{y}\rangle$ $(\mathbf{x} \neq \mathbf{y})$ if $\beta(\mathbf{x},\mathbf{y}) = 0$. Show that

1. G has diameter 2, degree $q+1$ and q^2+q+1 vertices.
2. if q is even then G can be extended to a graph of diameter 2, degree $q+1$ with q^2+q+2 vertices.
3. G can be extended to a graph of degree $q+1+2r$, diameter 2 with $q^2+q+1+r(q+1)$ vertices.

12. Let U be a (maximal totally isotropic) subspace of dimension $r-1$ of a polar space of rank r. Let P be a point not in U. Prove that the set of (totally isotropic) lines joining P to U form a subspace W of dimension $r-1$ and that $U \cap W$ is a hyperplane in both U and W.

13. Label the points of Figure 3.1 with points of $PG(3,2)$ so that the lines are totally isotropic with respect to the alternating form

$$\beta(\mathbf{u},\mathbf{v}) = u_1v_2 + v_1u_2 + u_3v_4 + v_3u_4.$$

What should we do to the labels if we wish the points of Figure 3.1 to be the points of $Q(4,2)$ defined with respect to the quadratic form

$$Q(\mathbf{u}) = u_1u_2 + u_3u_4 + u_5^2 \ ?$$

14. A *partial m-system* $\mathcal{M}$ of ρ, a polar space of rank r, is a set of subspaces of ρ of dimension m with the property that a subspace of ρ of rank $r-1$ containing an element of $\mathcal{M}$ is disjoint from any other element of $\mathcal{M}$.

Prove that $|\mathcal{M}| \leq q^{r+\epsilon}+1$.

[Hint: Let τ_i be the number of points of ρ, not contained in an element of $\mathcal{M}$, that are incident with exactly i totally isotropic subspaces of dimension $m+1$ containing an element of $\mathcal{M}$. By counting similar to that in Proposition 2.1.2 and repeated use of Corollary 3.5.5 and Corollary 3.5.6 expand the inequality

$$\sum_i (i-1-(x-1)/q^{m+1})^2 \tau_i \geq 0,$$

where $x = |\mathcal{M}|$.]

A partial m-system of size $q^{r+\epsilon}+1$ is called an m-system.

Using the inequality above show that a point of ρ that is not incident with an element of an m-system $\mathcal{M}$ is incident with precisely $q^{r-m-1+\epsilon}+1$ subspaces of ρ of dimension $m+1$ that contain an element of $\mathcal{M}$.

Chapter 4

Generalised quadrangles and inversive planes

4.1 Generalised quadrangles

A polar space of rank 1 contains only points and so has no geometrical structure.

A polar space ρ of rank 2 consists of points and lines. Let $\langle \mathbf{u} \rangle$ be a point and let $\langle \mathbf{v}, \mathbf{w} \rangle$ be a line of ρ such that $\mathbf{u} \notin \langle \mathbf{v}, \mathbf{w} \rangle$. Since $\langle \mathbf{u} \rangle^{\perp}$ is a hyperplane of the ambient space, there is a point of ρ in the intersection of $\langle \mathbf{u} \rangle^{\perp}$ and $\langle \mathbf{v}, \mathbf{w} \rangle$. Moreover there are no other points of $\langle \mathbf{v}, \mathbf{w} \rangle$ collinear with $\langle \mathbf{u} \rangle$ since ρ does not contain planes. So a polar space of rank 2 satisfies the following axioms of a generalised quadrangle.

A *generalised quadrangle* is an incident structure of points and lines with the following properties.

(GQ1) Every two points are incident with at most one line.

(GQ2) For every point P and a line l such that $P \notin l$ there exists a unique point $Q \in l$ collinear with P.

(GQ3) There is no point collinear with all the others.

The axioms of a generalised quadrangle are self-dual; the dual of a generalised quadrangle is also a generalised quadrangle. Hence for any result we prove the dual statement also holds.

The trivial examples of generalised quadrangles are the complete bipartite graphs and their duals, the grids.

Proposition 4.1.1. *If G is a generalised quadrangle with a line incident with at least 3 points or a point incident with at least 3 lines then there exist constants s*

and t such that every line is incident with $s+1$ points and every point is incident with $t+1$ lines.

Proof. By duality we can assume that G is a generalised quadrangle with a point P incident with at least 3 lines,. We will prove that such an s exists. The proof that such a t exists is similar.

By (GQ2), there is a bijection between the points on two skew (non-concurrent) lines, hence the number of points on these lines must be the same. Next we show that the number of points on two concurrent lines (say, e and f) not containing P, are the same. By (GQ2), there are exactly one line through P, meeting e and exactly one through P, meeting f. Hence there is a line through P which is skew to both e and f; so from the above e and f have the same number of points too.

It follows from the above argument that the lines not through P must have the same number of points. To show that each line incident with the same number of points, we only left to show that if f is a line through P then there is a line skew to f; hence (from the above) each line contains the same number of points. Let m be a line through P different from f. By (GQ3), there exists a point Q, not collinear with P and by (GQ2), there is a line g through Q meeting m. Again by (GQ2), g is skew to f.

□

We shall be only interested in finite generalised quadrangles here, so we restrict our attention to s and t finite.

The polar spaces of rank 2 are generalised quadrangles; their orders and isomorphisms are given in Table 4.1. There are other finite generalised quadrangles known, however either they have the same parameters as one of the polar spaces of rank 2 examples or they have order $(q-1,q+1)$ or $(q+1,q-1)$ where $q=p^h$ for some prime p.

Name	Polar Space	(s,t)	
Symplectic	$W(3,q)$	(q,q)	dual of $Q(4,q)$
Unitary	$U(3,q^2)$	(q^2,q)	dual of $Q^-(5,q)$
Unitary	$U(4,q^2)$	(q^2,q^3)	
Hyperbolic	$Q^+(3,q)$	$(q,1)$	a grid
Parabolic	$Q(4,q)$	(q,q)	dual of $W(3,q)$
Elliptic	$Q^-(5,q)$	(q,q^2)	dual of $U(3,q^2)$

Table 4.1: The polar spaces of rank 2.

PROPOSITION 4.1.2. *In a finite generalised quadrangle $\mathcal{G}$ of order (s,t) there are $(st+1)(s+1)$ points and $(st+1)(t+1)$ lines.*

Proof. Let l be a line of $\mathcal{G}$. There are $(s+1)t$ lines that meet l and by (GQ2) every point not on l is incident with exactly one of these lines. Hence the total number of points is $(s+1)ts+s+1$. The number of lines follows by duality. □

We say that two points P and Q are adjacent or collinear and write $P \sim Q$ if there is a line joining them. Likewise for two lines l and m, we write $l \sim m$ if they are concurrent.

If $P \not\sim Q$, then the number of points adjacent to both P and Q is $t+1$. When $P \sim Q$, the number of points adjacent to both P and Q are the number of points on the line $\langle P, Q\rangle$ and there are $s+1$ of them.

Proposition 4.1.3. *If there is a finite generalised quadrangle $\mathcal{G}$ of order (s,t) then $s+t$ divides $st(s+1)(t+1)$.*

Proof. Let the points of $\mathcal{G}$ be $\{P_i \,|\, i = 1, 2, \ldots, (st+1)(t+1)\}$ and define a matrix $A = (a_{ij})$ by

$$a_{ij} = \begin{cases} 1 & \text{if } P_i \sim P_j \text{ or } i = j, \\ 0 & \text{if } P_i \not\sim P_j. \end{cases}$$

Let $A^2 = (b_{ij})$. Then

$$b_{ii} = \sum_j a_{ij}a_{ji} = (t+1)s+1,$$

the number of the points adjacent with the point P_i including P_i. If $P_i \not\sim P_k$ then

$$b_{ik} = \sum_j a_{ij}a_{jk} = t+1,$$

the number of points adjacent to both P_i and P_k. If $P_i \sim P_k$ then

$$b_{ik} = \sum_j a_{ij}a_{jk} = s+1,$$

the number of points on the line joining P_i and P_k. Hence

$$A^2 - (s-t)A - stI = (t+1)J,$$

where I is the identity matrix and J is the all-one matrix. The matrix J has rank 1 so its null-space has co-rank 1. For any $\mathbf{u}$ in this null-space

$$(A - sI)(A + tI)\mathbf{u} = 0,$$

so $\mathbf{u}$ is an eigenvector of A with eigenvalue s or $-t$. The all-one vector $\mathbf{j}$ is also an eigenvector of A with eigenvalue $s(t+1)+1$. Let x be the multiplicity of the eigenvalue s and let y be the multiplicity of the eigenvalue $-t$, so x and y are positive integers satisfying $x + y = (st + 1)(s + 1) - 1$. The trace $\mathrm{Tr}(A)$ of the matrix A is equal to the sum of its eigenvalues and so

$$(st + 1)(s + 1) = \mathrm{Tr}(A) = s(t + 1) + 1 + sx - ty.$$

Therefore $y = s^2(st+1)/(s+t)$ and $x = st(s+1)(t+1)/(s+t)$, both are integers and so the divisibility holds. □

Proposition 4.1.4. *Suppose there is a finite generalised quadrangle $\mathcal{G}$ of order (s, t). If $t > 1$ then $s \leq t^2$ and if $s > 1$ then $t \leq s^2$.*

Proof. Let P and Q be non-adjacent points, and let x_n be the number of points R adjacent to neither P nor Q for which there are exactly n points adjacent to P, Q and R.

Every point R not adjacent to P nor Q is adjacent to some number of points adjacent to both P and Q so

$$\sum x_n = (st + 1)(s + 1) - 2(t + 1)s + t + 1 = s^2t - st - s + t.$$

The $t + 1$ points adjacent with both P and Q are adjacent to $s(t - 1)$ points that are not adjacent to either P nor Q so counting (S, R) where $S \sim P$, $S \sim Q$ and $S \sim R$

$$\sum nx_n = s(t + 1)(t - 1).$$

Counting triples (S_1, S_2, R) where $S_i \sim P$, $S_i \sim Q$ and $S_i \sim R$ and $S_1 \neq S_2$ gives

$$\sum n(n - 1)x_n = (t + 1)t(t - 1).$$

Clearly

$$\sum \left(n - \frac{s(t^2 - 1)}{(s^2t - st - s + t)}\right)^2 x_n \geq 0.$$

Expanding this and using the equations above yields the inequality

$$(s - 1)t(s^2 - t) \geq 0.$$

Hence if $s > 1$ then $t \leq s^2$ and the other inequality follows by duality. □

4.2 Dualities, polarities and ovoids

A *duality of a generalised quadrangle* π is a map from the points to the lines that preserves incidence. Note that a generalised quadrangle which possesses a duality is of order (s, s) for some s. The map π induces a map π^* from the lines to the points given by

$$\pi^*(\langle P, Q\rangle) = \pi(P) \cap \pi(Q).$$

If $\pi\pi^*$ is the identity map we say the duality π is a *polarity*.

Proposition 4.2.1. *Let π be a polarity and let P be a point of a generalised quadrangle $\mathcal{G}$ for which $P \not\sim \pi(P)$. The point Q incident with the line $\pi(P)$ and adjacent to P is the point $\pi^*(m)$ where m is the line joining P and Q.*

Proof. Now $Q \in \pi(P)$ if and only if $\pi^*(\pi(P)) \in \pi(Q)$ if and only if $P \in \pi(Q)$. Moreover $P \sim Q$ if and only if $\pi(P) \sim \pi(Q)$, so $\pi(Q)$ is incident with P and meets $\pi(P)$. Therefore $\pi(Q) = m$ and $Q = \pi^*(m)$. Note $Q \in \pi(Q)$. □

A point Q is called *absolute* with respect to a polarity π if $Q \in \pi(Q)$. A line l is called *absolute* with respect to a polarity π if $\pi^*(l) \in l$.

Proposition 4.2.2. *Let $\mathcal{G}$ be a generalised quadrangle with a polarity π. Every line is incident with exactly one absolute point and every point is incident with exactly one absolute line.*

Proof. The statements are dual so it is sufficient to prove the first.

Let l be an absolute line, $\pi^*(l) \in l$. If $P \in l$ then $\pi^*(l) \in \pi(P)$ and if $\pi(P) \neq l$ then $\pi^*(l) = \pi(P) \cap l$. If P is absolute then $P \in \pi(P)$ and so $P = \pi^*(l)$.

Now assume that l is not an absolute line. Proposition 4.2.1 implies that the point $P \in l$ collinear with $\pi^*(l)$ is absolute. If $Q \in l$ is an absolute point then $\pi^*(l) \in \pi(Q)$ and $Q \in \pi(Q)$, so $Q = P$. □

An *ovoid of a generalised quadrangle* $\mathcal{O}$ is a set of points with the property that every line is incident with exactly one point of $\mathcal{O}$.

Proposition 4.2.2 tells us that the absolute points of a polarity form an ovoid.

Theorem 4.2.3. *If $\mathcal{G}$ is a generalised quadrangle of order (s, s) with a polarity π, then $2s$ is a square.*

Proof. Let the points of $\mathcal{G}$ be $\{P_i \,|\, i = 1, 2, \dots, (st+1)(t+1)\}$ and define a matrix $D = (d_{ij})$ by

$$d_{ij} = \begin{cases} 1 & \text{if } P_i \in \pi(P_j), \\ 0 & \text{if } P_i \notin \pi(P_j). \end{cases}$$

Let $D^2 = (e_{ij})$.

$$e_{ii} = \sum_j d_{ij} d_{ji} = s + 1,$$

the number of the lines incident with the point P_i.

$$e_{ik} = \sum_j d_{ij} d_{jk} = a_{ij},$$

where a_{ij} is defined as in Proposition 4.1.3. Hence $D^2 = A + sI$. In the proof of Proposition 4.1.3 we saw that A had eigenvalues $s^2 + s + 1$, s and $-s$ with multiplicities 1, $s(s+1)^2/2$ and $s(s^2+1)/2$ and so D^2 has eigenvalues $(s+1)^2$, $2s$ and 0 with the same multiplicities. The matrix D has eigenvalues $\pm(s+1)$, $\pm\sqrt{(2s)}$ and 0 with total multiplicities 1, $s(s+1)^2/2$ and $s(s^2+1)/2$. Now D has constant row sum $s+1$, so $s+1$ is certainly an eigenvalue. Let u and w be the multiplicities of the eigenvalues $\sqrt{(2s)}$ and $-\sqrt{(2s)}$ respectively, so $u+w = s(s+1)^2/2$. The trace of D is equal to the number of absolute points of π. These points form an ovoid so there are s^2+1 of them. The sum of the eigenvalues of D counted with multiplicity is equal to the trace of D and so we have

$$s^2+1 = s+1+u\sqrt{2q}-(s(s+1)^2/2-u)\sqrt{(2s)},$$

which implies $u = s(s+1)^2/4 + s(s-1)/(2\sqrt{(2s)})$. Since u is an integer, $2s$ is a square. □

PROPOSITION 4.2.4. *If the generalised quadrangle $\mathcal{G}$ is a polar space of rank 2 that contains a polar space ρ of rank 1 as a hyperplane H of the ambient space then the points of ρ are an ovoid of $\mathcal{G}$.*

Proof. Every line of $\mathcal{G}$ is a totally isotropic line of the ambient space which meets the hyperplane H in a point of ρ. □

In Section 3.3 we constructed large graphs of degree $q+1$ and diameter 2 from a polarity of $PG(2,q)$.

PROPOSITION 4.2.5. *Let $\mathcal{G}$ be a generalised quadrangle of order (s,s) with a polarity π. The graph whose vertices are the points of $\mathcal{G}$ and whose vertices P and $Q \neq P$ are joined by an edge iff $P \in \pi(Q)$, has degree $s+1$ and diameter 3.*

Proof. If P is an absolute point then $P \in \pi(P)$ and the degree of the vertex in the graph corresponding to P is s and if P is not an absolute point then it has degree $s+1$.

Let $Q \neq P$ be such that $P \notin \pi(Q)$ then by (GQ2) there exists an $R \in \pi(Q)$ such that $P \sim R$. Let l be the line joining P and R. In the graph there is a path $P - \pi^*(l) - R - Q$. □

4.3 The symplectic generalised quadrangle

Given a line of $PG(3,q)$

$$\langle (x_0,x_1,x_2,x_3),(y_0,y_1,y_2,y_3)\rangle$$

the *Plücker coordinates* are defined as

$$p_{ij} = x_i y_j - x_j y_i.$$

The Plücker coordinates are independent of the points on the line chosen

$$x_i(y_j + \lambda x_j) - x_j(y_i + \lambda x_i) = x_i y_j - x_j y_i = p_{ij},$$

and satisfy

$$p_{03}p_{12} + p_{13}p_{20} + p_{01}p_{23} = 0. \tag{4.1}$$

The *Klein correspondence* takes a line of $PG(3,q)$ to a point of $Q^+(5,q)$ by mapping the line of $PG(3,q)$ to the point $\langle p_{03}, p_{12}, p_{13}, p_{20}, p_{01}, p_{23} \rangle$ which lies on the quadric (4.1).

Let β be the non-degenerate alternating form defined by

$$\beta(\mathbf{x}, \mathbf{y}) = x_0 y_1 - y_1 x_1 + x_2 y_3 - y_2 x_3.$$

The lines of the $W(3,q)$ that this form defines satisfy $p_{01} + p_{23} = 0$, so the Klein correspondence takes a line of $W(3,q)$ to a point of $Q(4,q)$ defined by the quadratic form

$$p_{03}p_{12} + p_{13}p_{20} - p_{01}^2 = 0.$$

THEOREM 4.3.1. *The generalised quadrangle* $W(3,q)$ *is the dual of* $Q(4,q)$.

Proof. Let $\langle \mathbf{x}, \mathbf{y} \rangle$, $\langle \mathbf{x}, \mathbf{z} \rangle$ and $\langle \mathbf{x}, \mathbf{y}+\lambda\mathbf{z} \rangle$ be three concurrent lines of $W(3,q)$. Define the point of $Q(4,q)$

$$P_{\mathbf{x},\mathbf{y}} := \langle (p_{03}, p_{12}, p_{13}, p_{20}, p_{01}) \rangle.$$

The points $P_{\mathbf{x},\mathbf{y}}$, $P_{\mathbf{x},\mathbf{z}}$ and $P_{\mathbf{x},\mathbf{y}+\lambda\mathbf{z}}$ are collinear since

$$P_{\mathbf{x},\mathbf{y}+\lambda\mathbf{z}} = P_{\mathbf{x},\mathbf{y}} + \lambda P_{\mathbf{x},\mathbf{z}}.$$

Therefore the line $\langle P_{\mathbf{x},\mathbf{y}}, P_{\mathbf{x},\mathbf{z}} \rangle$ is totally isotropic. Hence if two lines are concurrent in $W(3,q)$ their images under the Klein correspondence are collinear in $Q(4,q)$. □

THEOREM 4.3.2. *If* q *is even then* $W(3,q)$ *is self-dual.*

Proof. Let the generalised quadrangle $Q(4,q)$ be defined by the quadratic form

$$Q(\mathbf{x}^*) := x_0^* x_1^* + x_2^* x_3^* + (x_4^*)^2,$$

where $x_0^* = p_{03}$, $x_1^* = p_{12}$, $x_2^* = p_{13}$, $x_3^* = p_{20}$ and $x_4^* = p_{01}$. Assuming q is even, two points $\langle \mathbf{x}^* \rangle$ and $\langle \mathbf{y}^* \rangle$ of $Q(4,q)$ are collinear iff

$$Q(\mathbf{x}^* + \mathbf{y}^*) - Q(\mathbf{x}^*) - Q(\mathbf{y}^*) = x_0^* y_1^* + x_1^* y_0^* + x_2^* y_3^* + x_3^* y_2^* = \beta(\mathbf{x}^*, \mathbf{y}^*) = 0.$$

Therefore the map

$$\langle (x_0^*, x_1^*, x_2^*, x_3^*, x_4^*) \rangle \mapsto \langle (x_0^*, x_1^*, x_2^*, x_3^*) \rangle,$$

from the points of $Q(4,q)$ to the points of $W(3,q)$, preserves collinearity. Theorem 4.3.1 implies that $Q(4,q)$ is the dual of $W(3,q)$ and so we conclude that $W(3,q)$ is self-dual. □

Let q be even for the rest of this section.

Let $\mathbf{x} = \langle (x_0, x_1, x_2, x_3) \rangle$ be a point of $W(3, q)$ defined from the alternating form

$$\beta(\mathbf{x}, \mathbf{y}) = x_0 y_1 + y_0 x_1 + x_2 y_3 + y_2 x_3.$$

Let $\mathbf{y} = \langle (0, x_2, 0, x_0) \rangle$ and $\mathbf{z} = \langle (0, x_3, x_0, 0) \rangle$. The lines $\langle \mathbf{x}, \mathbf{y} \rangle$ and $\langle \mathbf{x}, \mathbf{z} \rangle$ are totally isotropic (and hence lines of $W(3, q)$), distinct if $x_0 \neq 0$, and the other $q - 1$ totally isotropic lines containing $\mathbf{x}$ are $\langle \mathbf{x}, \mathbf{y} + \lambda \mathbf{z} \rangle$. The Plücker coordinates $\langle (x_0^*, x_1^*, x_2^*, x_3^*) \rangle$ of these lines are given by $P_{\mathbf{x},\mathbf{y}} = \langle (x_0^2, x_2^2, x_0 x_1 + x_2 x_3, 0) \rangle$, $P_{\mathbf{x},\mathbf{z}} = \langle (0, x_0 x_1 + x_2 x_3, x_3^2, x_0^2) \rangle$ and the formula from Theorem 4.3.1

$$P_{\mathbf{x},\mathbf{y}+\lambda\mathbf{z}} = P_{\mathbf{x},\mathbf{y}} + \lambda P_{\mathbf{x},\mathbf{z}}.$$

According to Theorem 4.3.1 these points are collinear in the dual quadrangle so we can calculate the Plücker coordinates

$$\langle (p_{03}^*, p_{12}^*, p_{13}^*, p_{20}^*) \rangle = \langle (x_0^4, x_1^2 x_0^2, x_2^2 x_0^2, x_3^2 x_0^2) \rangle = \langle (x_0^2, x_1^2, x_2^2, x_3^2) \rangle$$

of the line that joins them.

If $x_0 = 0$ then at least one of the other x_i must be non-zero. If $x_1 \neq 0$ then put $\mathbf{y} = \langle (x_3, 0, x_1, 0) \rangle$ and $\mathbf{z} = \langle (x_2, 0, 0, x_1) \rangle$. If $x_2 \neq 0$ then put $\mathbf{y} = \langle (x_2, 0, 0, x_1) \rangle$ and $\mathbf{z} = \langle (0, x_2, 0, x_0) \rangle$. If $x_3 \neq 0$ then put $\mathbf{y} = \langle (0, x_3, x_0, 0) \rangle$ and $\mathbf{z} = \langle (x_3, 0, x_1, 0) \rangle$. In all cases the Plücker coordinates of the line $\langle P_{\mathbf{x},\mathbf{y}}, P_{\mathbf{x},\mathbf{z}} \rangle$ are the same.

Let π^* be the duality from the lines of $W(3, q)$ to the points of $W(3, q)$ given by

$$\langle \mathbf{x}, \mathbf{y} \rangle \mapsto P_{\mathbf{x},\mathbf{y}}^{\sigma/2} = \langle (p_{03}^{\sigma/2}, p_{12}^{\sigma/2}, p_{13}^{\sigma/2}, p_{20}^{\sigma/2}) \rangle,$$

where σ is an automorphism of $GF(q)$. The map π from points to lines induced by the incidence preserving π^* is then

$$\langle \mathbf{x} \rangle \mapsto \langle P_{\mathbf{x},\mathbf{y}}^{\sigma/2}, P_{\mathbf{x},\mathbf{z}}^{\sigma/2} \rangle = \langle P_{\mathbf{x}^{\sigma/2},\mathbf{y}^{\sigma/2}}, P_{\mathbf{x}^{\sigma/2},\mathbf{z}^{\sigma/2}} \rangle.$$

and so

$$\pi^* \pi(\langle \mathbf{x} \rangle) = \langle ((x_0^{\sigma^2/4})^2, x_1^{\sigma^2/2}, x_2^{\sigma^2/2}, x_3^{\sigma^2/2}) \rangle.$$

Hence π is a polarity if and only if $x^{\sigma^2} = x^2$ for all $x \in GF(q)$. To find such an automorphism it is necessary and sufficient that $q = 2^{2h+1}$ and we take $\sigma : x \mapsto x^{2^{h+1}} = x^{\sqrt{2q}}$.

The duality π takes the points

$$\langle (1, z, y, x) \rangle \mapsto \langle (1, y^\sigma, (xy + z)^{\sigma/2}, 0), (0, (xy + z)^{\sigma/2}, x^\sigma, 1) \rangle$$

and so the absolute points satisfy $y = (xy + z)^{\sigma/2} + x^{\sigma+1}$ and hence

$$z = xy + y^\sigma + x^{\sigma+2}.$$

The point $\langle(0,1,0,0)\rangle$ is also an absolute point since

$$\pi(\langle(0,1,0,0)\rangle) = \langle(0,0,1,0),(0,1,0,0)\rangle.$$

Proposition 4.2.2 implies that

$$\{\langle(0,1,0,0)\rangle\} \cup \{\langle(1, xy + y^{\sigma} + x^{\sigma+2}, y, x)\rangle \mid x, y \in GF(q)\}$$

is an ovoid of $W(3, 2^{2h+1})$. This ovoid was discovered by Tits and its stabiliser group is a simple group called the Suzuki group, which is denoted $Sz(2^{2h+1})$.

4.4 Inversive planes

Theorem 3.9.2 implies that the polar space of rank one $Q^-(3,q)$ consists of q^2+1 points of the ambient space $PG(3,q)$. Each point P of this elliptic quadric is the unique point on its tangent plane T_P since the polar space contains no lines. The plane sections consist of either a single point or a conic $Q(2,q)$ and the lines of the ambient space are incident with 0, 1 or 2 points of the elliptic quadric.

An *ovoid of* $PG(3,q)$ $\mathcal{O}$ is a set of points with the following properties.

(O1) Any line is incident with at most two points of $\mathcal{O}$.

(O2) For any $P \in \mathcal{O}$, the union of all lines l with $l \cap \mathcal{O} = \{P\}$ is a plane.

The elliptic quadrics $Q^-(3,q)$ are examples of ovoids of $PG(3,q)$.

It follows immediately from the axioms that the plane sections of an ovoid consists of either a single point or an oval, and so an ovoid consists of q^2+1 points. The following theorem of Barlotti and Panella follows from repeated use of Segre's Theorem 2.2.1.

Theorem 4.4.1. *If q is odd then an ovoid of $PG(3,q)$ is an elliptic quadric.*

However when q is even the following theorem implies that we have already seen that there is another example, the Tits ovoid; there are no other examples known.

Theorem 4.4.2. *If q is even then an ovoid of $W(3,q)$ is an ovoid of $PG(3,q)$.*

Proof. We have to show that the lines of $PG(3,q)$ that are not totally isotropic meet an ovoid $\mathcal{O}$ of $W(3,q)$ in 0 or 2 points. Let l be a line such that $l \cap \mathcal{O} \neq \emptyset$. For any point $P \in l \cap \mathcal{O}$ the intersection $P^{\perp} \cap \mathcal{O} = \{P\}$ and hence $l^{\perp} \cap \mathcal{O} = \emptyset$. Hence at most half of the $q^2(q^2+1)$ non-totally isotropic lines are incident with points of $\mathcal{O}$. Let τ_i be the number of non-totally isotropic lines lines that meet $\mathcal{O}$ in $i > 0$ points. Then

$$\sum_{i \geq 1} \tau_i \leq q^2(q^2+1)/2.$$

Counting pairs (P, l) where P is a point of $\mathcal{O}$ and l is a non totally isotropic line gives

$$\sum_{i \geq 1} i\tau_i = q^2(q^2+1),$$

and counting triples (P, Q, l) where $P \in l$ and $Q \in l$ are points of $\mathcal{O}$ and l is a non totally isotropic line gives

$$\sum_{i \geq 1} i(i-1)\tau_i = q^2(q^2+1).$$

These equations imply that

$$\sum_{i \geq 1} (i-2)(i-1)\tau_i \leq 0$$

and hence that $\tau_i = 0$ unless $i = 1$ or 2. And

$$\sum_{i=1}^{2} i(i-2)\tau_i = 0$$

implies $\tau_1 = 0$. □

For q odd, we have the following theorem. For a proof see [14].

Theorem 4.4.3. *If q is odd then there is no ovoid of $W(3, q)$.*

The known ovoids of $W(3, q)$ and of $PG(3, q)$, for q even are the elliptic quadric $Q^-(3, q)$ and the Suzuki-Tits ovoid (which exists when $q = 2^{2e+1}$).

Consider the incidence structure whose points are the points of an ovoid $\mathcal{O}$ of $PG(3, q)$ and whose "circles“ are the subsets of points on an oval section of $\mathcal{O}$. If we take a point P of $\mathcal{C}$, an oval section and another point $Q \notin \mathcal{C}$, then q of the $q+1$ planes containing the line $\langle P, Q \rangle$ contain one point of $\mathcal{C} \setminus \{P\}$ and the remaining oval section $\mathcal{D}$ meets $\mathcal{C}$ in the single point P. Hence this incidence structure satisfies the following axioms.

An *inversive plane* is an incidence structure of points and circles which satisfy the following axioms.

(I1) Three distinct points are incident with exactly one circle.

(I2) If P and Q are two points and c is a circle incident with P and not incident with Q then there is exactly one circle d incident with P and Q such that $c \cap d = \{P\}$.

(I3) There are at least two circles and every circle has at least three points.

For any point P of an inversive plane the points $\neq P$ and the circles incident with P form an affine plane. If the inversive plane is finite then all these affine planes have the same order; this integer is the order of the inversive plane. An inversive plane of order n has n^2+1 points and $n(n^2+1)$ circles; every circle is incident with $n+1$ points and any two points are incident with $n+1$ circles.

As we have seen the points and the non-tangent plane sections of an ovoid of $PG(3,q)$ form an inversive plane. Dembowski [6] proved the following partial converse:

THEOREM 4.4.4. Every inversive plane of even order n is isomorphic to the incidence structure of points and plane sections of an ovoid in $PG(3,n)$. □

Note Dembowski's theorem implies that if there is an inversive plane of even order n then n is a power of 2. There are no inversive planes known of odd order that do not come from elliptic quadrics of $PG(3,q)$.

4.5 Exercises

15. Prove that any pair of points of an inversive plane of order n are contained in $n+1$ circles.

Let P be a point in an inversive plane of even order n and $\mathcal{C}$ be a circle not containing P. Prove that there is a point $Q \neq P$ such that the $n+1$ circles containing P and tangent to $\mathcal{C}$ are the $n+1$ circles that contain P and Q.

16. Let $\mathcal{O}$ be an oval of $\pi = PG(2,q)$ embedded in a $PG(3,q)$. Let $T_2(\mathcal{O})$ be the incidence structure whose points are
(a) the q^3 points of $PG(3,q) \setminus \pi$,
(b) the q^2+q planes that meet π in a tangent of $\mathcal{O}$,
(c) a point (∞)
and whose lines are
(i) the lines of $PG(3,q) \setminus \pi$ incident with a point of $\mathcal{O}$.
(ii) the $q+1$ points of $\mathcal{O}$,
and where incidence is inherited from the incidence of $PG(3,q)$ and the point (∞) is incident with all the lines of type (ii).

Prove that $T_2(\mathcal{O})$ is a generalised quadrangle.

17. Let $\mathcal{O}$ be an ovoid of $Q(4,q)$, defined by the quadratic form $x_0x_1+x_2x_3+x_4^2$, that contains the point $\langle(1,0,0,0,0)\rangle$. Prove that if $\langle(-y^2-cx,1,c,x,y)\rangle$ and $\langle(-y^2-dx,1,d,x,y)\rangle$ are points of $\mathcal{O}$ then $c=d$. Hence argue that O is necessarily of the form

$$\{\langle(1,0,0,0,0)\rangle\} \cup \{\langle(-y^2-xg(x,y),1,g(x,y),x,y)\rangle : x,y \in GF(q)\}$$

for some polynomial $g(x,y)$.

Suppose that $g(x, y) = -mx^\sigma$ where σ is an automorphism of $GF(q)$. Prove that $\mathcal{O}$ is an ovoid of $Q(4, q)$ if and only if m is a non-square in $GF(q)$.

Chapter 5

Appendix

5.1 Solutions to the exercises

Chapter 1

1. The isomorphism is $x \mapsto x+1$.

2. Let the perspective point be the origin $(0,0)$.
Let $A=(x_1,y_1)$, $A'=(\lambda_1 x_1, \lambda_1 y_1)$, $B=(x_2,y_2)$, $B'=(\lambda_2 x_2,\lambda_2 y_2)$, $C=(x_3,y_3)$ and $C'=(\lambda_3 x_3,\lambda_3 y_3)$. Then the point $p_{12}:=AB$ is the point

$$\left(\frac{\lambda_1(1-\lambda_2)x_1+\lambda_2(\lambda_1-1)x_2}{\lambda_1-\lambda_2},\frac{\lambda_1(1-\lambda_2)y_1+\lambda_2(\lambda_1-1)y_2}{\lambda_1-\lambda_2}\right),$$

and it remains to check that

$$(\lambda_1-\lambda_2)(1-\lambda_3)p_{12}+(\lambda_1-\lambda_3)(1-\lambda_2)p_{13}+(\lambda_2-\lambda_3)(1-\lambda_1)p_{23}=0.$$

3. The number of points of $PG(3,q)$ is

$$\begin{bmatrix}4\\1\end{bmatrix}_q=(q^4-1)/(q-1)=q^3+q^2+q+1.$$

The number of points on a line is $q+1$. Each point is incident with a unique line so the number of lines in a spread is q^2+1. The intersection of two subspaces of rank 1 in $V(2,q^2)$ is the zero-vector.

4. Let P be a point of $\mathcal{K}$. There are $q+1$ lines incident with P that contain at most $r-1$ other points of $\mathcal{K}$, so

$$|\mathcal{K}|\leq(r-1)(q+1)+1.$$

If we have equality then every line is incident with 0 or r points of $\mathcal{K}$. If $r \leq q$ then we can take a point Q not in $\mathcal{K}$. There are $(rq - q + r)/r$ lines incident with Q that are incident with r points of $\mathcal{K}$. Hence r divides q.

Chapter 2

5. A conic is a set of points of $PG(2, q)$ that are the zeros of a quadratic form

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz.$$

The condition that a point P is contained in a conic imposes a linear condition on the coefficients of the quadratic form. Five points determines the quadratic point up to a scalar factor and hence determines the conic.

There are $q^2 + q + 1$ points in $PG(2, q)$, $(q^2 + q + 1)(q^2 + q)$ ordered pairs of points, $(q^2 + q + 1)(q^2 + q)q^2$ ordered triples of points no three collinear, $(q^2 + q + 1)(q^2 + q)q^2(q^2 - 2q + 1)$ ordered 4-tuples of points no three collinear, $(q^2 + q + 1)(q^2 + q)q^2(q^2 - 2q + 1)(q^2 - 5q + 6)$ ordered 5-tuples of points no three collinear. The number of conics is then

$$\frac{(q^2 + q + 1)(q^2 + q)q^2(q^2 - 2q + 1)(q^2 - 5q + 6)}{(q + 1)q(q - 1)(q - 2)(q - 3)}.$$

6. The dual of a linear MDS $[n, n-d+1, d]$-code $\mathcal{C}$ is a linear $[n, d-1, D]$-code for some D. We have to show that the minimum distance is $n - d + 2$. The minimum distance is equal to the minimum weight, so suppose there is a vector $\mathbf{x}$ in the dual code of weight at most $n - d + 1$. The q^{n-d+1} vectors of $\mathcal{C}$ differ on a set of $n - d + 1$ positions that include those where $\mathbf{x}$ has a non-zero entry. However since $\mathbf{x}$ is in the dual code the vectors of $\mathcal{C}$ satisfy a linear relation in these $n - d + 1$ positions, a contradiction.

7. The number of polynomials of degree at most $q - 1$ is q^q and so is the number of functions from $GF(q)$ to $GF(q)$. If two polynomials f and g of degree at most $q - 1$ define the same function from $GF(q)$ to $GF(q)$ then $f - g$ is zero on all elements of $GF(q)$ and has degree at most $q - 1$ and hence $f \equiv g$.

8. Since q is even $q + 1$ is odd so every point not in the oval is incident with a tangent. Every point in the oval is incident with a tangent, so the set of $q + 1$ tangents cover all points. A set of k lines covers at most $kq+1$ points with equality if and only if the lines are concurrent.

9.

1. If $\mathcal{H}$ contains the points $\langle(1, x, a)\rangle$ and $\langle(1, x, b)\rangle$ then, since the points $\langle(0, 0, 1)\rangle$, $\langle(1, x, a)\rangle$ and $\langle(1, x, b)\rangle$ are collinear, $a = b$.

2. The lines $Z = aX$ are incident with $\langle(0, 1, 0)\rangle$ and are therefore incident with distinct points of the set $\{\langle(1, x, f(x))\rangle \mid x \in GF(q)\}$, so $a = f(x)$ has a unique solution, i.e. f is a permutation.

3. The lines $Z + bY + (f(s) + sb)X = 0$ contain the point $\langle 1, s, f(s)\rangle$ and meet a point of the set $\{\langle 1, x+s, f(x+s)\rangle \mid x \in GF(q)^*\}$ whenever $f(x+s) + bx + f(s) = 0$. Hence $b = (f(x+s) + f(s))/x$ should have a unique solution for each b.

4. If f satisfies all the properties then every line is incident with 0 or 2 points of $\mathcal{H}$.

10. By choosing a basis we can assume that an arc contains the $r+2$ points

$$\langle(1,0,\ldots,0)\rangle, \langle(0,1,0,\ldots,0)\rangle, \langle(0,\ldots,0,1)\rangle, \text{ and } \langle(1,1,\ldots,1)\rangle.$$

Let $P = (p_i)$ be any other point. Since $r \geq q-1$ there must be two positions say j and k such that $p_j = p_k$ or a position l such that $p_l = 0$. In the first case the hyperplane $X_j = X_k$ contains $r+1$ points and in the second case the hyperplane $X_l = 0$ contains $r+1$ points.

Chapter 3

11.

1. $$\langle \mathbf{x}\rangle^\perp = \{\langle \mathbf{y}\rangle \mid \beta(\mathbf{x}, \mathbf{y}) = 0\}$$
   is a line of $PG(2,q)$ so $\mathbf{x}^\perp$ and $\mathbf{z}^\perp$ have a common point.

2. Add a vertex to G and join it to all the singular points. Then the degree remains $q+1$ and since every point lies on a tangent (q is even) it is at most at distance 1 in the graph to a singular point.

3. Add a vertex to G corresponding to each singular point and join it to the neighbours of the singular point. Then the degree increases by 2, there are at most two tangents incident with a given point. The diameter is still 2. Repeat this process.

12. The subspace $P^\perp$ is a hyperplane in the ambient space and so $X = P^\perp \cap U$ is subspace of rank $r-2$ of isotropic vectors perpendicular to P. The totally isotropic lines joining a point of X to P is the subspace $W = P \oplus X$. Clearly $X = W \cap U$ is a hyperplane of both W and U.

13. A solution would be Figure 5.1.

If we put $u_1u_2 + u_3u_4$ as the fifth coordinate to each point $\langle u_1, u_2, u_3, u_4\rangle$ we get $Q(4,2)$.

14. We can assume that $m < r-1$ since otherwise the bound follows from Theorem 3.9.2.

Let N be the set of points that are incident with an element of $\mathcal{M}$.

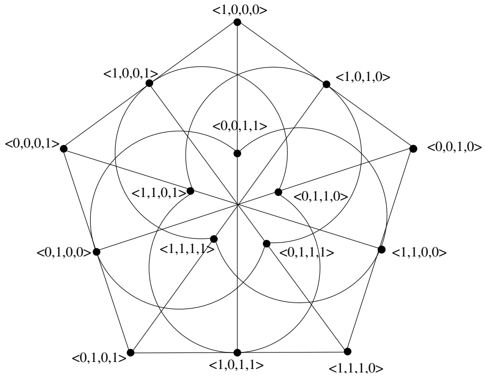

Figure 5.1: $W(3,2)$

Let τ_i be the number of points of $\rho \setminus N$ that are incident with exactly i totally isotropic subspaces of dimension $m+1$ containing an element of $\mathcal{M}$. Clearly

$$\sum \tau_i = F(r) - x(q^{m+1}-1)/(q-1),$$

where $x = |\mathcal{M}|$.

Corollary 3.5.6 used repeatedly implies that the number of totally isotropic subspaces of dimension $m+1$ containing a fixed subspace of dimension m is the number of points in a polar space of rank $r-m-1$. Counting pairs (P,U) where P is a point of $\rho \setminus N$ and U is a totally isotropic subspace of dimension $m+1$ containing an element of $\mathcal{M}$ gives

$$\sum i\tau_i = xq^{m+1}F(r-m-1),$$

where $F(r)$ is the number of points in a polar space of rank r as in Theorem 3.9.2.

Corollary 3.5.5 used repeatedly implies that the number of pairs totally isotropic subspaces of dimension $m+1$ containing fixed subspace of dimension m is the number of points in a polar space of rank $r-m-1$. Counting triples (P,T,U) where P is a point of $\rho \backslash N$ and U and V are totally isotropic subspaces of dimension $m+1$ containing an element of $\mathcal{M}$ gives

$$\sum i(i-1)\tau_i = x(x-1)F(r-m-1),$$

Now expand

$$\sum (i-1-(x-1)/q^{m+1})^2\tau_i \geq 0$$

and conclude that

$$-(x+q^r-1)(x-q^{r+\epsilon}-1) \geq 0.$$

Chapter 4

15. The quotient geometry at any point P is an affine plane $\mathcal{A}_P$ of a fixed order n. So there are n^2+1 points. If we fix P and Q then there are $n+1$ lines incident with Q in the plane $\mathcal{A}_P$ so there are $n+1$ circles incident with both P and Q. A circle c that does not contain P is an oval in the affine plane $\mathcal{A}_P$. If n is even then this oval has a nucleus. If this nucleus lies on the line at infinity the circles that joining a single point of c to P coincide only at P, which would imply that there are at least $n(n+1)+1$ points in the inversive plane, which there are not. So the nucleus Q lies in $\mathcal{A}_P$ and all the circles that join a single point of c to P contain the point Q.

16. We have to check axiom (GQ2).

A point P a type (a) and a line l of type (i) span a plane π' which meets π in a line say m.

If m is not as a tangent to $\mathcal{O}$ there is another line of type (i) l' such that $\langle l, l'\rangle$ meets π in m and $a \in l'$. The lines l and l' are contained in the plane π' so have a common point.

If m is a tangent then π' is a point of type (b) and the line of type (i) joining P to the point $m \cap l$ has the required property. One uses similar arguments to check the other point and line types.

17. For any two points of $\mathbf{u}, \mathbf{v} \in \mathcal{O}$ the bilinear form

$$\beta(\mathbf{u},\mathbf{v}) = u_0v_1 + v_0u_1 + u_2v_3 + u_3v_2 + 2u_4v_4 \neq 0.$$

However

$$\beta((-y^2-cx,1,c,x,y),(-y^2-dx,1,d,x,y)) = -2y^2-cx-dx+cx+dx+2y^2 = 0.$$

Now $\mathcal{O}$ contains $\langle 1,0,0,0,0\rangle$ so $u_1 \neq 0$ for any point in the ovoid. An ovoid of $Q(4,q)$ has q^2+1 points and so there is a unique point of $\mathcal{O}$ of the form $(-y^2-cx,1,c,x,y)$ for each pair (x,y).

Let $g(x,y) = -mx^\sigma$ and check that

$$\beta((-y^2-mx^{\sigma+1},1,-m^\sigma,x,y),(-t^2-ms^{\sigma+1},1,-m^\sigma,s,t))$$

$$= -y^2+mx^{\sigma+1}-t^2+ms^{\sigma+1}-smx^\sigma-mxs^\sigma+2yt = m(x-s)^{\sigma+1}-(y-t)^2.$$

Now assuming that $(x,y) \neq (s,t)$ β will be always non-zero iff m is a non-square.

Bibliography

[1] S. Ball, On large subsets of a finite vector space in which every subset of basis size is a basis, *Journal of the European Mathematical Society (JEMS)*, to appear.

[2] A. Blokhuis, A. A. Bruen and J. A. Thas, Arcs in $PG(n,q)$, MDS-codes and three fundamental problems of B. Segre - some extensions, *Geom. Dedicata* **35** 1–11 (1990).

[3] P. J. Cameron, *Combinatorics: topics, techniques, algorithms*, Cambridge University Press, 1994.

[4] P. J. Cameron, *Projective and polar spaces* 2nd edition, QMW Lecture Note Series, available from `www.maths.qmw.ac.uk/∼pjc/pps.`

[5] C. Chevalley, *The algebraic theory of spinor and Clifford algebras*, Springer-Verlag, Berlin, 1997.

[6] P. Dembowski, Inversive planes of even order, *Bull. Amer. Math. Soc.*, **69** 850–854 (1963).

[7] P. Dembowski, *Finite geometries*, Springer-Verlag, New York, 1968.

[8] J. W. P. Hirschfeld, *Finite projective spaces of three dimensions*, Oxford University Press, New York, 1985.

[9] J. W. P. Hirschfeld, *Projective geometries over finite fields*, Second Edition, Oxford University Press, New York, 1998.

[10] J. W. P. Hirschfeld and J. A. Thas, *General Galois geometries*, Oxford university Press, New York, 1991.

[11] D. R. Hughes and F. C. Piper, *Projective planes*, Springer-Verlag, New York-Berlin, 1973.

[12] C. W. H. Lam, L. H. Thiel and S. Swiercz, The non-existence of finite projective planes of order 10, *Canadian J. Math.* **41** 1117–1123 (1989).

[13] R. Lidl and H. Niederreiter, *Finite fields* 2nd Edition, Cambridge, 1997.

[14] S. E. Payne, J. A. Thas, *Finite generalised quadrangles*, Research Notes in Mathematics, 110. Pitman, Boston, 1984.

[15] B. Segre, Ovals in a finite projective plane, *Canad. J. Math.*, **7** 414–416 (1955).

[16] B. Segre, Curve normali e k-archi negli spazi finiti, *Ann. Mat. Pura Appl.* **39** 357-379 (1955).

[17] D. E. Taylor, *The geometry of the classical groups*, Heldermann Verlag, Berlin, 1992.