Finite Geometries
MA5980

Jürgen Bierbrauer

April 29, 2004

Contents

Chapter 1

Finite Fields

Finite fields and the principles of linear algebra will be fundamental for everything we do in this class. Recall the definition of a field:

1.1 Definition. *A* **field** *F is a set with two binary operations $+$ (addition) and $\cdot$ (multiplication) such that the following hold:*

- *$(F, +)$ is an abelian group (the* **additive group** *of F). Denote the neutral element by 0.*
- *$(F \setminus \{0\}, \cdot) = F^*$ is an abelian group (the* **multiplicative group** *of F),*
- *$0 \cdot a = a \cdot 0 = 0$ for all $a \in F$,*
- *$a(b + c) = ab + ac$ for all $a, b, c \in F$ (***distributive law***).*

Observe a condition, which is hidden in the second axiom: the product of two nonzero elements is nonzero: a field has no divisors of 0.

Let p be a prime number. Then $\mathbb{Z}/p\mathbb{Z}$ is a finite field. We denote it by $\mathbb{F}_p$. If n is a composite number (not a prime), then $\mathbb{Z}/n\mathbb{Z}$ cannot be a field as it has divisors of 0. Why is $\mathbb{Z}/p\mathbb{Z}$ a field? As p is prime there are no divisors of 0 (the product of two integers, both of which are not divisible by p is again not divisible by p). The only field axiom which could possibly be in doubt is the existence of multiplicative inverses. So let an integer a be given, which is not divisible by p. As p is prime it follows $gcd(a, p) = 1$. The main theoretical consequence of the Euclidean algorithm is that $gcd(n, m)$ can always be written as a linear combination of n and m. In our situation we obtain integers b, c such that $ab + cp = 1$. This shows that b represents the inverse of a in $\mathbb{F}_p$. A different proof is sketched in the first problem. For example, as $3 \cdot 4 = 12$ we have $4 = 1/3$ in $\mathbb{F}_{11}$. The inverse of 4 in $\mathbb{F}_{13}$ is 10 as $4 \cdot 10 = 40 \equiv 1(\bmod 13)$.

Let F be any finite field. Denote the sum of n copies of 1 by $n \cdot 1 \in F$. As F is finite the $n \cdot 1 \in F$ cannot all be different. So there must be some $m < n$ such that $n \cdot 1 = m \cdot 1$. It follows $(n - m) \cdot 1 = 0$. Denote by a the smallest natural number such that $a \cdot 1 = 0$. As F has no zero divisors we conclude that $a = p$ must be prime. It follows from the minimality that p is the only prime with this property and that $n \cdot 1 = 0$ if and only if n is a multiple of p. We see that the $i \cdot 1, i = 0, 1, \ldots p - 1$ form a subfield of F, which is isomorphic to $\mathbb{F}_p$. We call p the **characteristic** of F and $\mathbb{F}_p$, the subfield of F generated by 1, its **prime field.** So every finite field F may be described as an extension of its prime field $\mathbb{F}_p$. As F is by definition a vector space over $\mathbb{F}_p$ its number of elements is p^n for some n.

1.2 Theorem. *Every finite field has p^n elements for some prime p. The subfield generated by the element 1 is $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.*

In order to generate finite fields we use irreducible polynomials. So let $f(X) \in \mathbb{F}_p[X]$ be an irreducible polynomial of degree n. Take $f(X)$ to be monic (its leading coefficient is 1), so that $f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$. We claim that $F = \mathbb{F}_p[X]/(f(X))$, the factor ring of the polynomial ring over the ideal generated by $f(X)$, is a field with p^n elements:
denote by x the image of X mod the ideal $(f(X))$. Remark that $(f(X))$ simply is the set of all polynomials, which are divisible by $f(X)$. At first we see that F is a vector space of dimension n over $\mathbb{F}_p$, so $|F| = p^n$. The elements of F can be uniquely represented in the form $u = \sum_{i=0}^{n-1} c_i x^i$. In fact, as $x^n = -a_{n-1}x^{n-1} - \cdots - a_1 x - a_0$, every element of F has this form. On the other hand, the $x^i, i = 0, 1, \ldots, n - 1$ are linearly independent as otherwise $f(X)$ would divide a nonzero polynomial of degree $< n$, which is impossible. So every element of F can be written in a unique way as a polynomial of degree $< n$ with coefficients in $\mathbb{F}_p$. Assume $g(x), h(x)$ are such polynomials and $g(x)h(x) = 0$. This means that $f(X)$ divides $g(X)h(X)$. As $f(X)$ is irreducible it must divide either $g(X)$ or $h(X)$. Thus either $g(x) = 0$ or $h(x) = 0$. We have shown that F has no zero divisors. It remains to show that every nonzero element of F has a multiplicative inverse. So let $g(X)$ a nonzero polynomial of degree $< n$. As $f(X)$ is irreducible, it must be coprime to $g(X)$. We use the Euclidean algorithm again. We apply it here to

the polynomial ring $\mathbb{F}_q[X]$. We have $(g(X), f(X)) = 1$ as $f(X)$ is irreducible. We can therefore find polynomials such that $1 = g(X)h(X)+f(X)l(X)$. If we read this mod $(f(X))$ we get $1 = g(x)h(x)$ and have found the multiplicative inverse. In fact, these arguments are valid for any ground field. We have shown the following:

1.3 Theorem. *Let K be a field and $f(X)$ an irreducible monic polynomial of degree n over K. Then $F = K[X]/(f(X))$ is a field. It contains K as a subfield and is a vector space of dimension n over K. If x denotes the image of X mod $f(X)$, then the $x^i, i = 0, 1, \ldots, n-1$ form a basis of F as a K-vector space. We call n the* **degree** *of F over K.*

In order to illustrate this mechanism we construct the field $\mathbb{F}_4$ of 4 elements.

1.4 Example. *The only irreducible polynomial of degree 2 over $\mathbb{F}_2$ is $f(X) = X^2 + X + 1$. Let us check that $f(X)$ is indeed irreducible: if it was reducible it would have a root. As $f(0) = f(1) = 1 \neq 0$ we conclude that $f(X)$ is irreducible. We have $\mathbb{F}_2[X]/(f(X)) = \{0, 1, x, x + 1\}$, where x is the image of X mod $(f(X))$. We have $1 = f(X)+X(X+1)$. Reading this mod $(f(X))$ we obtain $1 = x(x+1)$. It follows that x and $x+1$ are multiplicative inverses. We conclude that $\mathbb{F}_2[X]/(f(X)) = \{0, 1, x, x+1\} = \mathbb{F}_4$ is a field with 4 elements.*

This method of generating extension fields is not limited to finite fields. In the second problem you are asked to use this method to construct the complex number field as a quadratic extension of the reals.

We accept from field theory the fact that an **algebraic closure** always exists and is uniquely determined. Denote by $\overline{\mathbb{F}}_p$ a fixed algebraic closure of $\mathbb{F}_p$. Recall that this means two things: firstly every element $a \in \overline{\mathbb{F}}_p$ is **algebraic** over $\mathbb{F}_p$, that is it satisfies a polynomial equation with coefficients in $\mathbb{F}_p$. Secondly, $\overline{\mathbb{F}}_p$ is **algebraically closed,** equivalently every polynomial with coefficients in $\overline{\mathbb{F}}_p$ splits into linear factors over that same field. Consider the polynomial $X^{p^n} - X$. Assume a field with p^n elements exists. As it is finite it must be algebraic over $\mathbb{F}_p$, so it can be considered as a subfield of $\overline{\mathbb{F}}_p$. As the multiplicative group of this field has order $p^n - 1$, each nonzero element u satisfies $u^{p^{n-1}} = 1$. Thus every element of a field of p^n elements is a root of our polynomial. We see that a field of order p^n is uniquely determined as a subfield of the algebraic closure, if it exists. On the other hand, the polynomial $X^{p^n} - X$ has p^n different roots. It suffices to check that these do form a field. As we are working inside a field it is sufficient to prove that sums, products and multiplicative inverses of roots are roots. For products and inverses this is obvious. For sums this is a consequence of the following lemma:

1.5 Lemma (Frobenius automorphism). *Let F be a field of characteristic p. Then the mapping σ, where $\sigma(x) = x^p$, is a field automorphism from F onto the field F^p of p-th powers. In the case of a finite field we have $F^p = F$. The fixed field of σ is $\mathbb{F}_p$.*

Proof. It suffices to prove that σ is linear with respect to addition and multiplication. In the case of multiplication this is obvious. Consider addition: obviously the binomial theorem applies, hence $(x+y)^p = \sum_{i=0}^{p} \binom{p}{i} x^i y^{p-i}$. Here the binomial coefficients are field elements, hence in $\mathbb{F}_p$. We see that $\binom{p}{i}$ is divisible by p and hence $= 0$ unless $i = 0$ or $i = p$. Hence the sum simplifies, giving the desired result: $(x+y)^p = x^p + y^p$. □

We conclude that our field of p^n elements exists and is uniquely determined.

1.6 Theorem. *For every prime p and natural number n there is a field with p^n elements. Moreover a fixed algebraic closure $\overline{\mathbb{F}_p}$ contains precisely one subfield with p^n elements, consisting of the roots of the polynomial $X^{p^n} - X$. We denote this field by $\mathbb{F}_{p^n}$.*

Once a field $\mathbb{F}_{q^n}$ is constructed we can go through the same process and construct fields of q^{nk} elements as extensions of $\mathbb{F}_{q^n}$, for every k. As we saw that these fields are uniquely determined we conclude that $\mathbb{F}_{q^n} \subset \mathbb{F}_{q^m}$ provided n divides m. On the other hand, assume $\mathbb{F}_{q^n} \subset \mathbb{F}_{q^m}$. Then the big field is a vector space over the small field. It follows that q^m must be a power of q^n, so n divides m. We have seen the following:

1.7 Theorem. *We have $\mathbb{F}_{q^n} \subseteq \mathbb{F}_{q^m}$ if and only if n divides m.*

1.8 Theorem. *The multiplicative group of a finite field is cyclic.*

Proof. Let $F = \mathbb{F}_{p^n}$. The multiplicative group of F is abelian of order $p^n - 1$. Assume it is not cyclic. Then for some prime l there must be a subgroup $Z_l \times Z_l$, an elementary abelian subgroup of order l. This is impossible as it would yield l^2 roots of the polynomial $X^l = 1$. □

1.9 Corollary. *Let q be an odd prime-power. Then -1 is a square in $\mathbb{F}_q$ if and only if $q \equiv 1 (mod\ 4)$.*

Proof. We have just seen that the multiplicative group of $\mathbb{F}_q$ is cyclic. The element -1 is the unique involution (= element of order 2) in this group. It is a square if and only if elements of order 4 exist. □

To sum up: we have found, for every $q = p^m$, a uniquely determined field $\mathbb{F}_q$ of q elements. Consider its extension field $\mathbb{F}_{q^n}$.

1.10 Lemma. *Consider the field extension $\mathbb{F}_{q^n} \supset \mathbb{F}_q$ for some prime-power q. The mapping σ, where $\sigma(x) = x^q$, is a field automorphism of $\mathbb{F}_{q^n}$* **over** *$\mathbb{F}_q$, this last term meaning that each element of the ground field is fixed under σ. More precisely we have $\sigma(x) = x$ if and only if $x \in \mathbb{F}_q$. The powers of σ form a group of automorphisms of order n. We call this group the Galois group $G(\mathbb{F}_{q^n}|\mathbb{F}_q)$.*

Proof. Our σ is a power of the Frobenius automorphism introduced in Lemma 1.5, so it certainly is a field automorphism of $\mathbb{F}_{q^n}$. As the elements of $\mathbb{F}_q$ satisfy $x^q = x$ we see that each element of $\mathbb{F}_q$ is fixed by σ. For the same reason we see that σ^n acts as the identity mapping on $\mathbb{F}_{q^n}$, and this is not the case for any smaller power of σ. As the polynomial $X^q - X$ of degree q cannot have more than q roots we conclude that the fixed points of σ are precisely the elements of $\mathbb{F}_q$. □

1.11 Definition (trace). *Let σ be the generator of the Galois group of $\mathbb{F}_{q^n}|\mathbb{F}_q$ as introduced in Lemma 1.10. Then the* **trace** *$tr : \mathbb{F}_{q^n} \longrightarrow \mathbb{F}_q$ is defined by*

$$tr(x) = \sum_{i=0}^{n-1} \sigma^i(x)$$

So $tr(x)$ is defined as the sum of the images of x under the elements of the Galois group. It is obvious that tr is an $\mathbb{F}_q$-linear mapping. As another application of σ permutes the elements of the Galois group we see that $\sigma(tr(x)) = tr(x)$. It follows $tr(x) \in \mathbb{F}_q$, as stated in Lemma 1.10. Moreover tr is not identically 0, as otherwise a polynomial of degree q^{n-1} would have q^n roots. We conclude that the kernel of tr is a hyperplane (an $(n-1)$-dimensional subspace) of $\mathbb{F}_{q^n}$, seen as a vector space over $\mathbb{F}_q$). For every $x \in \mathbb{F}_{q^n}$ the mapping $y \longrightarrow tr(xy)$ is a linear functional. As the space of linear functionals of an n-dimensional vector space clearly has dimension n we can describe every linear functional in this way:

1.12 Proposition. *Let F be an n-dimensional vector space over $\mathbb{F}_q$. Impose the structure of $\mathbb{F}_{q^n}$ on F. Then the linear functionals of F are in bijection with the elements of F, each $x \in F$ yielding the linear functional $y \longrightarrow tr(xy)$.*

We also see that **dual bases** always exist. In fact, let $v_1, v_2, \ldots, v_n$ be a basis of $\mathbb{F}_q^n$. Consider the linear functionals ϕ_i, where $\phi_i(\sum_j \alpha_j v_j) = \alpha_i$. Choose $x_i \in \mathbb{F}_{q^n}$ such that $tr(x_i y) = \phi_i(y)$. Then $tr(x_i v_j) = \delta_{ij}$. We have seen the following:

1.13 Theorem. *Let F be an n-dimensional vector space over $\mathbb{F}_q$. Impose the structure of $\mathbb{F}_{q^n}$ on F. For every basis $v_1, v_2, \ldots, v_n$ of $F|\mathbb{F}_q$ there exists another basis $x_1, x_2, \ldots, x_n$ (the* **trace-dual basis***) such that*

$$tr(x_i v_j) = \delta_{ij}.$$

Let us have a look at the smallest fields of non-prime order. It is in general handy to fix the multiplicative structure of the field (simply a cyclic group, as we know) and to determine the additive structure afterwards, using the irreducible polynomial. Write $\mathbb{F}_q = \{0\} \cup \{\epsilon^i \mid i = 0, 1, \ldots, q-1\}$. We know that the field is independent of the irreducible polynomial $f(X)$ chosen to describe it. However, there are good and bad choices. It is for example obviously advantageous to choose a polynomial of maximal exponent, meaning that the image of X is a generator of the multiplicative group. We will always do this. The addition in $\mathbb{F}_q$ will be completely known once the $1+\epsilon^i$ are known for all i.

$\mathbb{F}_4$: The only irreducible $\mathbb{F}_2$-polynomial of degree 2 is $f(X) = X^2+X+1$. It follows $1+\epsilon = \epsilon^2$. This determines the addition. For example $1+\epsilon^2 = \epsilon, \epsilon+\epsilon^2 = \epsilon(1+\epsilon) = \epsilon\epsilon^2 = 1$.

$\mathbb{F}_8$: We choose $f(X) = X^3 + X^2 + 1$, hence $1 + \epsilon^2 + \epsilon^3 = 0$. Further $1 + \epsilon^4 = (1+\epsilon^2)^2 = (\epsilon^3)^2 = \epsilon^6$, and then necessarily $1+\epsilon = \epsilon^5$. We know that these relations:

$$1+\epsilon^2+\epsilon^3 = 0,\ 1+\epsilon^4 = \epsilon^6,\ \text{and } 1+\epsilon = \epsilon^5$$

determine the field structure completely. As an example, $\epsilon^3 + \epsilon^4 = \epsilon^3(1+\epsilon) = \epsilon^3\epsilon^5 = \epsilon$.

$\mathbb{F}_9$: Take $X^2 - X - 1$ as irreducible polynomial. This leads to the relation $\epsilon^2 = \epsilon + 1$. Then $\epsilon^3 = -\epsilon + 1, \epsilon^4 = -1$, as it should be.

1. For every prime-power q there is a finite field $\mathbb{F}_q$ with q elements.
2. If p is a prime, then $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.
3. If $q = p^f$, then $\mathbb{F}_q$ can be constructed as an extension of $\mathbb{F}_p$, with the help of an irreducible polynomial.
4. The fields $\mathbb{F}_4, \mathbb{F}_8$ and $\mathbb{F}_9$ have been given in detail.
5. Among the most important notions concerning finite fields are the **trace** and the **Frobenius automorphism.**

Problems

1. Let $0 \neq a \in \mathbb{F}_p$. Show that a has an inverse in $\mathbb{F}_p$ by proving that the function f_a defined by $f_a(x) = ax$ is an injective and therefore bijective mapping : $F^* \to F^*$. This proof uses the finiteness of F.
2. Construct the field of complex numbers as a quadratic extension of the field $\mathbb{R}$ of real numbers. What is the natural choice of a quadratic irreducible polynomial?
3. Prove that the polynomial $X^3 + X^2 + 1$ is irreducible over $\mathbb{F}_2$.
4. Determine the product of all nonzero elements of $\mathbb{F}_q$.
5. Determine the sum of all elements of $\mathbb{F}_q$.
6. Prove the following: if $a, b \in \mathbb{F}_9$ are nonsquares and $a \neq \pm b$, then $a + b$ and $a - b$ are squares.
   Prove that every element of $\mathbb{F}_9$ is an $\mathbb{F}_3$-linear combination of two nonsquares.

Chapter 2

Projective geometries

We use basic linear algebra to construct projective geometries.

Let V be an $(n+1)$-dimensional vector space over the field K. We simply think of this object in geometrical terms. Call the 1-dimensional subspaces of V **points,** the 2-dimensional subspaces **lines,** the 3-dimensional subspaces **planes** and so on. Finally the n-dimensional subspaces of V are **hyperplanes.** This is justified by observations like the following:

Any two points are on precisely one common line.

Indeed, two different 1-dimensional subspaces (points) generate precisely one 2-dimensional space (a line). We begin with the smallest geometrically interesting case:

2.1 Definition. *The* 2*-dimensional projective geometry* $PG(2, K)$ *over the field* K, *also called* **classical projective plane** *over* K, *is based on a fixed* 3*-dimensional vector space* V.

The points of $PG(2, K)$ *are the* 1*-dimensional subspaces of* V, *the lines of* $PG(2, K)$ *are the* 2*-dimensional subspaces of* V. *A point* P *is on a line* l *(*P *and* l *are* **incident***) if the corresponding subspaces are contained in each other.*

2.2 Proposition. *The following hold for* $PG(2, K)$:

- *Any two points are on precisely one line, and dually*
- *any two lines have precisely one point in common.*

Proof. We have convinced ourselves of the validity of the first claim already. As for the second: two different 2-dimensional vector spaces generate the

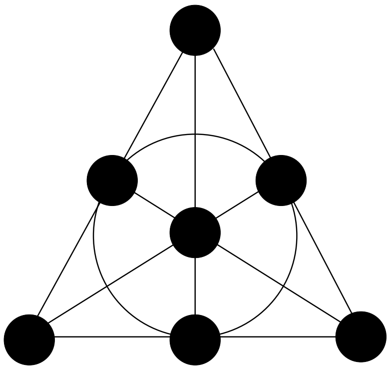

Figure 2.1: The Fano plane

ambient 3-dimensional vector space, and, by a familiar dimension formula of linear algebra, they intersect in a 1-dimensional vector space. □

Observe the shift of 1 in dimension from linear algebra to geometry. A point is a 0-dimensional geometric object but a 1-dimensional vector space, a line (seen as 1-dimensional in geometry) is a 2-dimensional vector space, and so forth.

We are mostly interested in finite geometries, so we choose $K = \mathbb{F}_q$ and write $PG(2, q)$ for $PG(2, \mathbb{F}_q)$. Here q is called the **order** of the projective plane. In the finite case it must be possible to count the basic objects. Indeed, this is easy to do: the number of points in $PG(2, q)$ is $(q^3 - 1)/(q - 1) = q^2 + q + 1$, the number of lines is the same. The number of points on a line is $(q^2 - 1)/(q - 1) = q + 1$, equal to the number of lines through a point.

2.3 Lemma. *$PG(2, q)$ has $q^2 + q + 1$ points and equally many lines. Any line has $q + 1$ points, any point is on $q + 1$ lines.*

As $\mathbb{F}_2$ is the smallest field, the classical projective plane $PG(2, 2)$ of order 2 is the smallest projective plane. We have seen the following: $PG(2, 2)$ has 7 points and 7 lines. Each line has 3 points, each point is on three lines, each pair of distinct points is on a unique line, each pair of distinct lines meets in a unique point. It should be possible to draw this little structure. In fact, the reader can convince himself or herself that the basic properties uniquely determine this structure. It is also known as the **Fano plane.**

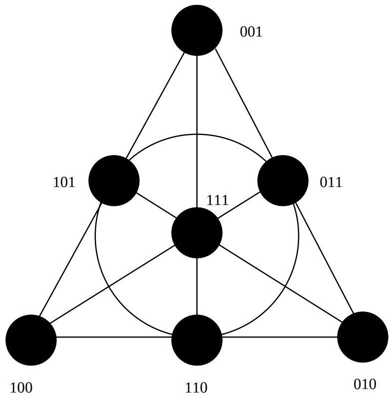

Figure 2.2: The Fano plane

If we need to calculate with the points of $PG(2,q)$ we use **homogeneous coordinates:** let $V = \mathbb{F}_q^3$ be the underlying vector space. Its elements are the triples (a,b,c) of field elements. Let $P = \mathbb{F}_q(a,b,c)$ be the 1-dimensional vector space generated by the nonzero triple (a,b,c). We write $P = (a:b:c)$. As any nonzero scalar multiple of (a,b,c) generates the same point P we have $(a:b:c) = (\lambda a : \lambda b : \lambda c)$ for every $0 \neq \lambda \in \mathbb{F}_q$. In the binary case ($K = \mathbb{F}_2$) there are no different nonzero scalar multiples. Here are the points of $PG(2,2)$, written in homogeneous coordinates:

$$P_1 = (0:0:1),\ P_2 = (0:1:0),\ P_3 = (1:0:0),$$

$$P_4 = (1:1:0),\ P_5 = (1:0:1),\ P_6 = (0:1:1),\ P_7 = (1:1:1).$$

In Figure 2.2 we labelled the points with their homogeneous coordinates (writing 001 for $(0:0:1)$ and so on). This labelling is far from being uniquely determined. There are as many labellings of our picture as there are symmetries of the Fano plane (see the Problems section).

Back to the projective geometry $PG(n,K)$ of arbitrary dimension n. It is based on an $(n+1)$-dimensional vector space V and therefore has n types of objects (from points=1-dimensional subspaces to hyperplanes=n-dimensional subspaces).

2.4 Definition. *The n-dimensional projective geometry $PG(n,K)$ over the field K, is based on a fixed $(n+1)$-dimensional vector space V.*

The r-dimensional objects described by $(r+1)$-dimensional subspaces of V are also known as r-flats.

As before write $PG(n,q)$ for $PG(n,\mathbb{F}_q)$. The number of points $PG(n,q)$ clearly is $(q^{n+1}-1)/(q-1) = q^n + q^{n-1} + \cdots + 1$.

As an example consider $PG(3,2)$. It has $2^4-1=15$ points and also 15 planes. The number of lines is $15 \cdot 14/6 = 35$. Each point is on 7 lines, each line has 3 points, each plane has 7 points.

The main reason why projective geometries are so important is the following: let $f(X_0,X_1,\dots,X_n)$ be a **homogeneous** polynomial of degree d with coefficients in the field K. Recall that a polynomial is called homogeneous of degree d if each of its monomials has degree $= d$. For example, $XY^3 + YZ_2^3 + ZX^3$ is homogeneous of degree 4, whereas X^2+Y^3 is not homogeneous. Let $f(x_0,\dots,x_n)=0$ for our homogeneous polynomial f (where $x_0,\dots,x_n \in K$). For each $\lambda \in K$ we have $f(\lambda x_0,\dots,\lambda x_n) = \lambda^d f(x_0,\dots,x_n)$. It follows that whenever $f(x_0,\dots,x_n)=0$ for some nonzero vector $(x_0,\dots,x_n) \in K^{n+1}$, then f vanishes on the whole 1-dimensional space generated by that vector. It is therefore natural to interpret the zeroes of f as points in $PG(n,K)$.

2.5 Definition. *Let $f = f(X_0,X_1,\dots,X_n)$ be a homogeneous polynomial of degree d with coefficients in the field K. The* **variety** *$V(f) \cap PG(n,K)$ consists of all points in $PG(n,K)$ on which f vanishes.*

Here is an interesting example: let $f(X,Y,Z) = XY^3 + YZ^3 + ZX^3$, where we see the coefficients as elements of $\mathbb{F}_2$. The points of $PG(2,2)$ on which f vanishes are $(1:0:0),\ (0:1:0),\ (0:0:1)$. This shows

$$V(f) \cap PG(2,2) = \{(1:0:0),\ (0:1:0),\ (0:0:1)\}.$$

As $\mathbb{F}_2 \subset \mathbb{F}_4$ we can also consider f as a polynomial with coefficients in $\mathbb{F}_4$. We have

$$V(f) \cap PG(2,4) = (V(f) \cap PG(2,2)) \cup \{(1:\omega:\omega^2),\ (1:\omega^2:\omega)\}.$$

Let us consider f as defined over $\mathbb{F}_8$. We use the description from Chapter 1:

$$\mathbb{F}_8 = \mathbb{F}_2(\epsilon),\ \epsilon^2+\epsilon^3 = \epsilon+\epsilon^5 = \epsilon^4+\epsilon^6 = 1.$$

Let $P = (x:y:z) \in V(f) \cap PG(2,8)$. The points with 0-coordinates are the three points from $PG(2,2)$, which we know already. From now on all coordinates are nonzero. We can choose $x = 1$. If $y = 1$ we obtain 3 solutions:

$$(1 : 1 : \epsilon^3),\ (1 : 1 : \epsilon^5),\ (1 : 1 : \epsilon^6).$$

The remaining points are easy to find once we observe the symmetry $\sigma : (x : y : z) \mapsto (x : \epsilon y : \epsilon^3 z)$ on $V(f) \cap PG(2, 8)$. In fact, σ maps $xy^3 + yz^3 + zx^3 \mapsto \epsilon^3(xy^3 + yz^3 + zx^3)$. In particular it maps points from $V(f) \cap PG(2, 8)$ to points from $V(f) \cap PG(2, 8)$. The reader will doubtless already have observed another symmetry of $V(f)$, the cyclic permutation $\rho : (x : y : z) \mapsto (y : z : x)$. We see that $|V(f) \cap PG(2, 8)| = 24$

A homogeneous polynomial of degree d with 3 variables describes what is known as an algebraic curve of degree d. The algebraic curve described by $XY^3 + YZ^3 + ZX^3$ is known as the **Klein quartic** (quartic because it has degree 4). The points in $V(f) \cap PG(2, K)$ are also known as the K-rational points of f. The mathematical discipline which studies the algebraic varieties defined by homogeneous polynomials is **algebraic geometry.** This is not the topic of our lecture. However it is important to note that homogeneous polynomials with coefficients in finite fields define highly structured interesting sets of points in projective geometries. We are going to study the case of degree 2 (quadrics) in detail. This is interesting in itself and it gives us rich material for everything to come.

1. The **projective plane** $PG(2,q)$ is based on a 3-dimensional vector space $V = V(3,q)$ over $\mathbb{F}_q$. Its points are the 1-dimensional subspaces, its lines are the 2-dimensional subspaces of V.

2. $PG(2,q)$ has q^2+q+1 points and equally many lines. Any two points are on precisely one line, any two lines intersect in one point. Each line has $q+1$ points, each point is on $q+1$ lines.

3. $PG(2,2)$ is the **Fano plane.**

4. The n-dimensional projective geometry $PG(n,q)$ is based on an $(n+1)$-dimensional vector space $V = V(n+1,q)$ over $\mathbb{F}_q$. Its elements are the subspaces of V (points, lines, ... hyperplanes).

5. $PG(n,q)$ has $(q^{n+1}-1)/(q-1)$ points and equally many hyperplanes.

6. **Homogeneous coordinates** are used for calculations with points and hyperplanes.

7. Let $f(X_0,\ldots,X_n)$ be a homogeneous polynomial in $n+1$ variables, with coefficients in $\mathbb{F}_q$. The roots of f can be seen as points in $PG(n,q)$. They form the variety $V(f) \cap PG(n,q)$.

Problems

1. Determine as many **symmetries** of the Fano plane as possible. Here a symmetry (or **automorphism**) is a permutation of the points (an element of the symmetric group S_7) having the property that the image of any line is a line again: the set of lines is unchanged under the permutation. Can we find out exactly how many symmetries there are? At least give a reasonable upper and lower bound.

2. Give a list of the points of $PG(2, 3)$ in homogeneous coordinates.

3. Prove that the number of hyperplanes of $PG(n, q)$ equals the number of points of $PG(n, q)$.

4. Set up a list of all 24 $\mathbb{F}_8$-rational points of the Klein quartic.

5. Let F be the group generated by the symmetries σ and ρ of the $\mathbb{F}_8$-rational points of the Klein quartic. Determine $|F|$.

Chapter 3

The link to codes

We want to understand in this chapter that linear error-correcting codes can equivalently be described by sets of points in projective spaces. This also provides an important application of our finite geometries. We start with a brief introduction to the scenario that led to the definition of error-correcting codes.

The objective of coding theory is the transmission of messages over noisy channels. Below is the standard picture visualizing the situation:

We assume that in the channel information gets destroyed according to certain probabilistic laws. This is a physical problem unlike the scenario in cryptography where we have to fight an opponent.

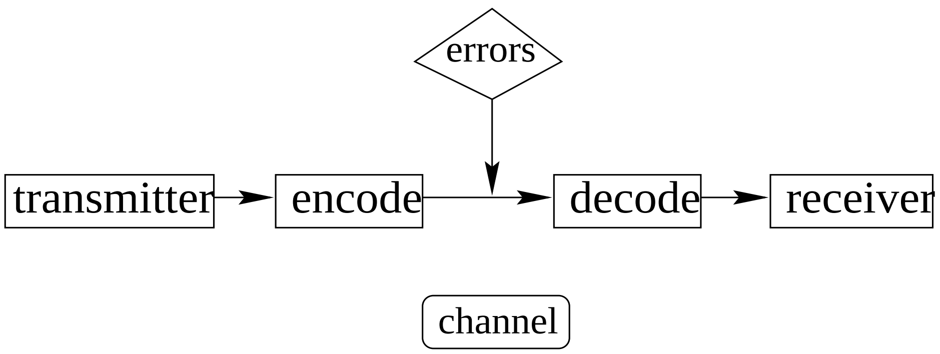

Figure 3.1: Information transmission over a noisy channel

3.1 Definition. *A* **linear code** *is a linear subspace of* $\mathbb{F}_q^n$.

The elements of a code are also known as code-words. Observe that the space $\mathbb{F}_q^n$ is an n-dimensional space over the field $\mathbb{F}_q$. The parameter n (the dimension of the ambient space) is the **length** of the code. A second basic parameter is the dimension k of the code. A third parameter allows the application we have in mind.

3.2 Definition. *The* **Hamming distance** *between two tuples* $x = (x_1, x_2, \ldots, x_n)$ *and* $y = (y_1, y_2, \ldots, y_n)$ *is defined as follows:* $d(x, y)$ *is the number of coordinates* i *where* $x_i \neq y_i$.

It is easy to see that d is in fact a metric. This is a natural notion in our context as $d(x, y)$ is simply the minimal number of errors in the coordinates that can transform the sent vector x into the received vector y.

The **minimum distance** of the linear code $\mathcal{C}$ is the largest number d such that any two different vectors from $\mathcal{C}$ have distance at least d. As we deal with linear codes this notion can be further simplified. The distance to the 0-tuple is known as the weight $wt(v)$. As each distance is also a weight ($d(x, y) = wt(x - y)$) the minimum distance equals the minimum weight among the nonzero code-words. We summarize:

3.3 Definition. *An* $[n, k, d]_q$*-code is a linear* k*-dimensional subspace of* $\mathbb{F}_q^n$ *such that each nonzero code-word has weight at least* d.

How can a code be used to allow reliable communication? Observe that any two different code-words are very different. For example, imagine we have $d = 7$ and we sent a code-word. Imagine further not more than three coordinates get corrupted during transmission. The received vector will resemble the sent code-word more (Hamming distance 3) than any code-word (Hamming distance ≥ 4). Of course, transmitter and receiver know the code. In particular the receiver knows that a code-word was sent. He will decode the received tuple as the closest code-word.

Imagine the plaintext as a sequence with entries in $\mathbb{F}_q$ (in this business it is often assumed that $q = 2$). On the transmitter side this stream is divided into blocks of length k. Fix a linear isomorphism $\alpha : \mathbb{F}_q^k \rightarrow \mathcal{C}$. The image of this encoding function is a code-word, in particular an n-tuple. This code-word is sent. If not too many errors occur (if the channel is not all that bad) the receiver can correct the errors. A final application of α^{-1} yields the original message.

It is clear that the code has to be adapted to the channel. There is a trade-off involved. k-tuples are sent as n-tuples, where $n > k$, in order to allow error-correction. The factor k/n (the **information rate**) represents the fraction of the coded message that actually represents information. The fraction representing additional costs is the **redundancy** $1 - k/n$. We have to introduce redundancy in order to allow error correction, but we want to keep it low to avoid additional cost.

The basic problem of coding theory is the construction of linear codes $[n, k, d]_q$, where d is maximized when the other parameters are given. A data base of the best known linear code parameters and upper bounds in maintained by A.E.Brouwer:

`http://www.win.tue.nl/~aeb/voorlincod.html`

It is an advantage of linear codes that k code-words suffice to describe a k-dimensional code (which has q^k elements). We write the elements of such a basis as rows of a matrix. This matrix describes the code.

3.4 Definition. *A* **generator matrix** G *of a k-dimensional q-ary linear code is a (k, n)-matrix whose rows form a basis of the code.*

This allows us to see the promised connection between sets of points in projective spaces and linear codes. As an example consider the $[7, 3]_2$-code with generator matrix

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Here comes the trick: we read this matrix columnwise instead of rowwise. Consider the 7 columns as generators of 1-dimensional subspaces of $\mathbb{F}_2^3$, hence as points in $PG(2, 2)$, the Fano plane. In our example each point of the Fano plane occurs precisely once as a column. What are the weights of codewords? Let v_1, v_2, v_3 the rows of G (a basis of the code). A generic code-word is $v = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$. It is described by the triple $(\lambda_1, \lambda_2, \lambda_3)$. What does it mean geometrically that v has entry 0 in coordinate i? Let the point P_i corresponding to coordinate i be $(a : b : c)$. The entry of v in coordinate i is then $\lambda_1 a + \lambda_2 b + \lambda_3 c$. This is 0 if point P_i satisfies the linear equation with coefficients $\lambda_1, \lambda_2, \lambda_3$, that is if it is in a certain hyperplane (line). So

$wt(v)$ equals the number of points among $P_1, \ldots, P_7$ which are outside the line described by the coefficients λ_j.

In our example we took all points of $PG(2,2)$ to describe columns. Outside each line there are 4 points of $PG(2,2)$. This shows that every nonzero code-word has weight 4. In particular we have a code $[7,3,4]_2$.

It is rather clear how this generalizes. In order to describe k-dimensional q-ary codes we work in $PG(k-1,q)$. The length of the code is the number of points chosen from $PG(k-1,q)$. Observe that there is no reason why a point should not occur more than once as a column of G. We should therefore not speak of a set of projective points but rather of a multiset. The formal way to describe this is by a weight function which assigns a non-negative integer weight $w(P)$ to each point $P \in PG(k-1,q)$. If $w(P) = 0$, then P does not occur among the columns, if $w(P) = 3$ say, then 3 of the columns of the generator matrix correspond to P.

3.5 Theorem. *The following are equivalent:*

- *A linear $[n,k,d]_q$-code such that in every coordinate there is a code-word with nonzero entry, and*
- *A function w assigning non-negative integer values $w(P)$ to the points of $PG(k-1,q)$ such that $\sum_{P \in PG(k-1,q)} w(P) = n$ and for every hyperplane H we have $\sum_{P \in H} w(P) \leq n-d$.*

Proof. As this is such an important theorem let us go through the formal proof. Describe the code by a generator matrix $G = (a_{ij})$, where $i = 1, \ldots, k; j = 1, \ldots, n$. Let $P_j = (a_{1j} : a_{2j} : \cdots : a_{kj}) \in PG(k-1,q)$ and $v_i = (a_{i1}, a_{i2}, \ldots, a_{in})$. The multiset $\{P_1, P_2, \ldots, P_n\}$ describes the function w : the value $w(P)$ is the multiplicity of P in this multiset. It is clear that this procedure is reversible: given a weight function w with $\sum_P w(P) = n$ we construct a corresponding (k,n)-matrix, which we use as generator matrix of a code. The code will have dimension k if and only if the points of the multiset are not contained in a hyperplane.

The basic point is the description of the weights of codewords. Let $x \neq 0$ be an arbitrary nonzero code-word. As the v_i form a basis there are uniquely determined scalars $\lambda_1, \ldots, \lambda_k$ (not all $= 0$) such that $v = \sum_{i=1}^k \lambda_i v_i$.
The entry in coordinate j is

$$v_j = \sum_{i=1}^{k} \lambda_i a_{ij}.$$

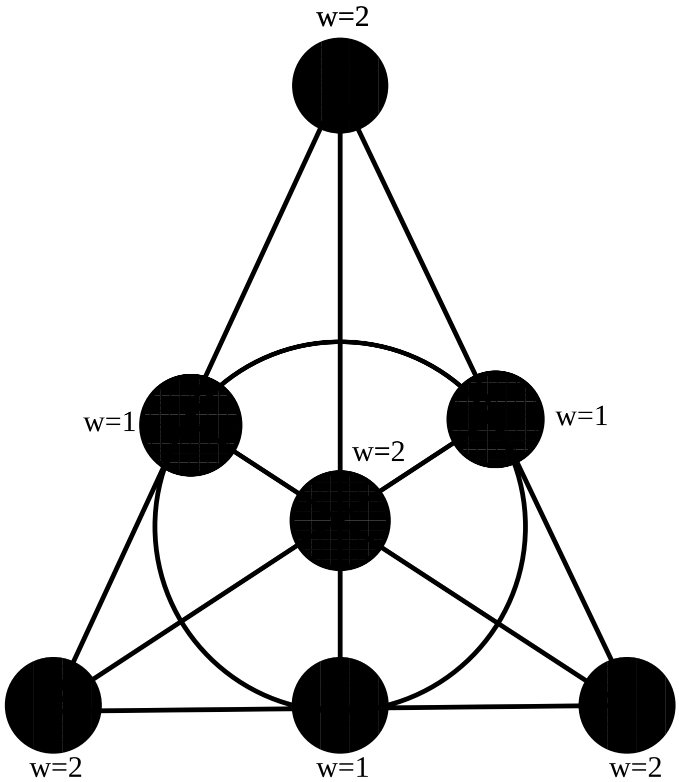

Figure 3.2: **A binary code [11,3,6]**

We have $v_j = 0$ if and only if P_j is in the hyperplane given by the λ_i (the elements of the vector space satisfying one nontrivial linear equation form a subspace of codimension 1). We conclude that v determines a hyperplane H and $wt(v) = n - |H \cap \{P_1, P_2, \dots, P_n\}|$, where the intersection has to be taken in the sense of multisets, formally $wt(v) = n - \sum_{P \in H} w(P)$. $\square$

Here is another illustration in the case of projective planes, where hyperplanes are lines. Use the points of the Fano plane as indicated in Figure 3.2. The sum of all weights is $n = 11$. We see that the sum of weights along lines does not exceed 5. This shows $d = 11 - 5 = 6$.

In general one speaks of **projective codes** if in a generator matrix no two columns are multiples of each other. In the terminology of Theorem 3.5 this means that $w(P) = 0$ or $= 1$ for each P. Our $[7, 3, 4]_2$ is projective, the $[11, 3, 6]_2$ is not.

A more solid example is obtained from the Klein quartic, see Chapter 2.

We saw that it has 24 rational points over $\mathbb{F}_8$. As the homogeneous polynomial has degree 4 it can be shown that no more than 4 of these points are on a line (are **collinear**). This shows that the $\mathbb{F}_8$-rational points of the Klein quartic determine a code $[24, 3, 20]_8$.

We know how to calculate with points of $PG(k-1, q)$, assigning homogeneous coordinates to them. The proof of Theorem 3.5 suggests a natural way to assign homogeneous coordinates to hyperplanes as well. In fact, each hyperplane can be described as the set of points satisfying a nontrivial linear equation. The coefficients of this equation are uniquely determined up to scalar multiples.

3.6 Definition (homogeneous coordinates). *The homogeneous coordinates $(x_1 : x_2 : \ldots, x_k)$ denote the point of $PG(k-1, q)$ determined by $(x_1, x_2, \ldots, x_k)$. The homogeneous coordinates $[y_1 : y_2 : \ldots, y_k]$ determine the hyperplane of $PG(k-1, q)$ concisting of all points $(x_1 : x_2 : \ldots, x_k)$ such that*

$$x_1y_1 + x_2y_2 + \cdots + x_ky_k = 0.$$

1. **Codes** are used to transmit messages over noisy channels.
2. Linear q-ary codes of length n are subspaces of $\mathbb{F}_q^n$.
3. A linear code $[n, k, d]_q$ is a k-dimensional subspace of $\mathbb{F}_q^n$ such that each nonzero word (element) has a nonzero entry in at least d of the n coordinates.
   d is the **minimum distance.**
4. A **generator matrix** of a linear q-ary code is a (k, n)-matrix with entries from $\mathbb{F}_q$ whose rows form a basis of the code (=subspace).
5. A code $[n, k, d]_q$ can equivalently be described as a multiset of n points in $PG(k-1, q)$, which has the property that there are at least d points outside any given hyperplane.

Problems

1. Assume you have 24 points in $PG(11, 2)$ such that each hyperplane contains at most 16 of these points. Determine the parameters of the corresponding code. What is the number of codewords?

2. Find the generator matrix of a code $[13, 3, 7]_2$.

3. Use representatives for all points of $PG(k-1, q)$ as columns of a generator matrix. Determine the parameters of this **Simplex code.**

4. Fix a line l in $PG(2, q)$. Use representatives for all points of $PG(2, q)$ outside l as columns of a generator matrix. Determine the parameters of the code.

5. The **Singleton bound** states that $k + d \leq n + 1$ for each linear code $[n, k, d]_q$. Prove this bound.

Chapter 4

An application: resilient functions

Let $A = (a_{ij})$ be a generator matrix of an $[n, k, d]_q$-code $\mathcal{C}$, just as in the preceding chapter. Let $P_1, P_2, \ldots, P_n$ be the points in $PG(k-1, q)$ described by the columns of A (eventually with multiplicities). We recall that the minimum distance d can be expressed in terms of the points P_i as follows: outside each hyperplane of $PG(k-1, q)$ there are at least d of the points P_i, again counted with multiplicities.

Consider the function $F : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^k$ defined by

$$F(x) = Ax$$

Here we see $x = (x_1, x_2, \ldots, x_n)^T$ and y as column vectors. In order to be perfectly clear: $F(x)$ is a linear combination of the columns of the generator matrix (the points P_i) with the x_i as coefficients. What are the properties of this mapping F? First of all it is a linear mapping. As $\mathcal{C}$ has dimension k (equivalently: A has rank k), F is onto. It follows that each $y \in \mathbb{F}_q^k$ has the same number of preimages x (this number is q^{n-k}). In the applications we view x as an input vector and y as an output vector.

Now fix the input in some t coordinates, say the coordinates $1, 2, \ldots t$. Let these input values be $a_1, \ldots, a_t$. We consider the function g obtained from F by making this substitution, to be precise $g : \mathbb{F}_q^{n-t} \longrightarrow \mathbb{F}_k$ is defined by

$$g(x_1, \ldots, x_{n-t}) = F(a_1, \ldots, a_t, x_1, \ldots, x_{n-t}).$$

We wish to guarantee that g is onto, no matter which set of t input coordinates are fixed and no matter what the fixed value $a_1, \ldots, a_t$ are. In the situation above, when will that be violated? If the points $P_{t+1}, \dots, P_t$ are on a hyperplane, in general if some $n-t$ of the points P_i are on a hyperplane, equivalently if some $n-t$ columns of A form a matrix of rank $< k$. Our geometric description of the minimum distance shows that this cannot happen if we choose $t = d-1$.

4.1 Definition. *A q-ary* **linear t-resilient function** *$RF_q(n,k,t)$ is a linear mapping $F : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^k$ with the property that whenever the value of some t input variables are fixed the resulting mapping $: \mathbb{F}_q^{n-t} \longrightarrow \mathbb{F}_q^k$ still is onto.*

Our discussion shows that linear resilient functions are linear codes in disguise:

4.2 Theorem. *The following are equivalent:*

- *A linear $RF_q(n,k,t)$,*
- *a linear code $[n,k,t+1]_q$.*

We also saw how to construct the resilient function from a generator matrix of the code.

Why is this interesting and what is resilient about a resilient function? Imagine a situation in cryptography where x is a cryptographic key. We suspect that a certain number of input values may have leaked to an opponent. In other words, the key x has been compromised and should not be used any more. The idea is to apply a function $F : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^k$ for some $k < n$ and use the shorter string $y = F(x)$ as key. Which conditions will F have to satisfy? Even if x itself is perfectly safe, there is one condition that F has to satisfy: each y has to have the same number of preimages under F : the function F must be **balanced.** If this was not the case, an opponent could exploit the fact that some values y are more probable than others. The need of balance also explains the main axiom of t-resiliency: even if up to t input values have leaked this does not give the opponent any information about the output value y, not even in terms of probabilities. Using linear functions is a cheap way of generating balanced functions: each surjective linear function is automatically balanced. The term **resilient** refers to an application in the construction of ciphers, which offer resistance to a certain type of attack, a correlation attack.

As an example consider the following generator matrix of a binary code:

$$A = \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{array} \right)$$

The code generated by A is known as the **extended binary Hamming code.** It is easy to see that its parameters are $[8,4,4]_2$. The corresponding resilient function is

$$F(x) = (x_1+x_6+x_7+x_8, x_2+x_5+x_7+x_8, x_3+x_5+x_6+x_8, x_4+x_5+x_6+x_7),$$

an $RF_2(8,4,3)$. In practice resilient functions are considered only in the binary case $q=2$. Resiliency is not the only design criterion. Satisfying other cryptographic criteria as well leads to subtle tradeoff questions. One important criterion is **non-linearity.** As the linearity can be used in attacks as well one wants to design functions which are in a sense far from being linear. This sounds at first like bad news for our linear construction. However, the usual way to deal with this problem is to base oneself on linear resilient functions, which are then twisted in some way to obtain non-linearity. Virtually all constructions involve linear codes or, equivalently, point sets in projective spaces.

1. Let $F : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^k$ be linear.
2. F is a linear resilient functions $RF_q(n,k,t)$ if all functions obtained by fixing t input variables are onto.
3. Linear $RF_q(n,k,t)$ are equivalent with $[n,k,t+1]_q$-codes.
4. They have multiple applications in cryptography and theoretical computer science.

Problems

1. Prove that the matrix A given above generates an $[8,4,4]_2$-code.
2. Give the geometric description for an $[8,4,4]_2$-code.
3. Construct an $RF_2(11,3,5)$.

Chapter 5

Arcs in projective planes

We use homogeneous coordinates to calculate in $PG(2,q)$. What we want to study are quadrics. Later on it will turn out that it does not make much difference which quadratic polynomial we use. For now let us consider $V(XZ-Y^2)$. Let $(x:y:z)$ be a point on this quadric. If $x=0$, then $y=0$. The corresponding point is $(0:0:1)$. The remaining points have $x\neq 0$. We can choose $x=1$. The equation is $z=y^2$. Clearly we have q such points: y is an arbitrary element of $\mathbb{F}_q$ and $z=y^2$. All in all we count $q+1$ points.

5.1 Proposition. *The quadric $V(XZ-Y^2)\cap PG(2,q)$ (also known as a* **conic***) has $q+1$ points. These are $P_\infty=(0:0:1)$ (also known as* **point at infinity***) and the points $P_\alpha=(1:\alpha:\alpha^2)$, where $\alpha\in\mathbb{F}_q$.*

We want to study the structure of this conic. It has as many points as a projective line, so at first one may suspect that all points are on a line. This is far from being true. Consider the lines joining pairs of points of the conic. The lines through P_∞ have the form $[a:b:0]$. This line contains P_α if $a+b\alpha=0$. It follows that $b\neq 0$. We choose $b=1$ and obtain $a=-\alpha$. This line $[-\alpha:1:0]$ contains P_β if and only if $\beta=\alpha$. This shows that the line from P_∞ to P_α contains only these two points of the conic.

Let now α,β be different field elements and $[a:b:c]$ the line joining them. As before it is clear that $c\neq 0$. Choose $c=1$. The equations for a,b are $a+b\alpha+\alpha^2=0,\ a+b\beta+\beta^2=0$. Subtract, divide by $\beta-\alpha$. This yields $b=-(\alpha+\beta), a=\alpha\beta$. The connecting line is $[\alpha\beta:-(\alpha+\beta):1]$. This is easy to check (the dot product of $(1,\alpha,\alpha^2)$ and $(\alpha\beta,-(\alpha+\beta),1)$ is $\alpha\beta-\alpha^2-\alpha\beta+\alpha^2=0$).

Which points P_γ are on this connecting line? The condition is $0 = \alpha\beta - \gamma(\alpha + \beta) + \gamma^2 = (\gamma - \alpha)(\gamma - \beta)$. This is satisfied only if $\gamma = \alpha$ or $\gamma = \beta$. We conclude that there are no further points of the conic on that line.

5.2 Definition. *A k-**arc** of $PG(2, q)$ is a set of k points such that no 3 are on a line (collinear). An **oval** is a $(q+1)$-arc, a **hyperoval** is a $(q+2)$-arc.*

Our calculation with coordinates has shown the following:

5.3 Theorem. *The conic $V(XZ - Y^2) \cap PG(2, q)$ is an oval.*

Let us do some combinatorial work. What is the maximum conceivable size of an arc $K \subset PG(2, q)$? Fix a point $P \in K$. Every point is on $q+1$ lines. Each such line contains at most one further point of K. It follows $|K| \leq q+2$. Conics give us ovals. This raises the question if hyperovals exist. Let K be a hyperoval. The counting argument based on $P \in K$ shows that each line which intersects K nontrivially must contain precisely 2 points of K. Let now X be a point, which is not in K. Each line through X meets K either in 0 or in 2 points. This shows that $|K| = q + 2$ must be even, in other words this can happen only in characteristic 2.

5.4 Theorem. *When q is odd, then the maximum number of points of an arc in $PG(2, q)$ is $q+1$. When q is a power of 2, then every oval is embedded in precisely one hyperoval. In particular the maximum size of an arc in $PG(2, q)$ is $q + 2$ when q is even.*

We have not proved yet the statement concerning the existence of hyperovals in $PG(2, q)$. In the problems section the reader is asked to complete the proof. We start with some easy observations. Let K be an oval and $P \in K$. With obvious terminology we call a line a **tangent** to K if it meets K in precisely one point. On how many tangents is P? There are q lines joining P to the remaining points of the oval. This shows that P is on precisely one tangent. We have seen that there are precisely $q + 1$ tangents to K, one through each point of K.

Let $N \notin K$. Which conditions will N have to meet such that $K \cup \{N\}$ is a hyperoval? Clearly each line NP for $P \in K$ must be a tangent to K. As there are $q+1$ points in K all lines through N must be tangents. This yields a very strong condition: in order for K to be embeddable in a hyperoval, all tangents to K must meet in a common point N. If this is satisfied N is the only point such that $K \cup \{N\}$ is a hyperoval. Point N is also known as the **nucleus** of K.

In order to check that our conic is an oval we could have proceeded in a different way. Observe that three points are collinear if and only if they satisfy a linear equation. This means that the $(3,3)$-matrix whose columns are representatives of the points has determinant 0. The statement that the conic $V(XZ-Y^2) \cap PG(2,q)$ is an oval is therefore equivalent to the following: whenever α, β, γ are different, the matrix which has $P_\alpha, P_\beta, P_\gamma$ as columns is non-singular (determinant $\neq 0$). If P_∞ is involved this is clear, so we can assume $\{\alpha, \beta, \gamma\} \subset \mathbb{F}_q$. The corresponding matrix is $\begin{pmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{pmatrix}$. Matrices of this type are known as **Vandermonde matrices.** It is easy to see that the determinant is $\neq 0$.

Vandermonde determinants are an essential ingredient in the study of **Reed-Solomon codes** and **cyclic codes.**

1. A k-arc in $PG(2,q)$ is a set of k points no 3 of which are collinear (on a line).
2. An **oval** is a $(q+1)$-arc, a **hyperoval** is a $(q+2)$-arc in $PG(2,q)$.
3. The set of roots of $XZ - Y^2$ (a **conic**) is an oval in $PG(2,q)$.
4. Hyperovals exist only in characteristic 2.

Problems

1. Determine homogeneous coordinates of the tangents to the conic $V(XZ - Y^2) \cap PG(2,q)$.
2. Let q be a power of 2. Find the nucleus of $V(XZ - Y^2) \cap PG(2,q)$.
3. Let q be a power of 2. Prove that each oval in $PG(2,q)$ can be embedded in a hyperoval. Hint: study at first the intersections of tangents and external lines.
4. Find a generator matrix of a code $[6,3,4]_4$.

Chapter 6

Symmetric bilinear forms

Let $V = V(n,q)$ be an n-dimensional vector space over $\mathbb{F}_q$. Consider mappings

$$(,) : V \times V \longrightarrow \mathbb{F}_q.$$

We speak of a **bilinear form** (or scalar product) if it satisfies the following conditions:

- Biadditivity: $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$ and $(x, y_1 + y_2) = (x, y_1) + (x, y_2)$ (everything in V, in particular $(0, y) = (x, 0) = 0$ always),
- $$(\lambda x, y) = (x, \lambda y) = \lambda \cdot (x, y) \text{ for all } x, y \in V, \lambda \in \mathbb{F}_q.$$

A bilinear form is **non-degenerate** if the only element x_0 satisfying $(x_0, y) = 0$ for all $y \in V$ is $x_0 = 0$ and analogously the only $y_0 \in V$ satisfying $(x, y_0) = 0$ for all $x \in V$ is $y_0 = 0$. A vector $v \in V$ is **isotropic** if $(v, v) = 0$. A subspace $U \subset V$ is **totally isotropic** if $(u, u') = 0$ for all $u, u' \in U$. Given the scalar product $(,)$ and a basis $\{v_1, v_2, \ldots, v_n\}$ of V we can form the **Gram matrix** $A = ((v_i, v_j))_{i,j}$. Clearly the bilinear form is non-degenerate if and only if the Gram matrix is invertible.

We can represent bilinear forms in matrix notation. Let $V = \mathbb{F}_q^n$ and $e_1, \ldots, e_n$ the standard basis. Choose constants a_{ij} such that $(e_i, e_j) = a_{ij}$. Let $x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n)$. Then

$$(x, y) = \sum_{i,j} x_i y_j a_{ij}.$$

This means that an n-dimensional bilinear form is described by choosing n^2 coefficients a_{ij}. The bilinear form can then also be written in matrix notation as follows: define the matrix $A = (a_{ij})$ whose entries are the chosen coefficients. Then

$$(x, y) = \sum_{i,j=0}^{n} a_{ij} x_i y_j = xAy^t.$$

Then A is the Gram matrix with respect to the standard basis.

As an example consider the following 3-dimensional bilinear form over $\mathbb{F}_q$ for odd q :

$$(x, y) = \frac{1}{2}x_1y_3 + \frac{1}{2}x_3y_1 - x_2y_2.$$

Its Gram matrix is $A = \begin{pmatrix} 0 & 0 & 1/2 \\ 0 & -1 & 0 \\ 1/2 & 0 & 0 \end{pmatrix}$ of determinant $1/4$. In particular it is non-degenerate.

As indicated earlier we are particularly interested in the scalar products of vectors with themselves. These are the values (v, v). In our example we have $(x, x) = x_1x_3 - x_2^2$. This leads to a homogeneous quadratic polynomial. It is precisely the polynomial considered in Chapter 5 (here only in odd characteristic). The Gram matrix A of the example has another special property: it is symmetric. This has the following effect on the bilinear form: $(x, y) = (y, x)$ for all x, y. Such bilinear forms are called symmetric.

6.1 Definition. *A bilinear form on V is* **symmetric** *if $(x, y) = (y, x)$ for all $x, y \in V$. This is equivalent to the Gram matrix being a symmetric matrix.*

6.2 Definition. *Let $(,)$ be a symmetric bilinear form on V. The corresponding* **quadratic form** *is $Q(x) = (x, x)$. Here $Q : V \to \mathbb{F}_q$ has the property $Q(\lambda x) = \lambda^2 Q(x)$ for all $x \in V$.*

It seems that the quadratic form Q carries less information than the symmetric bilinear form it was derived from. However, this is not so. We can recover the symmetric bilinear form from the quadratic form. Let us see why:

$$Q(x + y) = (x + y, x + y) = Q(x) + Q(y) + 2(x, y)$$

Here we used that the bilinear form is symmetric. This shows

$$(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y)) \text{ for all } x, y \in V.$$

However, this works only in odd characteristic (we divided by 2). Let us collect all this, starting from the abstract definition of a quadratic form in odd characteristic.

6.3 Definition. *Let q be odd and $V = V(n, q)$. A* **quadratic form** *on V is a mapping $Q : V \rightarrow \mathbb{F}_q$ such that $Q(\lambda x) = \lambda^2 Q(x)$ for all $\lambda \in \mathbb{F}_q, x \in V$ and such that*

$$(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y))$$

is a bilinear form (by force symmetric). Call Q non-degenerate if the corresponding bilinear form is.

Let us check this for our example. Starting from Q we obtain

$$2(x, y) = (x_1 + y_1)(x_3 + y_3) - (x_2 + y_2)^2 - x_1x_2 + x_2^2 - y_1y_3 + y_2^2 =$$

$$= x_1y_3 + x_3y_1 - 2x_2y_2$$

as predicted.

The quadratic form corresponding to a symmetric bilinear form is always described by a homogeneous quadratic polynomial. We have seen that in odd characteristic quadratic forms are equivalent with symmetric bilinear forms.

For the remainder of the chapter we concentrate on symmetric bilinear forms in odd characteristic. So let $V = V(n, q)$ for odd q and $(,)$ a non-degenerate symmetric bilinear form defined on V. The structure we have in mind is the quadric consisting of the isotropic points, that is the points in $PG(n - 1, q)$ generated by nonzero vectors x such that $(x, x) = 0$. One also says that vectors x, y are **orthogonal** provided $(x, y) = 0$. In this terminology a vector is isotropic if it is orthogonal to itself.

6.4 Definition. *Let $(,)$ be a symmetric bilinear form on V. For every subset $W \subseteq V$ define $W^\perp = \{v | v \in V, (v, W) = 0\}$. Here $(v, W) = 0$ stands short for $(v, w) = 0$ for all $w \in W$.*

6.5 Proposition. *Let $(,)$ be a symmetric bilinear form on V. For every subset $W \subseteq V$ we have that $W^\perp$ is a subspace. If $(,)$ is non-degenerate and W is a subspace, then $dim(W) + dim(W^\perp) = dim(V)$.*

Proof. The first statement is obvious. Consider now the non-degenerate case. By definition we have $v^\perp \neq V$ for all $v \neq 0$. Let $dim(V) = n$. We claim $dim(v^\perp) = n - 1$. It suffices to show that $v^\perp$ intersects every 2-dimensional subspace of V in dimension ≥ 1. Let v_1, v_2 be a basis of a 2-dimensional subspace, and $(v, v_1) = \lambda, (v, v_2) = \mu$. Then $(v, \mu v_1 - \lambda v_2) = 0$, hence $\mu v_1 - \lambda v_2 \in v^\perp$.

We have shown that the orthogonal of a 1-dimensional subspace has dimension $n-1$. Let now $v_1, v_2, \ldots, v_n$ be a basis of V. We have

$$\langle v_1 \rangle^\perp \supseteq \langle v_1, v_2 \rangle^\perp \supseteq \ldots \langle v_1, v_2, \ldots, v_n \rangle^\perp = V^\perp = 0.$$

As at each step the dimension decreases by at most one (intersection with a hyperplane), and in the last step we reach the 0-space, the dimension must decrease by precisely one at each step. In particular $dim(\langle v_1, \ldots, v_m \rangle^\perp) = n - m$ for each m. As this is true for an arbitrary basis the statement is proved. □

The 1-dimensional case

It may sound silly, but we start with case $n = 1$. Let $x \neq 0$. Then $(x, x) \neq 0$ because of non-degeneracy. Recall $(\lambda x, \lambda x) = \lambda^2 (x, x)$. This means that by changing the "basis" we can introduce an arbitrary quadratic factor in the "Gram matrix". This shows that there are two non-equivalent bilinear forms in the case of dimension 1. Either we can find a "basis" v such that $(v, v) = 1$ or we can find a basis such that $(v, v) = \nu$ is our favorite non-square in $\mathbb{F}_q$.

The 2-dimensional case

Let $V = V(2, q)$ with a non-degenerate form $(,)$. Assume there is a vector $v \neq 0$ such that $(v, v) = 0$. Let w be linearly independent from v. We have $(v, w) \neq 0$. Multiplying w by a suitable scalar we can assume $(v, w) = 1$. Upon replacing w by $w + sv$ for suitable s we can assume $(w, w) = 0$. This shows that we can find a basis such that the Gram matrix is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ in this case. These 2-dimensional spaces are also known as **hyperbolic planes.**

Assume there is no nonzero isotropic vector. Can this happen? Such spaces, where no nonzero vector is isotropic, are called **anisotropic.** Euclidean (real) spaces are examples of anisotropic spaces. Assume at first (v, v) is a square for all $v \neq 0$. Choose $w \neq 0$ such that $(v, w) = 0$. As $(sv + tw, sv + tw) = s^2(v, v) + t^2(w, w)$ must be a square unless $s = t = 0$

we see in particular that the sum of two squares must be a square. This leads to the contradiction that the squares together with 0 form a subfield of $\mathbb{F}_q$. Assume now (v, v) is a non-square for all $v \neq 0$. Replace the bilinear form $(,)$ by $\nu(,)$, where ν is a fixed non-square. This new form is still symmetric bilinear anisotropic, and each value $\nu(v, v)$ for $v \neq 0$ is a square, by assumption. This case has just been excluded.

We have seen that we can choose $(v, v) = 1$ and $(v, w) = 0$. As $(sv + tw, sv + tw) = s^2 + t^2(w, w) \neq 0$ unless $s = t = 0$ it must be that $-(w, w)$ is a non-square, without restriction $(w, w) = -\nu$, where ν is our favorite non-square. The Gram matrix is $\begin{pmatrix} 1 & 0 \\ 0 & -\nu \end{pmatrix}$ As $Q(sv + tw) = s^2 - t^2\nu$ we see that indeed this describes an anisotropic space. Observe also that every field element is represented. In fact, we saw that it is the uniquely determined anisotropic 2-dimensional space. Multiplying the Gram matrix by a constant c yields another anisotropic space, where c is represented as often as 1 is represented in the original space. As these spaces are equivalent we see that every nonzero element is represented the same number of times.

It is also interesting to consider the degenerate 2-dimensional case. If the radical is 2-dimensional, the bilinear form is identically 0. In particular all points of the corresponding projective line are isotropic.

The remaining case is when $Rad(,) = \langle v \rangle$ is 1-dimensional. Then V is the orthogonal sum of the radical and a 1-dimensional non-degenerate space. In particular the line contains precisely one isotropic point in this case.

6.6 Theorem. *Let $(,)$ be a symmetric bilinear form on $V = V(2, q)$, for odd q. The number of isotropic points on the projective line $PG(1, q)$ corresponding to V is $0, 1, 2$ or $q + 1$. It is $q + 1$ if the form is identically zero, it is 1 if the radical has dimension 1.*

There are up to equivalence only two non-degenerate bilinear forms, the hyperbolic plane with Gram matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and the anisotropic space with Gram matrix $\begin{pmatrix} 1 & 0 \\ 0 & -\nu \end{pmatrix}$

If the Gram matrix is not given in one of the standard forms of Theorem 6.6, how can we decide on the fly what the type is? For that purpose it is good to know what happens to the Gram matrix under change of basis.

6.7 Theorem. *Let $v_1, \dots, v_n$ be a basis of V and $A = (a_{ij})$ the Gram matrix of a bilinear form with respect to the v_i ($a_{ij} = (v_i, v_j)$). Let $w_1, \dots, w_n$ be another basis of V with Gram matrix $B = (b_{ij})$. Let $w_i = \sum_{k=1}^{n} t_{ik} v_k$. Let $T = (t_{ij})$ be the matrix describing this change of basis. Then $B = TAT^t$.*

Proof. Simply compute

$$b_{ij} = (w_i, w_j) = (\sum_k t_{ik} v_k, \sum_l t_{jl} v_j) = \sum_{k,l} t_{ik} a_{kl} t_{jl}.$$

This is the (i, j)-entry of TAT^t. □

The most important point about the base change is what remains invariant. Consider the determinants: $det(B) = det(A)det(T)^2$. The determinant certainly changes but squares are mapped to squares and non-squares are mapped to non-squares.

6.8 Definition. *Let A be an (n, n)-matrix with entries from $\mathbb{F}_q$, where q is odd. The* **discriminant** *of A is $disc(A) = +1$ if $det(A)$ is a square, it is $disc(A) = -1$ if $det(A)$ is non-square, it is 0 if $det(A) = 0$. The discriminant of a bilinear form is the discriminant of its Gram matrix.*

In fact, Theorem 6.7 shows that the discriminant of the Gram matrix is invariant under change of basis. It is therefore an invariant of the bilinear form. This shows us how to distinguish hyperbolic planes from anisotropic 2-dimensional spaces. As the determinants of the Gram matrices in standard form are -1 and $-\nu$, respectively, they have different discriminant.

Our analysis of the anisotropic 2-dimensional case suggests that anisotropic spaces of higher dimensions cannot exist.

6.9 Proposition. *There is no anisotropic symmetric bilinear form in odd characteristic on a vector space of dimension $n > 2$.*

Proof. The proof in Section 6 shows that we can find v_1 such that $(v_1, v_1) = 1$. The space V is the orthogonal sum of $\langle v_1 \rangle$ and $\langle v_1 \rangle^\perp$, and the latter space still has dimension ≥ 2. We can find v_2 such that $(v_1, v_2) = 0$ and $(v_2, v_2) = -1$. However $Q(v_1 - v_2) = 0$, contradiction. □

We can now complete the classification in arbitrary dimension. Let $n \geq 3$. By Proposition 6.9 we can find $v_1 \neq 0$ such that $(v_1, v_1) = 0$. Because of non-degeneracy we can find $w_1' \neq 0$ such that $(v_1, w_1') \neq 0$. After multiplying w_1'

by a suitable constant we have $(v_1, w_1') = 1$. Let $w_1 = w_1' + sv_1$. We have $(v_1, w_1) = 1$ independent of the value of s. As $Q(w_1) = Q(w_1') + 2s$ we can choose s such that $Q(w_1) = 0$. It follows that $H_1 = \langle v_1, w_1 \rangle$ is a hyperbolic plane. As H_1 is itself non-degenerate we obtain $V = H_1 \perp H_1^\perp$, and the $(n-2)$-dimensional space $H_1^\perp$ is non-degenerate. Apply induction. We can repeat this procedure until we are left with a space of dimension 1 (if n is odd) or 2 (if n is even). As we have already done the classification in dimensions 1 and 2 we obtain the complete picture.

6.10 Theorem. *Let $V = V(n, q)$ be an n-dimensional vector space for odd q. There are up to equivalence precisely two non-degenerate symmetric bilinear forms $(,)$ on V. These have different discriminants.*

Let $n = 2m$ be even. Then either $V = H_1 \perp \cdots \perp H_m$ is orthogonal sum of m hyperbolic planes (a Gram matrix has determinant $(-1)^m$), or $V = H_1 \perp \cdots \perp H_{m-1} \perp A$ is orthogonal sum of $m-1$ hyperbolic planes and a 2-dimensional anisotropic space (a Gram matrix has determinant $(-1)^m\nu$, where $\nu \in \mathbb{F}_q$ is a non-square).

Let $n = 2m+1$ be odd. Then $V = H_1 \perp \cdots \perp H_m \perp \langle v \rangle$, where either $(v, v) = 1$ or $(v, v) = \nu$ (and $\nu \in \mathbb{F}_q$ is a fixed non-square).

This yields interesting point sets in the corresponding projective spaces. As we know these points sets are quadrics, sets of roots of homogeneous polynomials of degree 2. When $n = 2m+1$ we have two different bilinear (quadratic) forms in Theorem 6.10. The set of isotropic points are identical for these two cases. This is clear as multiplication of the bilinear form by a non-square transforms one type in the other.

6.11 Definition. *Let q be odd. Denote by $Q(2m, q) \subset PG(2m, q)$ the set of isotropic points of a non-degenerate symmetric bilinear (quadratic) form in $V(2m+1, q)$.*

$Q^+(2m-1, q) \subset PG(2m-1, q)$ (a **hyperbolic quadric***) is the set of isotropic points of a non-degenerate symmetric bilinear (quadratic) form in $V(2m, q)$ such that $V(2m, q)$ is orthogonal sum of m hyperbolic planes.*

$Q^-(2m-1, q) \subset PG(2m-1, q)$ (an **elliptic quadric***) is the set of isotropic points of a non-degenerate symmetric bilinear (quadratic) form in $V(2m, q)$ such that $V(2m, q)$ is orthogonal sum of $m-1$ hyperbolic planes and an anisotropic space.*

Let us illustrate with a specific example. Start from the quadratic form

(homogeneous polynomial of degree 2)

$$Q(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

What is the corresponding (equivalent) bilinear symmetric form? Just apply the formula from Definition 6.3. With $x = (x_1, x_2, x_3, x_4), y = (y_1, y_2, y_3, y_4)$ we obtain

$$(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y)) =$$

$$= \frac{1}{2}(\sum_i (x_i + y_i)^2 - \sum_i x_i^2 - \sum_i y_i^2) = \sum_i x_i y_i.$$

This is the standard dot product. The corresponding Gram matrix A is the unit matrix. Its determinant is a square, so the discriminant is 1, just as for the orthogonal sum of two hyperbolic planes. It follows that $V = V(4, q)$ with the dot product is hyperbolic. We can write V as orthogonal sum of two hyperbolic planes. Really? Let v_1, v_2, v_3, v_4 be the basis which was used implicitly (each Gram matrix is with respect to a basis). This means $(v_i, v_i) = 1$ and $(v_i, v_j) = 0$ for all $i \neq j$. Can we find a basis w_1, w_2, w_3, w_4 such that w_1, w_2 is a standard basis (consisting of isotropic vectors) for the hyperbolic plane $H_1 = \langle w_1, w_2 \rangle$, likewise w_3, w_4 for the hyperbolic plane $H_2 = \langle w_3, w_4 \rangle$ and $H_1 \perp H_2$? The calculations depend on the field. Use the field $\mathbb{F}_5$. The vector $w_1 = v_1 + 2v_2$ is isotropic, just as $w_2 = v_1 + 2v_3$. As $(w_1, w_2) = 1$ we have a standard basis for $H_1 = \langle w_1, w_2 \rangle$. There is no choice for H_2 as $H_2 = H_1^{\perp}$. A vector $av_1+bv_2+cv_3+dv_4$ is in H_2 if $a+2b = a+2c = 0$, equivalently $b = c,\ a = 3b$. This shows that we can use v_4 (case $a = b = c = 0$) and $3v_1+v_2+v_3$ as basis for H_2. In order to find a standard basis, such that the corresponding Gram matrix is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ we need two linearly independent isotropic vectors in H_2. This is left as an exercise.

It is not hard to count the points on the quadrics from Definition 6.11. Recall the conic from Chapter 5. It is a quadric in $PG(2, q)$ and has $q + 1$ points. This shows $|Q(2, q)| = q + 1$. When dealing with our quadratic form in Chapter 5 there was no need to distinguish between the characteristic 2 and the odd characteristic case. This indicates that the mechanism of the present chapter should generalize to cover the characteristic 2 case as well.

6.12 Definition. *Let q be odd. Denote by $h_m(c)$ the number of vectors $v \in V(2m, q)$ such that $(v, v) = c$ for the hyperbolic quadratic form. Analogously $e_m(c)$ denotes the representation number of c for the elliptic quadratic form.*

Let $p_m(c)$ the number of vectors $v \in V(2m+1, q)$ such that $(v, v) = c$ when $V(2m+1, q)$ is the orthogonal sum of m hyperbolic planes and $\langle v_0 \rangle = 1$. The corresponding representation numbers for the case when $\langle v_0 \rangle = \nu$ is a non-square are denoted by $p'_m(c)$.

Let us compute the representation numbers in Definition 6.12. Observe at first that a scalar multiple of a symmetric bilinear form is a symmetric bilinear form again. If the original form is elliptic, then all scalar multiples are elliptic. The same holds for hyperbolic forms. This shows that $h_m(c) = h_m(1)$ and $e_m(c) = e_m(1)$ for all $c \neq 0$. Also $h_1(0) = 2q - 1$ (there are two isotropic points in a hyperbolic plane) and $h_1(1) = q - 1$, analogously $e_1(0) = 1$ and $e_1(1) = q + 1$.

Consider the 4-dimensional hyperbolic case. The corresponding representation numbers are $h_2(c)$. Here the space can be written as orthogonal sum of two hyperbolic planes $H_1 \perp H_2$. Consider vector $u+v$, where $u \in H_1, v \in H_2$. How often will it happen that $Q(u+v) = 0$? As $(u, v) = 0$ this is equivalent with $Q(u) + Q(v) = 0$. There are two different situations: if u is isotropic, then also v is isotropic. The number of choices is $h_1(0)h_1(0) = (2q-1)^2$. The second situation occurs when u is not isotropic (there are $q^2 - h_1(0) = (q-1)^2$ such vectors). Then $Q(v)$ must have a specific nonzero value. This occurs $h_1(1) = q - 1$ times. In the second situation we count $(q-1)^3$ possibilities, all in all

$$h_2(0) = (2q-1)^2 + (q-1)^3 = q^3 + q^2 - q.$$

The by now familiar argument shows that $h_2(c) = h_2(1)$ for all $c \neq 0$. This number is therefore

$$h_2(1) = (q^4 - h_2(0))/(q-1) = q^3 - q.$$

In the elliptic 4-dimensional case the arguments are very similar. This time the space is orthogonal sum of a hyperbolic plane and an anisotropic plane. We use the numbers $h_1(c)$ and $e_1(c)$. The number of isotropic vectors is the sum of $h_1(0)e_1(0) = h_1(0) = 2q-1$ (the first situation) and $(q^2 - h_1(0))e_1(1)$ (the second situation). We obtain

$$e_2(0) = (2q-1) + (q-1)^2(q+1) = q^3 - q^2 + q.$$

The general case is not harder. We have the following obvious recursive relation for the hyperbolic case:

$$h_m(0) = h_{m-1}(0)h_1(0) + (q^{2(m-1)} - h_{m-1}(0))h_1(1) = qh_{m-1}(0) + (q-1)q^{2(m-1)}.$$

We saw $h_2(0) = q(q^2 + q - 1),\ h_3(0) = q^2(q^3 + q - 1)$, by induction $h_m(0) = q^{m-1}(q^m + q - 1)$. The corresponding number of isotropic points is $(h_m(0) - 1)/(q-1)$. Fortunately $h_m(0) - 1 = (q^m - 1)(q^{m-1} + 1)$. For $m = 1$ we obtain two isotropic points as it should be. The next case is $h_2(0) = q(q^2 + q - 1)$, leading to $(q + 1)^2$ isotropic points.

6.13 Theorem. *Let q be odd. We have $h_m(0) = q^{m-1}(q^m + q - 1)$, $h_m(1) = q^{m-1}(q^m - 1)$ and*

$$|Q^+(2m - 1, q)| = \frac{(q^m - 1)(q^{m-1} + 1)}{q - 1}.$$

For the elliptic type we can use this result:

$$e_m(0) = h_{m-1}(0) \cdot 1 + (q^{2(m-1)} - h_{m-1}(0))(q + 1) = q^{2m-1} - q^m + q^{m-1}.$$

As $e_m(0) - 1 = (q^m + 1)(q^{m-1} - 1)$ we obtain the following result:

6.14 Theorem. *Let q be odd. We have $e_m(0) = q^{m-1}(q^m - q + 1)$, $e_m(1) = q^{m-1}(q^m + 1)$ and*

$$|Q^-(2m - 1, q)| = \frac{(q^m + 1)(q^{m-1} - 1)}{q - 1}.$$

By the same argument $p_m(0) = h_m(0) \cdot 1 + \dfrac{q^{2m} - h_m(0)}{2} \cdot 2 = q^{2m}$ and $p_m(1) = h_m(0) \cdot 2 + h_m(1)(q - 2) = q^m(q^m + 1)$.

6.15 Theorem. *Let q be odd. We have $p_m(0) = p'_m(0) = q^{2m}$, $p_m(1) = p'_m(\nu) = q^m(q^m + 1)$ and $p_m(\nu) = p'_m(1) = q^m(q^m - 1)$. Further*

$$|Q(2m, q)| = \frac{q^{2m} - 1}{q - 1}.$$

It is a natural question to determine the maximum dimension of totally isotropic subspaces.

6.16 Definition. *The* **Witt index** *of a symmetric bilinear form in odd characteristic is the largest dimension of a totally isotropic subspace.*

6.17 Proposition. *Let q be odd and $(,)$ a non-degenerate symmetric bilinear form. If W is an i-dimensional totally isotropic subspace, we can find i hyperbolic planes such that*

$$V = H_1 \perp H_2 \perp \cdots \perp H_i \perp R,$$

where R is non-degenerate. The Witt index is the largest such i.

Proof. Let $0 \neq v_1 \in W$. Find a hyperbolic plane $H_1 = \langle v_1, w_1 \rangle$. We have $V = H_1 \perp H_1^\perp$ and $H_1^\perp \cap W = w_1^\perp \cap W$ of dimension at least $m - 1$. The first claim follows by induction. The second claim is immediate as we see i-dimensional totally isotropic subspaces if we have a subspace of the form $H_1 \perp H_2 \perp \cdots \perp H_i$. $\square$

Proposition 6.17 shows that the hyperbolic and the elliptic quadric in $PG(2m-1, q)$ have different Witt indices. The hyperbolic quadric $Q^+(2m-1, q)$ has Witt index m (the space is orthogonal sum of m hyperbolic planes), whereas the elliptic quadric $Q^-(2m-1, q)$ has Witt index $m-1$ (we can split off $m-1$ hyperbolic planes).

1. Let V be a vector space over $\mathbb{F}_q$. A **bilinear form** is a mapping : $V \times V \rightarrow \mathbb{F}_q$, which is $\mathbb{F}_q$-linear in both arguments.

2. Choose a basis $e_1, \ldots, e_n$ of V. A bilinear form is determined by the values (e_i, e_j). The **Gram matrix** A with entries $a_{ij} = (e_i, e_j)$ determines the bilinear form (because of bilinearity).

3. In matrix notation $(x, y) = xAy^t$, where $x = (x_1, \ldots, x_n)$, analogously for y, and y^t is the transposed.

4. The bilinear form is **symmetric** ($(x, y) = (y, x)$ for all x, y) if and only if its Gram matrix is symmetric.

5. If $(x, y) = (y, x) = 0$ we call x, y **orthogonal.** x is **isotropic** if $(x, x) = 0$ (x is orthogonal to itself).

6. The **radical** of the bilinear form consists of all vectors, which are orthogonal to all vectors in V. The radical is a subspace. The bilinear form is **non-degenerate** if the radical is $\{0\}$.

7. The bilinear form is non-degenerate if and only if the Gram matrix is regular ($det(A) \neq 0$).

8. A symmetric bilinear form $(,)$ determines the **quadratic form** $Q(x) = (x, x)$.

9. A quadratic form is described (with respect to a fixed basis) by a homogeneous polynomial of degree 2 in n variables: $Q(x_1, \ldots, x_n) = \sum_{1 \leq i \leq j \leq n} q_{ij} x_i x_j$.

10. Let q be odd, $dim(V) = n$ over $\mathbb{F}_q$ and $(,)$ a symmetric bilinear form with Gram matrix A.

11. The quadratic form determines the equivalent symmetric bilinear form via $(x, y) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y))$ (this is also called the **polarization** of Q).

- The polarization in coordinates: each q_{ii} yields entry $a_{ii} = q_{ii}$, each $q_{ij}, i < j$, yields $a_{ij} = a_{ji} = q_{ij}/2$ in the Gram matrix.

- If $(,)$ is non-degenerate and $W \subset V$ is a subspace of dimension m, then $W^{\perp}$, the set of all vectors orthogonal to all of W, has complementary dimension $n - m$.

- A 2-dimensional non-degenerate space either is a hyperbolic plane (2 isotropic points) or is anisotropic (no isotropic point). In each case the space is uniquely determined.

- The **discriminant** is 1 if $det(A)$ is a square, it is -1 otherwise. It is an invariant of the space.

- The **Witt index** d is the largest dimension of a totally isotropic subspace (meaning that Q vanishes on the subspace).

- Let $n = 2m$. Then either V (non-degenerate) is orthogonal sum of m hyperbolic planes (the **hyperbolic case** or $+$ case, Witt index m) or it is orthogonal sum of $m - 1$ hyperbolic planes and an anisotropic 2-dimensional space (the **elliptic** case or $(-)$case , Witt index $m - 1$). The discriminants are different.

- Let $n = 2m+1$. Then V is orthogonal sum of m hyperbolic planes and a 1-dimensional (the **parabolic** case, Witt index m).

- The corresponding sets of isotropic points (quadrics) in $PG(n-1, q)$ are denoted $Q^+(2m-1, q)$ (hyperbolic), $Q^-(2m-1, q)$ (elliptic), $Q(2m, q)$ (parabolic).

- We determine the representation numbers of the quadratic forms (given $c \in \mathbb{F}_q$, how many vectors have $Q(x) = c$?) and the number of points of the quadrics.

Problems

1. Consider the bilinear form with Gram matrix $A = \begin{pmatrix} 1 & 0 & 2 \\ \mu & 1 & 0 \\ \lambda & 0 & 1 \end{pmatrix}$ in odd characteristic. For which values of λ, μ is $(,)$ degenerate?

2. Let $Q(x_1, x_2, x_3) = x_1^2 - x_3^2 + 2x_1x_3 + 3x_2x_3$ over $\mathbb{F}_p, p \neq 2$. Determine the corresponding bilinear form. In which characteristic is this non-degenerate?

3. Consider the preceding example in characteristic 5. Find a standard basis.

4. Consider the bilinear form with Gram matrix $A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & a \\ 1 & 1 & a & 1 \end{pmatrix}$ in odd characteristic. When is it degenerate?

5. Consider the preceding problem with $a = 1$. Over which fields is the bilinear form hyperbolic (respectively elliptic)?

6. Consider the preceding problem over $\mathbb{F}_5$. Find a totally isotropic line.

7. Let $V = V(4, 5)$ with the standard dot product. Find a standard basis such that the hyperbolic space V is written as orthogonal sum of two hyperbolic planes.

Chapter 7

Symplectic bilinear forms

Before covering the characteristic 2 case of quadratic forms it is natural to consider symplectic bilinear forms first. A bilinear form is **symplectic** if all vectors are isotropic:

$$(x, x) = 0 \text{ for all } x \in V.$$

As $0 = (x + y, x + y) = (x, x) + (y, y) + (x, y) + (y, x) = (x, y) + (y, x)$, a symplectic form is also skew-symmetric:

$$(y, x) = -(x, y) \text{ for all } x, y.$$

This shows that the **radical** of V is

$$Rad(V) = \{x|(x, V) = 0\} = \{x|(V, x) = 0\}.$$

A symplectic form is non-degenerate if $Rad(V) = \{0\}$. Symplectic forms are bilinear forms, which are described by skew-symmetric Gram matrices A. The form is non-degenerate if and only if $det(A) \neq 0$. Also, the same proof as in the symmetric case shows that the dual

$$W^{\perp} = \{x|(x, W) = 0\} = \{x|(W, x) = 0\}$$

of a subspace W has complementary dimension if $(,)$ is non-degenerate.

Clearly there can be no non-degenerate symplectic form in dimension $n = 1$. Let $n \geq 2$. Let $v_1 \neq 0$. Find w_1 such that $(v_1, w_1) = 1$. The 2-dimensional subspace $\langle v_1, w_1 \rangle$ is non-degenerate. It follows $V = \langle v_1, w_1 \rangle \perp \langle v_1, w_1 \rangle^{\perp}$. Proceeding inductively we see that $n = 2m$ must be even and V is the orthogonal sum of $\langle v_i, w_i \rangle, i = 1, 2, \ldots, m$ such that the Gram matrix with respect to this basis has $(2,2)$-submatrices $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ along the main diagonal. All other entries are 0. Such a basis $\{v_1, v_2, \ldots, v_m\} \cup \{w_1, w_2, \ldots, w_m\}$ is known as a **symplectic basis.**

1. A bilinear form is **symplectic** if all vectors are isotropic.
2. It is then **skew-symmetric** ($(x, y) = -(y, x)$ for all x, y).
3. A non-degenerate symplectic form on $V = V(n, q)$ exists only if $n = 2m$ is even.
4. V is then orthogonal sum of m 2-dimensional spaces, each with Gram matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
5. A corresponding basis is a **symplectic basis.**

Problems

1. Find the number of totally isotropic lines in $PG(3, q)$ of a non-degenerate symplectic bilinear form on $V(4, q)$.
2. Consider $V(4, 2)$ with a non-degenerate symplectic form. Count the symplectic bases.

Chapter 8

Quadratic forms in characteristic 2

A quadratic form on $V = V(n,q)$ is a homogeneous polynomial of degree 2 in n unknowns:

$$Q(x_1,\ldots,x_n) = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i<j} a_{ij}x_i x_j.$$

We can describe it by a symmetric matrix $A = (a_{ij})$. In Chapter 6 we saw that quadratic forms are equivalent to symmetric bilinear forms, provided the underlying field has characteristic $\neq 2$. In the characteristic 2 case this is not true. It can be expected that quadratic polynomials should behave in a special way in characteristic 2. The reason is that squaring is a field automorphism (see Lemma 1.5). In fact, let $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ be a symmetric $(2,2)$-matrix in characteristic 2 and $(,)$ the corresponding bilinear form. Let $x = (x_1, x_2)$. Then $(x,x) = ax_1^2 + dx_2^2$. The non-diagonal entry b does not show up at all. This shows that (x,x) does not give us the general case of a quadratic form. We should start from quadratic forms. In the case of dimension 2 matrix A describes the quadratic form $Q(x) = ax_1^2 + dx_2^2 + bx_1x_2$. We have $Q(x+y) + Q(x) + Q(y) = bx_1y_2 + bx_2y_1$, a symmetric bilinear form, which is also symplectic. This leads to a formal definition of quadratic forms in characteristic 2.

8.1 Definition. *Let $V = V(n,q)$, where q is a power of* 2. *A* **quadratic form** *on V is a mapping $Q : V \longrightarrow \mathbb{F}_q$ such that*

- *$Q(\lambda x) = \lambda^2 Q(x)$ for all $\lambda \in \mathbb{F}_q, x \in V$, and*
- *$(x, y) = Q(x + y) + Q(x) + Q(y)$ is a bilinear form.*

Observe that the bilinear form is by force symplectic. The quadratic form carries more information. The underlying symplectic form is uniquely determined by the quadratic form, but not the other way around. In fact, if the quadratic form is described by the symmetric matrix A, then the Gram matrix of the corresponding symplectic bilinear form is obtained by putting zeroes in the main diagonal of A.

8.2 Definition. *A vector v is* **singular** *with respect to the quadratic form Q in characteristic 2, if $Q(v) = 0$. A subspace is* **totally singular** *if Q vanishes on it. It is* **asingular** *if $Q(x) = 0$ only when $x = 0$. The dimension of the radical V_0 of the underlying symplectic bilinear form is the* **index** *$i = i(Q)$ of Q. The quadratic form Q is* **non-degenerate** *if Q is asingular on the radical of the bilinear form.*

The 1-dimensional case

In case $n = 1$ we have without restriction $Q(v_1) = 1$. Observe that in contrast to the odd characteristic case every field element is a square.

The 2-dimensional case

We know from Chapter 7 that $(,)$ is either identically 0 or non-degenerate. Assume it is $\equiv 0$. Then $Q(x + y) = Q(x) + Q(y)$ and $Q(\lambda x) = \lambda^2 Q(x)$ (Q is **semi-linear**). It follows that Q is degenerate in this case. Either Q is identically 0 on V or it has a radical of dimension 1 (and the projective line corresponding to $V = V(2, q)$ has one singular point).

So let $(,)$ be non-degenerate. Assume at first there is $v \neq 0$ such that $Q(v) = 0$. Choose w' such that $(v, w') = 1$. Consider $w = w' + tv$. We have $(v, w) = (v, w') = 1$ and (see Definition 8.1) $Q(w) = Q(w')+Q(tv)+(w', tv) = Q(w') + t$. Choosing $t = Q(w')$ we obtain $Q(w) = 0$. This shows that we can find a symplectic basis v, w such that $Q(v) = Q(w) = 0$. Call $\langle v, w\rangle$ a **quadratic hyperbolic plane.**

Now assume there is no singular nonzero vector. Choose v such that $Q(v) = 1$ and w such that $(v, w) = 1$. Let $Q(w) = a$. An arbitrary vector

$sv+tw$ has $Q(sv+tw) = s^2+t^2a+st$. This expression has to be $\neq 0$ whenever $(s,t) \neq 0$. When $t = 0$ this is satisfied. Let $t \neq 0$. Division by t shows that we can assume without restriction $t = 1$. We must have $s^2 + s \neq a$ for all $s \neq 0$.

8.3 Lemma. *Let $q = 2^f$ and $tr : \mathbb{F}_q \rightarrow \mathbb{F}_2$ the trace. The elements of $\mathbb{F}_q$, which can be written in the form $x + x^2$, are precisely the elements of trace 0.*

Proof. Recall Definition 1.11. The trace is a nonzero $\mathbb{F}_2$-linear function : $\mathbb{F}_q \rightarrow \mathbb{F}_2$. The kernel of tr is therefore an $(f-1)$-dimensional subspace of $\mathbb{F}_q$ (where we see $\mathbb{F}_q$ as an f-dimensional vectorspace over $\mathbb{F}_2$. Also $tr(x) = tr(x^2)$ (squaring is a field automorphism), therefore $tr(x^2+x) = tr(x) + tr(x) = 0$. On the other hand, the function $f : \mathbb{F}_q \rightarrow \mathbb{F}_q$ defined by $f(x) = x + x^2$ is linear over $\mathbb{F}_2$. Its kernel is $\{0, 1\}$. Its image is therefore a hyperplane. This proves the claim. $\square$

Lemma 8.3 is known as the additive version of **Hilbert's theorem 90**. Observe that method for solving quadratic equations is very different from the odd characteristic case (additive instead of multiplicative).

It follows from Lemma 8.3 that a satisfies the condition if and only if $tr(a) = 1$. Replacing w by $w' = w+sv$ we obtain $Q(w') = a+s^2+s$. When s varies $Q(w')$ varies over all elements of trace 1. This shows that Q is uniquely determined.

To sum up, we have seen that there are precisely two non-degenerate quadratic forms in dimension 2, the quadratic hyperbolic space and the **asingular** space. The general procedure is very similar to the odd characteristic case, but the role played by the distinction between squares and non-squares in the odd case is replaced by the distinction between field elements of traces 0 or 1. Next we show that asingular spaces have dimension ≤ 2 and use induction to describe all non-degenerate quadratic forms in arbitrary characteristic, just as in the odd characteristic case.

8.4 Proposition. *There is no asingular quadratic form on a vector space of dimension $n > 2$.*

Proof. Choose $0 \neq v \in V$ and $w \notin \angle v\rangle$ such that $(v,w) = 0$. Then $Q(v + tw) = Q(v) + t^2Q(w)$. Either $Q(w) = 0$ or we can choose t such that $Q(v + tw) = 0$. $\square$

The Witt index is defined and determined just as in the odd characteristic case.

8.5 Definition. *The* **Witt index** *of a quadratic form is the largest dimension of a totally singular subspace.*

8.6 Proposition. *Let Q be a non-degenerate quadratic form in characteriostic 2. If W is an i-dimensional totally singular subspace, we can find i quadratic hyperbolic planes such that*

$$V = H_1 \perp H_2 \perp \cdots \perp H_i \perp R,$$

where R is non-degenerate.

The proof is just as the proof of Proposition 6.17. We are in the same position as in the odd characteristic case. When $n \geq 3$ we can split off a quadratic hyperbolic plane. Let $n = 2m$. After splitting off $m-1$ quadratic hyperbolic planes the remaining 2-dimensional space is either another quadratic hyperbolic plane (this is the hyperbolic case) or it is asingular (we have an elliptic quadratic form in this case). If $n = 2m+1$ we can split off m quadratic hyperbolic planes and are left with a 1-dimensional space.

8.7 Definition. *Let q be even. Denote by $Q(2m,q) \subset PG(2m,q)$ the set of singular points of a non-degenerate quadratic form in $V(2m+1,q)$.*

$Q^+(2m-1,q) \subset PG(2m-1,q)$ (a **hyperbolic quadric***) is the set of singular points of a non-degenerate quadratic form in $V(2m,q)$ such that $V(2m,q)$ is orthogonal sum of m quadratic hyperbolic planes.*

$Q^-(2m-1,q) \subset PG(2m-1,q)$ (an **elliptic quadric***) is the set of singular points of a non-degenerate quadratic form in $V(2m,q)$ such that $V(2m,q)$ is orthogonal sum of $m-1$ hyperbolic planes and an asingular space. The representation numbers $h_m(c)$, $e_m(c), p_m(c)$ are defined as the number of vectors x such that $Q(x)=c$ in the respective cases.*

Observe that in characteristic 2 there is only one 1-dimensional quadratic form and therefore only one non-degenerate quadratic form in odd dimension. It follows from Proposition 8.6 that the two non-degenerate quadratic forms in even dimension have different Witt indices.

With respect to the standard basis for the quadratic hyperbolic plane we have $Q(sv+tw) = st$ and it follows $h_1(0) = 2q-1, h_1(c) = q-1$ for $c \neq 0$. The usual argument shows $e_1(c) = q+1$ for all $c \neq 0$. This shows that the representation numbers $h_m(c),\ e_m(c)$ are the same as in the odd case. For odd dimension things are easier yet. As the 1-dimensional non-degenerate quadratic form represents each field element precisely once we obtain $p_m(c) = q^{2m}$ for all c in characteristic 2. In particular we conclude that the formulas for the number of points on our quadrics are the same as in odd characteristic.

1. Let $q = 2^f$ and $V = V(n, q)$.
2. In characteristic 2 quadratic forms (described by homogeneous quadratic polynomials) and symmetric bilinear forms are not equivalent. The quadratic form carries more information.
3. The bilinear form underlying the quadratic form Q is $(x, y) = Q(x + y) + Q(x) + Q(y)$.
   It is symplectic (and symmetric).
4. In coordinates: if $Q(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le n} q_{ij} x_i x_j$, then each q_{ij} for $i < j$ contributes $a_{ij} = a_{ji} = q_{ij}$ in symmetric positions of the Gram matrix of $(,)$. Terms q_{ii} do not contribute to A at all. If A is given, then the coefficients $q_{ij}, i < j$ of the quadratic form are determined. The diagonal terms q_{ii} are arbitrary.
5. A vector x is **singular** if $Q(x) = 0$, a subspace is **totally singular** if all its vectors are singular.
   A space is **asingular** if $Q(x) \neq 0$ for all $x \neq 0$.
6. Q is **non-degenerate** if it is asingular on the radical of $(,)$.
7. A 2-dimensional non-degenerate space V is either a **quadratic hyperbolic plane** (2 singular points) or it is asingular (no singular point).
8. Let $tr : \mathbb{F}_q \to \mathbb{F}_2$ be the trace. In the analysis the distinction between elements of trace 0 and of trace 1 takes the place of the distinction between squares and non-squares in odd characteristic.
9. The general structure is the same as in odd characteristic, the words singular, asingular replacing isotropic, anisotropic.
10. The formulas for the number of singular points are the same as in odd characteristic.

Problems

1. Let $Q(x_1, x_2, x_3) = x_1^2 + x_1x_2 + \omega x_2x_3$, a quadratic form on $V(3,4)$. Determine the Gram matrix of the underlying symplectic form and its radical. Is Q degenerate?

Chapter 9

Unitary bilinear forms

$\mathbb{F}_{q^2} \mid \mathbb{F}_q$ has the involutory field automorphism $\alpha \mapsto \overline{\alpha} = \alpha^q$ (meaning that it has order 2 : applying the Frobenius twice is the identity). Clearly $\alpha = \overline{\alpha}$ if and only if $\alpha \in \mathbb{F}_q$. The situation is analogous to complex conjugation in the complex field. Observe that for $x \in \mathbb{F}_{q^2}$ we have $x + \overline{x} = tr(x)$, where $tr : \mathbb{F}_{q^2} \to \mathbb{F}_q$ is the trace (see Chapter 1). The mapping $N(x) = x\overline{x} = x^{q+1}$ is known as the **norm.** As $N(x)^{q-1} = x^{q^2-1} = 1$ whenever $x \neq 0$, we have $N(x) \in \mathbb{F}_q$. As the multiplicative group of $\mathbb{F}_{q^2}$ is cyclic of order $q^2 - 1$ each nonzero element of $\mathbb{F}_q$ is the norm of precisely $q + 1$ elements of $\mathbb{F}_{q^2}$.

$(,)$ is called a **unitary** (or **hermitian**) scalar product if $(,)$ is biadditive,

$$(w, v) = \overline{(v, w)}$$

and

$$(\alpha \cdot v, w) = \alpha(v, w),\ (v, \alpha w) = \overline{\alpha}(v, w)$$

always hold (one speaks of **sesquilinearity** with respect to the Frobenius, semilinearity in the second component).

Observe that $(v, v) = \overline{(v, v)}$, hence $(v, v) \in \mathbb{F}_q$ for all v. The scalar product is non-degenerate if no nonzero vector is orthogonal to the whole space. In the sequel we consider the case that $(,)$ is non-degenerate and $n \geq 2$. Assume at first $(x, x) = 0$ for all $x \in V$. Choose x_1, x_2 such that $(x_1, x_2) = 1$. Then $(x_1 + \lambda x_2, x_1 + \lambda x_2) = \lambda + \overline{\lambda}$. For a suitable choice of λ this will be $\neq 0$, leading to a contradiction. We have shown that we can find x_1 such that $(x_1, x_1) \neq 0$. As $(x_1, x_1) \in \mathbb{F}_q$ and $(\lambda x_1, \lambda x_1) = \lambda^{q+1}(x_1, x_1)$ we can choose $(x_1, x_1) = 1$. We get $V = \langle x_1 \rangle \perp x_1^\perp$. By induction we obtain an orthonormal basis, that is the Gram matrix can be chosen to be the identity matrix.

9.1 Proposition. *The number of $v \in V = V(n, q^2)$ such that $(v, v) = 0$ is $u_n(0) = q^{n-1}(q^n + (-1)^n(q-1))$. For every $0 \neq \alpha \in \mathbb{F}_q$ the number of vectors $v \in V$ satisfying $(v, v) = \alpha$ is $u_n(1) = q^{n-1}(q^n - (-1)^n)$.*

Proof. Denote by $f(n, \alpha)$ the number of $v \in V$ such that $(v, v) = \alpha$. Write $v = \sum_{i=1}^n a_i x_i$. Then $(v, v) = \sum_i a_i^{q+1}$. We see that each value $\alpha \neq 0$ occurs equally often, in other words $f(n, \alpha) = f(n, 1)$. Clearly $q^{2n} = f(n, 0) + (q-1)f(n, 1)$. Distinguishing according to the value of a_n we obtain: $f(n, 0) = f(n-1, 0) + (q^2-1)f(n-1, 1)$. The proposition follows by induction. □

The totally isotropic points form the Hermitian variety. Clearly the number of these points is $(f(n, 0) - 1)/(q^2 - 1)$. Distinguishing cases n even and n odd Proposition 9.1 yields the following:

9.2 Theorem. *Denote by $H(n-1, q^2) \subset PG(n-1, q^2)$ (the* **Hermitian variety***) the set of isotropic points of the non-degenerate unitary form on $V(n, q^2)$. If $n = 2d$ we have*

$$|H(2d-1, q^2)| = \frac{(q^{2d-1}+1)(q^{2d}-1)}{q^2-1}.$$

In case $n = 2d + 1$

$$|H(2d, q^2)| = \frac{(q^{2d}-1)(q^{2d+1}+1)}{q^2-1}.$$

Case $n = 2$

It follows from Theorem 9.2 that a line has exactly $q + 1$ isotropic points if the restriction of the unitary form to the corresponding 2-dimensional vector space W is non-degenerate. If the radical of W is 1-dimensional, the line has exactly one isotropic point. The final case is that the unitary form vanishes on the line. We see that the intersection of a Hermitian variety defined over $\mathbb{F}_{q^2}$ with a line has either $q + 1$ or 1 or $q^2 + 1$ points.

What is the Witt index (the dimension of the largest totally isotropic subspace)? Let $e_1, \ldots, e_n$ be an orthonormal basis and fix $c \in \mathbb{F}_{q^2}$ such that $N(c) = -1$. Then

$$\langle e_1 + ce_2, e_3 + ce_4, \ldots \rangle$$

is a totally isotropic subspace of dimension d, where $n = 2d$ or $n = 2d+1$. We claim that d is the Witt index. This is seen using a by now familiar argument.

Let $W = \langle v_1, \ldots, v_m \rangle$ be totally isotropic. Because of non-degeneracy we can find w_1 such that $(v_1, w_1) = 1$. Replacing w_1 by $w_1 + sv_1$ for suitable s we see that we can assume w_1 is isotropic. Then $H_1 = \langle v_1, w_1 \rangle$ is non-degenerate. It follows $V = H_1 \perp H_1^\perp$, the orthogonal complement $H_1^\perp$ is non-degenerate and $dim(W \cap H_1^\perp) = m - 1$. Proceeding inductively we see that we can split off m such 2-dimensional subspaces. In particular $n \geq 2m$. We have seen the following:

9.3 Theorem. *Let $n = 2d$ or $n = 2d + 1$ and $(,)$ a non-degenerate unitary form on $V(n, q^2)$. The Witt index (largest dimension of a totally isotropic subspace) is d.*

1. Let q be a prime-power, $V = V(n, q^2)$ and $\overline{x} = x^q$ the image of x under the Frobenius automorphism.
2. $tr(x) = x + x^q \in \mathbb{F}_q$ (the trace), $N(x) = x\overline{x} = x^{q+1}$ (the norm).
3. A biadditive scalar product is **sesquilinear** if
$$(w, v) = \overline{(v, w)}, \ (\alpha v, w) = \alpha(v, w), \ (v, \alpha w) = \overline{\alpha}(v, w).$$
4. The **unitary** (non-degenerate sesquilinear) scalar product is uniquely determined. There is an orthonormal basis.
5. The isotropic points form the **Hermitian variety.**
6. We determined the number of points of the Hermitian variety.
7. The Witt index is d, where $n = 2d$ or $n = 2d + 1$.

Chapter 10

Quadrics in $PG(2,q)$ and in $PG(3,q)$

Let us go back to the quadratic form $Q(x) = x_1x_3 - x_2^2$ from Chapter 5. In odd characteristic the corresponding symmetric bilinear form is

$$(x,y) = \frac{1}{2}x_1y_3 + \frac{1}{2}x_3y_1 - x_2y_2.$$

Its Gram matrix is $A = \begin{pmatrix} 0 & 0 & 1/2 \\ 0 & -1 & 0 \\ 1/2 & 0 & 0 \end{pmatrix}$. We saw this example in the opening stages of Chapter 6. As $det(A) = 1/4$, we have that Q is non-degenerate. Its isotropic points form $Q(2,q)$. We know that $Q(2,q)$ has $q+1$ points (this is the special case $m=1$ of Theorem 6.15). As Q has Witt index 1 there is no totally isotropic line. It follows that no more than 2 points of $Q(2,q)$ are collinear. Lines l containing just one point P of $Q(2,q)$ (tangents) correspond to degenerate 2-dimensional spaces. As P is in the radical of l we must have $l = P^\perp$. This shows that every point of $Q(2,q)$ is on precisely one tangent. All this confirms results obtained in Chapter 5 by calculations with coordinates and by combinatorial counting.

This example also illustrates the bijection between symmetric bilinear forms and quadratic forms in odd characteristic. Terms $c_{ii}X_i^2$ lead to entry $a_{ii} = c_{ii}$ in the Gram matrix, whereas each mixed term $c_{ij}X_iX_j$ yields entries $a_{ij} = a_{ji} = c_{ij}/2$ in the Gram matrix.

Let now q be a power of 2. The Gram matrix is $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. Recall that A describes the underlying symplectic bilinear form, not the quadratic form. We have $det(A) = 0$, so $(,)$ is degenerate. Recall that this is always so in odd dimension. Non-degenerate symplectic forms exist only in even dimension. The radical of $(,)$ is $V_0 = \langle e_2 \rangle$. We have $Q(e_2) = Q(0,1,0) = 1$, so Q is non-degenerate. As $(e_1, e_3) = 1$ and $Q(e_1) = Q(e_3) = 0$, the space $H_1 = \langle v_1, v_3 \rangle$ is a quadratic hyperbolic plane.

Recall from Chapter 5 that each oval in $PG(2,q)$ is embedded in a unique hyperoval when q is a power of 2. The tangents of the oval meet in a unique point, the nucleus N, which complements the oval to a hyperoval. Can we confirm that from our quadratic point of view? Well, the nucleus corresponds to the radical of the underlying symplectic form. This radical is 1-dimensional, so we talk about a point in $PG(2,q)$, and Q is non-degenerate on it, so $N \notin Q(2,q)$. More importantly all tangent lines meet in N as the tangent lines are the duals of the points of the quadric, and vectors from N are orthogonal to everything.

We have recovered all the information that we gathered earlier concerning $Q(2,q)$. Let us go one dimension higher.

Consider $Q(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2$. In characteristic 2 we have $Q(x) = (\sum_i x_i)^2$. As Q vanishes on a 3-dimensional subspace it clearly is degenerate. Let q be odd. Recall that in dimension $n = 4$ and odd characteristic the hyperbolic quadric has discriminant 1, the elliptic quadric has discriminant -1, meaning that the determinant of the Gram matrix is a square in the hyperbolic case, a non-square in the elliptic case. The Gram matrix of our form is the unit matrix. It follows that the quadric is hyperbolic. We know $|Q(3,q)| = (q+1)^2$. It is not all that easy to confirm this by concrete calculations with coordinates.

Use the standard form corresponding to a decomposition in hyperbolic planes, in odd characteristic. The Gram matrix is $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.

The quadratic form is $Q(x) = 2x_1x_2 + 2x_3x_4$. We can use just as well $Q(x) = x_1x_2 + x_3x_4$. There should be $(q+1)^2$ isotropic points.

There are $4q$ points such that $x_1x_2 = x_3x_4 = 0$. All remaining points must have all coordinates $\neq 0$. We can choose $x_1 = 1$. For arbitrary nonzero values of x_2, x_3 we can find precisely one value of x_4 such that $x_4x_3 = x_2$. This gives us $(q-1)^2$ points all of whose coordinates are $\neq 0$. We count $|Q^+(3,q)| = (q-1)^2 + 4q = (q+1)^2$, as predicted.

The standard form for an elliptic 4-dimensional quadratic form in odd characteristic corresponds to the Gram matrix $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\nu \end{pmatrix}$. The quadratic form is $Q(x) = 2x_1x_2 + x_3^2 - \nu x_4^2$. We expect q^2+1 isotropic points. This is in fact easy to count directly. Assume at first $x_3 = x_4 = 0$. As either $x_1 = 0$ or $x_2 = 0$ this gives us the two points $\langle e_1\rangle, \langle e_2\rangle$. Let now $(x_3, x_4) \neq (0,0)$. Then $x_3^2 - \nu x_4^2 \neq 0$. We can choose $x_1 = 1$ and obtain a unique solution x_2. This gives us $q^2 - 1$ points. We have confirmed that $|Q^-(3,q)| = q^2 + 1$.

As the Witt index is 1 there are no totally isotropic lines. It follows that each line intersects $Q^-(3,q)$ in at most 2 points (see Section 6). Consider planes and how they intersect the elliptic quadric. Let E be a plane. Let $P \in Q^-(3,q)$ and $E = P^\perp$. Then E intersects $Q^-(3,q)$ only in P. We see that each point $P \in Q^-(3,q)$ is contained in precisely one such **tangent plane.** If E is not one of the q^2+1 tangent planes, then the restriction of the quadratic form to E is non-degenerate. It follows that $|E \cap Q^-(3,q)| = q+1$ in this case, and the $q+1$ points of intersection form an oval. Observe that all this is true in any characteristic.

10.1 Proposition. *Let $Q^-(3,q) \subset PG(3,q)$ be the elliptic quadric. Each line contains at most 2 points of the quadric. Each point $P \in Q^-(3,q)$ is on precisely one plane $P^\perp$ meeting $Q^-(3,q)$ only in P. Any plane, which is not one of the q^2+1 tangent planes, meets $Q^-(3,q)$ in $q+1$ points, which form an oval.*

Consider the hyperbolic quadric in $PG(3,q)$ now. As the Witt index is 2, totally singular lines exist. Let us count them. Fix $P \in Q^+(3,q)$. Then $P^\perp$ is the orthogonal sum of P and a hyperbolic plane. It follows that P is on precisely 2 totally singular lines (in odd characteristic we can speak equivalently of isotropy instead of singularity). By double counting we see that the total number of totally singular lines is $2(q+1)^2/(q+1) = 2(q+1)$.

Can we determine the structure of this family of $2(q+1)$ lines? Fix one such line, l, let $P_1, \dots, P_{q+1}$ be the points of l and $g_i, i = 1, \dots, q+1$ the second totally singular line through P_i (aside of l). Assume lines g_i, g_j intersect in a point R. Then the plane E through the points P_i, P_j, R contains the lines l, g_i, g_j. It is clear that E must be totally singular, which is impossible as the Witt index is 2. We conclude that the lines g_i are mutually disjoint (one says: skew). It follows that the lines $g_i, i = 1, 2, \ldots, q+1$ partition the points of $Q^+(3, q)$. Starting from one of the g_i we see that l also is part of such a parallel class.

10.2 Proposition. *Let $Q^+(3, q) \subset PG(3, q)$ be the hyperbolic quadric. There are $2(q+1)$ totally singular lines. These come in two parallel classes of $q+1$ each. Each parallel class partitions the points of the quadric, whereas two lines from different parallel classes intersect in a point.*

Observe that the real work has been done in the preceding chapters. In the present chapter we are harvesting. As another example consider the non-degenerate symplectic form in $V(4, q)$. Again we would like to count the totally isotropic lines. Fix a point $P \in PG(3, q)$. We can choose $P = \langle v_1 \rangle$, where v_1 belongs to a symplectic basis. Then $P^\perp = P \perp \langle v_2, w_2 \rangle$. It follows that P is on precisely $q+1$ totally isotropic lines. Double counting shows that the number of totally isotropic lines equals the number of points of $PG(3, q)$, which is $(q^4-1)/(q-1)$. What type of geometry do the q^3+q^2+q+1 totally isotropic lines form? Clearly any two lines intersect in at most one point, and any two points are on at most one line. We can prove another important property: Let l be a totally isotropic line and $P \notin l$. Then $|P^\perp \cap l| = 1$. This means that there is precisely one point $Q \in l$ such that the line PQ is totally isotropic.

10.3 Proposition. *Consider the non-degenerate symplectic bilinear form on $V(4, q)$. The number of totally isotropic lines equals the number of points in $PG(3, q)$, and each point is on $q+1$ totally isotropic lines. Given a totally isotropic line l and a point $P \notin l$ there is precisely one point $Q \in l$ such that P and Q are collinear on a totally isotropic line.*

Chapter 11

Designs, projective planes and generalized quadrangles

In the previous chapter we drew some consequences from the properties of quadrics and bilinear forms. In this chapter we want to put this in perspective.

11.1 Definition. *A* t-**design***, more precisely a design* $t-(v,k,\lambda)$*, consists of a set (the ground set)* Ω *of* v *points and a family* $\mathcal{B}$ *of* k*-subsets of* Ω *called* **blocks.** *The defining property is:*

- *Each* t*-subset of* Ω *is contained in precisely* λ *blocks.*

A **Steiner** t**-design** *is a* t*-design with* $\lambda = 1$*. We denote it* $S(t,k,v)$*.*

Sometimes a more general notion of a design is used, where the blocks form a multiset. In this terminology the designs of Definition 11.1 would be called **simple designs.** Designs are a basic structure of modern combinatorial theory. We have encountered designs in several places already.

11.2 Theorem. *Let* q *be a prime-power. The projective plane* $PG(2,q)$ *is a Steiner* 2*-design* $S(2, q+1, q^2+q+1)$*.*

Here we use the lines as blocks. More generally we can consider the hyperplanes of $PG(n,q)$ as blocks.

11.3 Theorem. *Let* q *be a prime-power. Using the hyperplanes as blocks we obtain a design*

$$2-(\frac{q^{n+1}-1}{q-1}, \frac{q^n-1}{q-1}, \frac{q^{n-1}-1}{q-1}).$$

Consider the elliptic quadric in $PG(3,q)$. Any three points are in a unique plane, and this plane is of course not a tangent plane. This proves the following:

11.4 Theorem. *Let q be a prime-power and $\Omega = Q^-(3,q) \subset PG(3,q)$ the elliptic quadric. Use as blocks the intersections of Ω with planes, which are not tangent planes. This yields a Steiner 3-design $S(3, q+1, q^2+1)$.*

Although design theory has a rather long history by now, some of its basic problems remain unsolved. For example, no Steiner t-design is known for $t > 5$ and no infinite family of Steiner t-designs is known for $t > 3$. This indicates that the circle geometries of Theorem 11.4 are rather interesting.

Projective planes have found more interest than any other type of designs. Here is a definition.

11.5 Definition. *A* **projective plane** *of* **order** *n is a Steiner system $S(2, n+1, n^2+n+1)$.*

11.6 Proposition. *A projective plane of order n has n^2+n+1 lines (just as many as points). Any two lines intersect in precisely one point.*

Proof. This is a combinatorial triviality. The number of lines through a point is $(n^2+n)/n = n+1$ (fix a point P; the points on the lines through P must partition the n^2+n points $\neq P$.) By double counting the point-line incidences we see that the number of lines equals the number of points.

Fix a line l. Each of the $n+1$ points on l is on n lines $\neq l$. We count n^2+n lines, each of which intersects l in one point. These are all lines $\neq l$. $\square$

Proposition 11.6 justifies to call these designs projective planes. They have the same combinatorial properties as a 2-dimensional projective geometry. In particular the notion of a projective plane is self-dual in the following sense: if we start from a projective plane of order n and interpret the lines as points and the points as lines, then this dual structure is a projective plane of order n again.

In general 2-designs with equally many blocks as points are known as **symmetric designs.** Symmetric designs with $\lambda = 1$ are projective planes (see the Problems section).

The only examples of projective planes we know are the planes $PG(2,q)$ whose points are the 1-dimensional and whose lines are the 2-dimensional subspaces of a 3-dimensional vector space over $\mathbb{F}_q$. The terminology has been chosen such that the number of elements q of the underlying field is the order of the projective plane. Many constructions are known for projective planes whose order is a prime-power but not a prime. Not a single projective plane is known whose order n is composite. The existence of such projective planes constitutes a famous open problem.

Consider the Hermitian variety $H(2, q^2)$. By Theorem 9.2 it has $q^3 + 1$ points. As the Witt index is 1 there is no totally isotropic line. It follows that each line of $PG(2, q^2)$ intersects $H(2, q^2)$ either in 1 or in $q + 1$ points. It is clear how to obtain a design on the points of $H(2, q^2)$ (also called the **unital**). Let the intersections with lines containing $q+1$ points of the unital be the blocks. This yields a Steiner 2-design $S(2, q + 1, q^3 + 1)$ embedded in the projective plane $PG(2, q^2)$. As an example, in case $q = 3$ we obtain a design $S(2, 4, 28)$ embedded in $PG(2, 9)$.

We have encountered examples of another famous type of combinatorial structure as well.

11.7 Definition. *A* **generalized quadrangle** *of order (s, t) consists of a ground set Ω and a family of subsets (***lines***) of Ω such that the following hold:*

- *Each line has $s + 1$ points.*
- *Each point is on $t + 1$ lines.*
- *Any two points are on at most one common line.*
- *Given a line l and a point $P \notin l$ there is exactly one point $R \in l$ such that p and R are on a common line (collinear).*

11.8 Proposition. *A $GQ(s, t)$ has $(s+1)(st+1)$ points and $(t+1)(st+1)$ lines. The dual of a $GQ(s, t)$ is a $GQ(t, s)$.*

Proof. At first note that, just as in the case of projective planes, the notion of a generalized quadrangle is self-dual, and the dual of a $GQ(s, t)$ has order (t, s). Fix a line l of a $GQ(s, t)$. Each of its $s+1$ points is on t lines $\neq l$. Each of these lines contains s points not on l. All in all we count $(s+1)+(s+1)ts$ points. The axioms make sure that in this way we count each point precisely once. Because of duality the number of lines is obtained by substituting $s \mapsto t \mapsto s$ in this expression. $\square$

A GQ is called **thick** is both s and t are > 1, **thin** otherwise. Thick GQ are more interesting than thin ones. The thinnest of all GQ is $GQ(1,1)$, which is nothing but a quadrangle. This helps to explain the term *generalized quadrangle.* Going back to projective planes, a projective plane of order 1 simply is a triangle. It makes sense therefore to consider projective planes as *generalized triangles.*

We saw that the $(q+1)^2$ points of the hyperbolic quadric in $PG(3,q)$ and the totally singular lines form a grid. This is a (trivial) example of a $GQ(q,1)$. More interesting examples can be derived from quadrics and bilinear or sesquilinear forms. The basic observation is the following:

whenever the Witt index is 2, *the totally isotropic or totally singular lines will define a generalized quadrangle.* This is completely obvious for us. Checking all the cases when the Witt index is 2 we arrive at the following five families of examples.

The thick finite classical GQ

The symplectic case

The points of $PG(3,q)$ and the totally isotropic lines with respect to the symplectic bilinear form define a GQ, which is known as $W(q)$. Clearly each line has $q+1$ points, so $s=q$. Fix a point P. Then $P^{\perp}=P\perp l$, where l is a line. It follows that P is on $q+1$ totally isotropic lines, hence $t=q$. The GQ $W(q)$ has order (q,q).

In order to gain confidence, let us work out the case $W(2)$. As we know the Gram matrix can be chosen as $A=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$. Denote the corresponding basis of $V(4,2)$ by e_1,e_2,e_3,e_4. The points of $PG(3,2)$ can be identified with the nonzero vectors. As every vector is orthogonal to itself (this defines the symplectic form) a line $\langle v,w\rangle$ is totally isotropic if and only if $(v,w)=0$. An example of such a line would be $\langle e_1,e_3\rangle$, containing the points e_1,e_3,e_1+e_3. These totally isotropic lines are precisely the lines of $W(2)$. We know that there are 15 lines. Start from $l_1=\{e_1,e_3,e_1+e_3\}$. The lines intersecting l_1 are

$$\begin{aligned}
l_2 &= \{e_1, e_4, e_1 + e_4\} \\
l_3 &= \{e_1, e_3 + e_4, e_1 + e_3 + e_4\} \\
l_4 &= \{e_3, e_2, e_3 + e_2\} \\
l_5 &= \{e_3, e_1 + e_2, e_1 + e_2 + e_3\} \\
l_6 &= \{e_1 + e_3, e_2 + e_4, e_1 + e_2 + e_3 + e_4\} \\
l_7 &= \{e_1 + e_3, e_1 + e_2 + e_4, e_2 + e_3 + e_4\}
\end{aligned}$$

This completes the list of lines intersecting l_1. As we saw in the general counting argument all the points off l_1 appear, each exactly once. As $t = 2$, each of those 12 points must appear twice on the remaining 8 lines (fortunately $12 \times 2 = 8 \times 3$). Start from e_4. We have $e_4^\perp = \langle e_4, e_1, e_2 \rangle$. Aside of l_2 this yields the following lines through e_4 :

$$\begin{aligned}
l_8 &= \{e_4, e_2, e_2 + e_4\} \\
l_9 &= \{e_4, e_1 + e_2, e_1 + e_2 + e_4\}
\end{aligned}$$

It is clear how to continue.

The parabolic case

The non-degenerate quadric in $PG(4, q)$ has Witt index 2. This generalized quadrangle $Q(4, q)$ has of course $s = q$. Let P be a singular point. The orthogonal complement of P in $P^\perp$ is a conic in $PG(2, q)$ (with $q+1$ points). This shows $t = q$. The order of $Q(4, q)$ is (q, q). By Proposition 11.8 the number of its points and lines is $(q + 1)(q^2 + 1)$. This is in agreement with Theorem 6.15 in case $m = 2$ (it was noted earlier that this formula is true in any characteristic).

The elliptic case

Generalized quadrangle $Q(5, q)$ arises from the elliptic quadric in $PG(5, q)$. Clearly we have $s = q$. As $P^\perp = P \perp H_1 \perp A$, where A is anisotropic, and $H_1 \perp A$ is elliptic, we have $t = q^2$ (recall that an elliptic quadric in $PG(3, q)$ has $q^2 + 1$ points). The order of $Q(5, q)$ is (q, q^2).

The first Hermitian case

The non-degenerate Hermitian form in $PG(3, q^2)$ has Witt index 2, and therefore defines $H(3, q^2)$, a GQ with $s = q^2$. As a non-degenerate Hermitian form on a line has $q + 1$ points, we have $t = q$. The order of $H(3, q^2)$ is (q^2, q).

The second Hermitian case

The non-degenerate Hermitian form in $PG(4, q^2)$ has Witt index 2 as well. The corresponding GQ is known as $H(4, q^2)$. Clearly $s = q^2$. As a non-degenerate Hermitian variety in $PG(2, q^2)$ (in other words: a unital) has $q^3 + 1$ points, the order of $H(4, q^2)$ is (q^2, q^3).

The S_6-GQ

We give an elementary description of a GQ of order $(2, 2)$. By Proposition 11.8 it must have 15 points and 15 lines.

Start from a set $S = \{1, 2, 3, 4, 5, 6\}$ of 6 elements. The points of our GQ are the unordered pairs from S. There are $\binom{6}{2} = 15$ such pairs (points). Write a typical point as (12), for example. This is a shorter expression than $\{1, 2\}$. The lines are the fixed-point-free involutions in S_6, or, expressed differently, the partitions of S into three pairs. A typical line would be (12|34|56), containing the points (12), (34) and (56).

Basic combinatorial counting shows that there are 15 lines. Any two points are on at most one line. In fact, two points are not collinear if the corresponding pairs intersect. They are on precisely one line otherwise. For example, (12) and (13) are of course not on a common line (partition), whereas (12) and (35) are on the line (12|35|46). The main axiom of GQ is also easily verified. For example, consider the line $l = (12|34|56)$ and the point $P = (13) \notin l$. The unique point on l, which is collinear with P, is (56). It follows that we have indeed a GQ of order $(2, 2)$. Clearly every permutation of the underlying set S yields a symmetry of the GQ. This shows that the symmetric group S_6 is a subgroup of the automorphism group. It is in fact the complete automorphism group.

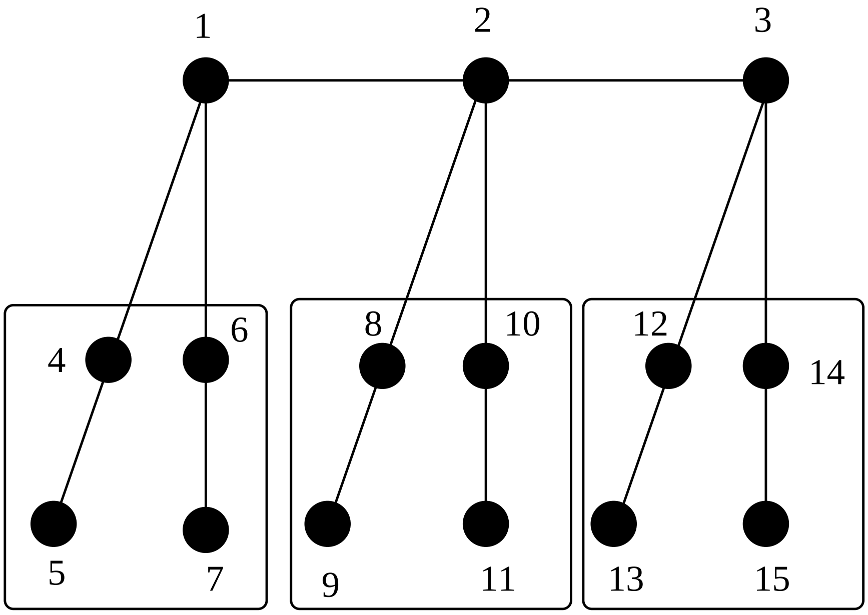

Figure 11.1: Towards $GQ(2, 2)$

The uniqueness of $GQ(2, 2)$

We have seen two constructions of a GQ of order $(2, 2)$, the symplectic $W(2)$ and the S_6-GQ. However, it is an elementary combinatorial fact that there is essentially only one GQ of order $(2, 2)$. Let us check on this. The points are $1, 2, \ldots, 15$. Denote the lines by $l_1, \ldots, l_{15}$. Choose $l_1 = \{1, 2, 3\}$. The main axiom implies that each of $1, 2, 3$ is on 2 further lines, and that the points on those lines partition all points outside l_1. Up to renumbering the points we can choose notation such that the first 7 lines (l_1 and those that intersect it) are as follows:

$$
\begin{aligned}
l_1 &= \{1, 2, 3\} \\
l_2 &= \{1, 4, 5\} \\
l_3 &= \{1, 6, 7\} \\
l_4 &= \{2, 8, 9\} \\
l_5 &= \{2, 10, 11\} \\
l_6 &= \{2, 12, 13\} \\
l_7 &= \{2, 14, 15\}
\end{aligned}
$$

Each of the remaining lines must pick exactly one point collinear with 1, one collinear with 2 and one collinear with 3, see Figure 11.1. In fact, assume l picks up two points collinear with 1, for example $4,6$. Then point 1 and line l violate the basic axiom (absence of triangles of collinear points). The general picture is: there are 2 more lines through each of the points off l_1, each such line picks up one point from each of the boxes of Figure 11.1. Up to renumbering the points we can assume that the points collinear with 4 are $8,10$ (from the second box) and $12,14$ (from the third box). The points collinear with 5 are then $9,11,13,15$. Again up to renumbering we can choose

$$l_8=\{4,8,12\},\ l_9=\{4,10,14\}.$$

We can choose the numbering of lines such that $l_{10}\supset\{5,9\}$ and $l_{11}\supset\{5,11\}$. The remaining points on these two lines are $13,15$. The situation is tight already. Assume

$$l_{10}=\{5,9,15\},\ l_{11}=\{5,11,13\}.$$

As $6,7$ can be interchanged we can choose notation such that

$$l_{12}=\{6,8,x\}$$

where x is from the last box. However, there is no suitable choice for x. 12 is excluded as the pair $8,12$ is on l_8 already, $x=13$ produces the triangle $8,12,13$ of collinear points, $x=14$ is impossible as l_{12} would form a triangle with l_8,l_9 and $x=15$ would produce the triangle $15,8,9$. It follows that we have reached a dead end. The only possibility is

$$l_{10}=\{5,9,13\},\ l_{11}=\{5,11,15\}.$$

As before we can choose $l_{12}=\{6,8,x\}$. All but one point of the last box is excluded: 12 is obviously impossible, $x=13$ gives the triangle $8,9,13$ and $x=14$ would lead to the triangle $4,8,14$. It follows $l_{12}=\{6,8,15\}$. The third line through 6 cannot contain 8 or 9, and 11 is excluded as well as this would produce the triangle $6,11,15$. It follows that $l_{13}=\{6,10,y\}$, and as before we can exclude all but one possibilities for $y:$ points $14,15$ are out for obvious reasons and $y=12$ would produce the triangle $4,10,12$. We have

$$l_{12}=\{6,8,15\},\ l_{13}=\{6,10,13\}.$$

The completion is uniquely determined:

$$l_{14}=\{7,9,14\},\ l_{15}=\{7,11,12\}.$$

11.9 Theorem. *There is up to equivalence only one GQ of order* $(2,2)$.

1. A **t-design** is described by a family of k-element subsets (blocks) from a fixed v-element set. The main axiom is: every t-element subset is contained in precisely λ blocks.
2. The special case of $\lambda = 1$ are the **Steiner designs** $S(t,k,v)$.
3. When $\lambda = 1, t = 2$ blocks are often called **lines.**
4. $S(2, n+1, n^2+n+1)$ are called **projective planes** of order n. Our $PG(2,q)$ is a projective plane of order q.
5. The main axiom of a **generalized quadrangle** is: for every line l and point $P \notin l$ there is precisely one point $Q \in l$, which is collinear (on a common line) with P.
6. Each quadric, bilinear forms or sesquilinear form of Witt index 2 defines a GQ. These are the classical GQ.
7. The S_6-GQ has 15 points and 15 lines.
8. It is the uniquely determined GQ of order $(2,2)$.

Problems

1. Complete the list of lines of $W(2)$.

2. Determine the number of blocks of an $S(3, q+1, q^2+1)$ (equivalently: the number of planes in $PG(3, q)$, which are not tangent planes of the elliptic quadric).

3. Let P be a point of the unital in $PG(2, q^2)$. Determine the number of blocks and the number of tangent lines to the unital that pass through P.

4. Prove the following: a symmetric Steiner 2-design is a projective plane.

5. Identify the S_6-GQ with one of the classical GQ.

Chapter 12

The small Witt designs

We have seen some Steiner 2-designs and 3-designs, see Theorem 11.4. In fact, it is very hard to construct Steiner designs with higher values of t. No Steiner t-design with $t > 5$ is known and only the following Steiner designs with $t > 3$ are known:

$$S(4,5,11),\ S(5,6,12),\ S(4,7,23),\ S(5,8,24).$$

These are the famous **Witt designs.** In the present chapter we want to construct the two smaller of those. This also serves as an application of the basic correspondence between codes and projective geometry.

Start from the elliptic quadric $Q^-(3,3)$. The 3-dimensional elliptic quadric $Q^-(3,q)$ has q^2+1 points, so we have 10 points in $PG(3,3)$ (see Theorem 6.14 and Chapter 10). We saw that these 10 points form a Steiner design $S(3,4,10)$ (this is the special case $q=3$ of Theorem 11.4).

Now use the basic relationship with codes, Theorem 3.5. Write representatives for the 10 points of $Q^-(3,3)$ as columns of a generator matrix. In order to determine the minimum distance of this ternary code we have to check the hyperplane intersections of $Q^-(3,3)$. Hyperplanes in $PG(3,3)$ are $PG(2,3)$, so hyperplanes are planes. They correspond to 3-dimensional vector subspaces. The symmetric bilinear form defining our $Q^-(3,3)$ gives a symmetric bilinear form on this subspace (subplane). As the Witt index is 1 no line is contained in $Q^-(3,3)$. The restriction to a plane either is itself non-degenerate (intersection $q+1=4$) or it has a radical of vector space dimension 1 (intersection size 1, the tangent planes). We used these facts in the proof of Theorem 11.4 already. As the hyperplane intersection size is ≤ 4, the minimum distance is $\geq 10-4=6$. This shows the following:

12.1 Proposition. *The points of the elliptic quadric* $Q^-(3,3)$ *define a* $[10,4,6]_3$*-code.*

Let us make this concrete. What we need is the Gram matrix A, a ternary $(4,4)$-matrix, which is symmetric (so as to define a symmetric bilinear form), invertible (this makes the form non-degenerate) and whose determinant is a non-square (this makes it elliptic). The only non-square in $\mathbb{F}_3$ is $2=-1$. One obvious choice is $A = diag(1,1,1,2)$, a diagonal matrix of determinant 2. As we want to continue working with this code we prefer a generator matrix in a standard form, containing the unit matrix. We want the elementary vectors $e_1=(1,0,0,0),\dots,e_4$ to be isotropic. This means A should have zeroes along the diagonal. After testing a couple of possibilities we see that the following choice can be made:

$$A = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix}$$

The corresponding quadratic form is

$$(x,x) = x_1x_4 - (x_1x_2 + x_1x_3 + x_2x_3 + x_2x_4 + x_3x_4)$$

The elementary vectors are on the corresponding $Q^-(3,3)$, and these are the only isotropic points with more than one coordinate $=0$. If $x_1=0$, then $x_2=x_3=x_4$, which we can choose to be $=1$. The mapping $x_1 \leftrightarrow x_4$ is a symmetry. This gives us the point $(1:1:1:0)$. Similarly we obtain one more point with $x_2=0$ and by symmetry also a point with $x_3=0$. So far we have 8 of the expected 10 points. The two remaining points have all their coordinates $\neq 0$. Write the quadratic form as

$$x_1x_4 - x_2x_3 - (x_1+x_4)(x_2+x_3)$$

If $x_1+x_4=0$ then $x_2+x_3=0$ and vice versa. These are the only remaining solutions. This yields the following matrix generating a $[10,4,6]_3$-code:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1 & 1 & 2 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 2 & 2 \end{pmatrix}$$

We ask now if this code can be extended to a $[12, 6, 6]_3$-code. This means that we are looking for a generator matrix of such a code, which has the following form:

$$\left(\begin{array}{cc|cccc|cccccc}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 2 & 1 & 2 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 2 & 0 & 2 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 2 & 2 \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\
1 & 0 & 0 & 0 & 0 & 0 & y_1 & y_2 & y_3 & y_4 & y_5 & y_6
\end{array}\right)$$

where we recognize G in the Northeast corner. Number the rows of this matrix $v_1, \dots, v_6$. Observe that the code is not changed if we add a linear combination of the first 4 rows to v_5 or v_6. This is why we can choose v_5, v_6 to have zeroes in columns $3, 4, 5, 6$. It remains to determine the vectors $x = (x_1, \dots, x_6),\ y = (y_1, \dots, y_6)$. Let us collect information about these vectors. As v_5, v_6 should have weight at least 6 we have that x and y have weight ≥ 5. Assume $x_2, \dots, x_6$ are all nonzero. Then we can find a linear combination of v_1 and v_5 with 3 zeroes in the last 5 columns. This word has then weight < 6, contradiction. It follows $wt(x) = wt(y) = 5$ (the same discourse is valid for v_6) and $x_1, y_1 \neq 0$. In which coordinate does x, y have entry 0? Coordinates $2, 3, 4$ are impossible as linear combinations with v_2, v_3, v_4 show. The same is true of y. It follows that x and y have their 0 in one of the coordinates 5 or 6. Linear combinations of v_5, v_6 show that those zeroes do not occur in the same coordinate. It follows that either $x_5 = y_6 = 0$ or $x_6 = y_5 = 0$. We claim that these choices are equivalent. In fact, if the second version is true interchange the role of v_5 and v_6, then flip the first and second column.

We can assume $x_5 = y_6 = 0$ and, eventually after replacing v_5 or v_6 by their negatives and doing the same to the first or send columns, $x_6 = y_5 = 1$. As $wt(v_1 - v_5) \geq 6$ it follows that at most one of x_2, x_3, x_4 can be $= 1$. Likewise $v_4 + v_5$ shows that at most one of x_1, x_3, x_4 can be $= 2$. These two facts together show $x_1 = 1, x_2 = 2$ and by symmetry then also $y_1 = 1, y_2 = 2$. It is clear now that everything is uniquely determined. As an example, $v_2 + v_5$ shows $x_4 = 2$. We arrive at the generator matrix given in the following theorem:

12.2 Theorem. *The following is a generator matrix of a self-dual* $[12, 6, 6]_3$-

code:

$$\left(\begin{array}{cc|cccc|cccccc}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 2 & 1 & 2 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 2 & 0 & 2 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 2 & 2 \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & 1 & 2 & 1 & 2 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 2 & 1 & 1 & 0
\end{array}\right)$$

In fact, it is easy to check that any two of the v_i have dot product 0. For example, $v_i \cdot v_i = 0$ follows from the fact that the v_i have weight 6. Let $\mathcal{G}_{12}$ be the code generated by G. As it is contained in its dual and because the dual space has complementary dimension it follows $\mathcal{G}_{12}^{\perp} = \mathcal{G}_{12}$. This makes it much easier to verify that the minimum weight is indeed 6. In fact, because of self-duality each codeword must have weight divisible by 3. It suffices therefore to show that the minimum weight is > 3. This is easy to do.

Let us come back to the geometric description. The columns of the generator matrix of $\mathcal{G}_{12}$ form a set $\mathcal{P}$ of points in $PG(5,3)$ with the property that at least 6 are outside any given hyperplane, equivalently at most $12-6=6$ are in any given hyperplane. The hyperplanes of $PG(5,3)$ are spaces $PG(4,3)$, and the hyperplanes of hyperplanes (such a subspace of codimension 2 is also known as a **secundum**) are $PG(3,3)$. Let S be a secundum. How many points from our set can be contained in a secundum? Observe that S is contained in precisely $q+1=4$ hyperplanes. If 5 points of $\mathcal{P}$ are in S, we count $|\mathcal{P}| \leq 5+1+1+1+1=9$, contradiction. Is $|S \cap \mathcal{P}| = 4$ possible? The same counting argument shows

$$12 = |\mathcal{P}| \leq 4+2+2+2+2 = 12.$$

This is possible, but every hyperplane containing S must intersect $\mathcal{P}$ in precisely the maximum number of 6 points. On the other hand, this situation occurs all the time: two points determine a line, together with a third point outside this line they determine a plane, together with a fourth point outside this plane they determine a secundum. This shows that any 3 points must determine a plane (they cannot be on a line), any 4 points must determine a secundum (they cannot be on a plane) and any 5 points of $\mathcal{P}$ determine a hyperplane (they are not on a secundum). But hold, we have here a Steiner system: any 5 points of $\mathcal{P}$ are on a hyperplane, and this hyperplane contains 6 points of $\mathcal{P}$. It follows that the points of $\mathcal{P}$ and the hyperplanes meeting $\mathcal{P}$ in 6 points form the points and blocks of a $S(5,6,12)$. If we omit a point and use only the blocks containing that point we obtain $S(4,5,11)$. These are the Steiner designs promised in the introduction to this chapter.

All these structures are important and have exceptional properties. The self-dual code $\mathcal{G}_{12}$ with parameters $[12,6,6]_3$ and a code $[11,6,5]_3$ that can be derived from it are the **ternary Golay codes.** The $S(4,5,11)$ and $S(5,6,12)$ are the **small Witt designs** and their groups of automorphisms are the small **Mathieu groups** M_{11} and M_{12}.

Chapter 13

Symmetry groups

13.1 Definition. *Let Ω be a finite set and $\mathcal{B}$ a family of subsets of Ω (the* **blocks***). An* **automorphism** *(or* **symmetry***) of the incidence structure $(\Omega, \mathcal{B})$ is a permutation π of Ω which respects the block structure, in other words: for every block B the image $\pi(B)$ is a block as well.*

All the finite geometries we considered thus far are special cases of incidence structures. Examples are designs, generalized quadrangles, projective geometries and totally singular subspaces on quadrics. It is always important to study the group of symmetries of a given incidence structure. We start from projective geometries.

Let $V = V(n, q)$ be an n-dimensional vector space over $\mathbb{F}_q$. Each bijective linear mapping : $V \rightarrow V$ is a symmetry of the incidence structure whose blocks consist of the subspaces of V. The corresponding automorphism group is the general linear group.

13.2 Definition. *The* **general linear group** *$GL(n, q)$ is the group of all bijective linear mappings from V to V.*

Using a fixed basis ($e_1, e_2, \ldots, e_n$ say), linear mappings : $V \rightarrow V$ are described by (n, n)-matrices. A matrix A describes an element of $GL(n, q)$ if A is invertible, equivalently if $det(A) \neq 0$. Recall from basic linear algebra that a linear mapping is uniquely determined by the images of a basis. The linear mapping $f : V \rightarrow V$ is bijective if and only if the image of a basis is a basis again. The **order** (number of elements) of $GL(n, q)$ is therefore the number of order bases of our vector space V.

13.3 Proposition. *The* **order** *(number of elements) of $GL(n,q)$ is*

$$|GL(n,q)| = (q^n - 1)(q^n - q) \dots (q^n - q^{n-1}).$$

This is clear, by basic counting principles: there are $q^n - 1$ choices for the first element v_1 of a basis $v_1, v_2, \dots, v_n$. Once v_1 has been chosen there are $q^n - q$ choices for v_2 (all vectors not belonging to the subspace $\langle v_1 \rangle$ generated by v_1). Continuing in this fashion the formula in Proposition 13.3 is obtained.

As an example, $|GL(3,2)| = (8-1)(8-2)(8-4) = 168$: there are 168 invertible $(3,3)$-matrices with entries in $\mathbb{F}_2$.

The points of $PG(n-1,q)$ are the 1-dimensional subspaces of V. It is clear that each element of $GL(n,q)$ gives us an automorphism of $PG(n-1,q)$. In the language of basic group theory we have a **permutation representation** of $GL(n,q)$ ob the points of $PG(n-1,q)$. In order to determine the order of the corresponding symmetry group we have to determine the kernel Z of the permutation representation. It consists of the elements $z \in GL(n,q)$ which map every point $P \in PG(n-1,q)$ to itself. Fix a basis and represent z by a matrix Z. Choosing $P = \langle e_i \rangle$ we see that Z must be a diagonal matrix. Points $\langle e_i - e_j \rangle$ show that all diagonal entries of Z must be the same, in other words Z must be a **scalar matrix** $diag(\lambda, \lambda, \dots, \lambda)$ for some $0 \neq \lambda \in \mathbb{F}_q$. Each such diagonal matrix does indeed map each point to itself. Clearly Z is a normal subgroup of $GL(n,q)$. It is isomorphic to the multiplicative group of the field, and therefore is cyclic (see Theorem 1.8). We have seen the following:

13.4 Theorem. *The group Z of invertible scalar matrices is a normal subgroup of $GL(n,q)$. It is cyclic of order $q-1$ and is the kernel of the permutation representation of $GL(n,q)$ on the points of $PG(n-1,q)$. The factor group $PGL(n,q) = GL(n,q)/Z$ (the* **projective general linear group***) is a symmetry group of $PG(n-1,q)$. We have*

$$|PGL(n,q)| = \frac{|GL(n,q)|}{q-1}.$$

When $q = 2$ we have $PGL(n,2) = GL(n,2)$. In particular $PGL(3,2)$ has order 168.

Recall from basic group theory the important operation of **conjugation:** the conjugate of g under h is

$$g^x = x^{-1}gx.$$

A subgroup $H \subseteq G$ is a **normal subgroup** if $H^g = g^{-1}Hg = H$ for all $g \in G$ (the conjugates of elements of H are in H again). In the case of $GL(n,q)$ we see that $det(A^B) = det(A)$ for all $A, B \in GL(n,q)$. This implies in particular that the matrices of determinant 1 form a normal subgroup, the special linear group. Also, det is a group homomorphism from $GL(n,q)$ onto the multiplicative group of $\mathbb{F}_q$. It follows from the first homomorphism theorem of basic group theory that the order of the special linear group is $\frac{1}{q-1}$ times the order of the general linear group.

13.5 Proposition. *The* **special linear group** $SL(n,q)$ *consists of all* (n,n)*-matrices of determinant* 1 *and is a normal subgroup of* $GL(n,q)$. *Its order is the same as that of* $PGL(n,q)$.

The kernel of the permutation representation of $SL(n,q)$ on the projective points is $SL(n,q) \cap Z$. When does a scalar matrix $diag(\lambda, \dots, \lambda)$ have determinant 1? When is $\lambda^n = 1$? Observe that λ is from a cyclic group of order $q-1$. It follows that $SL(n,q) \cap Z$ has order $gcd(n, q-1)$.

13.6 Theorem. *The intersection* $SL(n,q) \cap Z$ *has order* $gcd(n,q-1)$. *The factor group* $PSL(n,q) = SL(n,q)/(SL(n,q) \cap Z)$ *(the* **projective special linear group***) is the symmetry group of* $PG(n-1,q)$ *induced by* $SL(n,q)$. *Its order is*

$$|PSL(n,q)| = \frac{|SL(n,q)|}{gcd(n,q-1)}.$$

Nothing happens in the binary case: $GL(n,2) = SL(n,2) = PGL(n,2) = PSL(n,2)$. As another example consider $n = 2, q = 7$. We have $|GL(2,7)| = (7^2-1)(7^2-7)$, $|SL(2,7)| = (7+1)7(7-1)$ and $|PSL(2,7)| = \frac{1}{2}(7+1)7(7-1) = 168$, strangely the same order as $GL(3,2)$.

We note without proof that the groups $PSL_n(q)$ are in general simple groups. A group is **simple** if it has no normal subgroups (and therefore no factor groups) aside of the unit subgroup and the group itself. The only exceptions, $PSL_n(q)$ which are not simple, are $PSL_2(2)$ of order 6 and $PSL_2(3)$ of order 24. The smallest non-abelian simple group is A_5 of order 60. Both $PSL_2(5)$ and $PSL_2(4) = SL_2(4)$ are isomorphic to A_5.

We are not going to study the structure of these groups in detail. Just a hint that certain structural features are not that hard to see: consider the upper triangular matrices with entries 1 in the diagonal. These matrices are in $SL_n(q)$ and they form a subgroup P (the product of two such triangular matrices has this form again, as has the inverse of such a matrix). The order of P is the power of q with exponent the number of cells under the main diagonal. This number of cells is $1+2+\cdots+(n-1)=n(n-1)/2$. It follows $|P| = q^{n(n-1)/2}$. Comparison with the order of $GL_n(q)$ (Proposition 13.3) shows that this is the full power of q dividing the order. It follows that P is a **Sylow-p-subgroup** of $GL_n(q)$ and of $SL_n(q)$, where q is a power of the prime p. Clearly we could have worked with upper triangular matrices instead.

We know in principle that $PGL(n,q)$ is not in general the complete symmetry group of $PG(n-1,q)$. In fact, let $q=p^f$ for $f>1$. The **Frobenius automorphism** σ is a field automorphism (see Lemma 1.5). It induces a mapping : $V \rightarrow V$ and a mapping : $PG(n-1,q) \rightarrow PG(n-1,q)$, which clearly is an automorphism. The corresponding groups

$$P\Gamma L(n,q) = PGL(n,q)\langle\sigma\rangle, P\Sigma L(n,q) = PSL(n,q)\langle\sigma\rangle$$

have orders f times the order of the underlying linear group (which is a normal subgroup).

Each quadratic form, bilinear form or sesquilinear form defines a subgroup of $PGL(n,q)$, the subgroup consisting of those elements which respect the corresponding structure. This is a general principle, which describes an important family of finite groups.

Start with the symplectic case. We have $V=V(2m,q)$ equipped with a non-degenerate symplectic bilinear form, see Chapter 7.

13.7 Definition. *The* **symplectic group** $Sp(2m,q)$ *consists of the elements* $A \in GL(2m,q)$*, which satisfy* $(Ax,Ay)=(x,y)$ *for all* $x,y \in V$*.*

This expresses the idea that A should respect the symplectic structure. We can express this in a different way, as follows: fix a symplectic basis $\{v_1,\ldots,v_m\} \cup \{w_1,\ldots,w_m\}$. Then $A \in Sp(2m,q)$ if and only if the image of our symplectic basis is a symplectic basis again. This shows that the order $|Sp(2m,q)|$ equals the number of ordered symplectic bases. This leads to the same kind of counting argument that allowed us to determine the order of $GL(n,q)$. The number of (isotropic) nonzero vectors is $q^{2m}-1$. These are the candidates for v_1. Once v_1 is chosen, its partner w_1 must satisfy $(v_1,w_1)=1$. As $v_1^\perp$ is a hyperplane and all nonzero values of (v_1,x) occur equally often, the number of candidates for w_1 is $(q^{2m}-q^{2m-1})/(q-1) = q^{2m-1}$. The number of choices for the first pair v_1,w_1 of elements of a symplectic basis is $q^{2m-1}(q^{2m}-1)$. In case $m=1$ we are done, and therefore $|Sp(2,q)| = (q-1)q(q+1)$. If $m>1$ we can use induction ($\langle v_1, w_1\rangle^\perp$ is a $2(m-1)$-dimensional space with a non-degenerate symplectic bilinear form). As the sum of all odd natural numbers up to $2m-1$ is m^2 we obtain

13.8 Theorem.

$$|Sp(2m,q)| = q^{m^2}(q^{2m}-1)(q^{2(m-1)}-1)\cdot\dots\cdot(q^2-1).$$

The S_6-GQ again

We have $|Sp(4,2)| = 2^4\cdot 15\cdot 3 = 6!$, the same order as S_6. In fact, we saw in Theorem 11.9 of Chapter 11 that there is up to isomorphy only one GQ of order $(2,2)$. In particular the S_6-GQ is identical to the generalized quadrangle $W(2)$ derived from the symplectic 4-dimensional geometry on $V(4,2)$. As S_6 is an automorphism group of the S_6-GQ, $Sp(4,2)$ is a group of automorphisms of $W(2)$ and these groups happen to have the same order it is natural to expect that they are the same group (isomorphic). In order to prove this it suffices to show that S_6 is the full automorphism group of the S_6-GQ. Let us do this. Let $G\supseteq S_6$ be the automorphism group of the S_6-GQ. We know that S_6 is transitive on lines and the stabilizer of a line induces the symmetric group S_3 on its 3 points. The same is therefore true of the potentially larger group G. We can use the combinatorial work done in the previous chapter, where we now use the setup of the S_6-GQ, identifying points $1,2,3$ on line l_1 with $12,34,56$ and so on. As the stabilizer in S_6 of these three points is transitive on the 4 points off l_1, which are collinear with 12 (the points in the first box in the terminology of Chapter 11) the same is true of G. A symmetry fixing two points of a line must of course fix also the third point. Let H be the stabilizer in G of the points of l_1 and l_2. We need to show that H has order 2 (the order of G is then $15\times 6\times 4\times |H| = 6!$). We can identify 4 with 35 and 5 with 46. The group $H\cap S_6$ does of course have order 2. It is generated by the permutation (12). On the other hand each element of H must permute the elements of a given box (the second or third), which are collinear with 35, among themselves, the same with 46 replacing 35. This shows that the following sets of points are respected by the action of H :

$$\{16,26\},\ \{15,25\},\ \{14,24\},\{13,23\}$$

Also, the transposition (12) maps $16 \leftrightarrow 26$. Consider the stabilizer L of point 16 in H. We have to show that L is the identity group. This is clear now: each element $g \in L$ fixes 16 and 26, and therefore also 15 and 25 (being third points of lines two of whose points are fixed already). All points of the second box are fixed by g. Each point in the third box is the third point of a line whose remaining two points (from boxes one and two) are known to be fixed already. It follows that g is the identity.

13.9 Theorem. *The full automorphism group of the S_6-GQ is S_6. The groups S_6 and $Sp_4(2)$ are isomorphic.*

The unitary case

This is just as easy as the symplectic case..

13.10 Definition. *The* **general unitary group** $GU(n,q)$ *consists of the elements $A \in GL(2m,q)$, which satisfy $(Ax,Ay) = (x,y)$ for all $x,y \in V$.*

We have seen in Chapter 9 that we can find an orthonormal basis. The order of $GU(n,q^2)$ is the number of ordered orthonormal bases. In Proposition 9.1 we determined the numbers $u_n(c)$ of vectors satisfying $(v,v) = c$. In particular $u_n(1) = q^{n-1}(q^n - (-1)^n)$. We have
$|GU(n,q^2)| = u_n(1)u_{n-1}(1)\dots u_1(1)$. As the sum of all natural numbers up to $n-1$ equals $\binom{n}{2}$ the following formula is obtained:

13.11 Theorem.

$$|GU(n,q^2)| = q^{n(n-1)/2}(q+1)(q^2-1)\dots(q^n-(-1)^n).$$

The groups respecting non-degenerate quadratic forms are the **orthogonal groups.** We can derive formulas for their orders in much the same way as we did in the symplectic and in the unitary case. These groups, together with the alternating groups, form the family of finite simple groups which were known as the **classical groups** before the link to Lie algebras was discovered. We will briefly come back to this in the last chapter.

Chapter 14

Generators and Spreads

Totally singular (totally isotropic) subspaces of maximal dimension (= Witt index) are also known as **generators** in the geometric literature. It is an easy counting exercise to determine the number of generators in our geometries (although it is easy to go wrong).

14.1 Theorem. *The number of totally isotropic subspaces in the symplectic non-degenerate geometry on $V(2m,q)$ is*

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

Proof. As usual we count bases of totally isotropic subspaces $\langle v_1,\dots,v_m\rangle$. The number of choices for v_1 is $q^{2m}-1$. Once v_1 is fixed v_2 must be chosen from $v_1^\perp$ but outside $\langle v_1\rangle$. Counting in that way we obtain $(q^{2m}-1)(q^{2m-1}-q)(q^{2m-2}-q^2)\dots(q^{m+1}-q^{m-1})$ as the number of ordered bases of totally isotropic subspaces. Each such space has $(q^m-1)(q^m-q)\dots(q^m-q^{m-1})=|GL(m,q)|$ ordered bases. □

As an application we can count the self-dual binary codes.

14.2 Theorem. *The number of self-dual subspaces (self-dual binary codes) of $\mathbb{F}_2^{2m}$ with respect to the ordinary dot product equals the number of generators in the symplectic geometry on $V(2(m-1),2)$, see Theorem 14.1.*

Proof. Clearly self-dual codes can exist only when the length is even. Let $V=\mathbb{F}_2^{2m}$ with the dot product and $C=C^\perp\subset V$. Then $dim(C)=m$. The dot product is a non-degenerate symmetric bilinear form on V (it has the unit matrix as a Gram matrix). A vector is isotropic with respect to the dot product if and only if it has even weight. The even weight vectors form a hyperplane $V_0 \subset V$, known as the all-even code or as the augmentation ideal. In fact $V_0 = \mathbf{1}^\perp$, and $\mathbf{1} \subset V_0$. We have $\mathbf{1} \subset C \subset V_0$. Consider the factor space $W = V_0/\langle \mathbf{1} \rangle$. By definition V_0 is symplectic. As $\mathbf{1}$ is the radical of V_0 we have that W is a non-degenerate symplectic space of dimension $2(m-1)$. The self-dual codes in V are in bijection with the generators of W. □

14.3 Theorem. *The number of totally singular subspaces in the hyperbolic geometry on $V(2m, q)$ is*

$$2(q+1)(q^2+1)\dots(q^{m-1}+1).$$

Proof. The Witt index is m. At first count ordered bases: There are $h_m(0)-1$ choices for v_1. For fixed v_1 there are $qh_{m-1}(0)-q$ choices for v_2 and so forth. Here we use the notation from Chapter 6. We observed already that in characteristic 2 the numbers are the same. When a basis $v_1, \dots, v_i$ for a totally singular subspace has been chosen we have that $\langle v_1, \dots, v_i \rangle^\perp$ is the orthogonal sum of $\langle v_1, \dots, v_i \rangle$ and an $2(m-i)$-dimensional hyperbolic space. The number of choices for v_{i+1} is then $q^{2m-i}h_{m-i}(0) - q^{2m-i}$. We obtain $q^{m(m-1)/2}(h_m(0)-1)(h_{m-1}(0)-1)\dots(h_1(0)-1)$ as the number of ordered bases of totally singular subspaces. In order to count the subspaces we have to divide by $|GL(m, q)|$ again. This kills the factor $q^{m(m-1)/2}$. Also, in Chapter 6 we found $h_m(0)-1 = (q^m-1)(q^{m-1}+1)$. The first factor cancels against the denominator. This yields the formula. □

The factor 2 in the formula of Theorem 14.3 is $(h_1(0)-1)/(q-1) = q^0+1$. Proposition 10.2 is the special case $m = 2$ of Theorem 14.3. We observed the curious factor 2 there already: there are $2(q+1)$ totally isotropic lines in the 4-dimensional hyperbolic geometry. These come in two parallel classes. It may be expected that there is a similar phenomenon in all dimensions. This is indeed the case. The generators in hyperbolic space come in two equivalence classes. Two generators belong to the same equivalence if they intersect in a space whose dimension has the same parity as the Witt index itself. The proof is not that easy. The algebraic proof involves the spinorial norm and the Clifford algebra. A geometric proof is in [8]. We will not do this here.

It is a natural question to ask if the isotropic (singular) points of our geometries can be partitioned into generators. Naturally this is conceivable only if the number of points on a generator divides the number of singular points. Looking back at the formulas for the numbers of points on our quadrics we see that this condition is always satisfied.

14.4 Definition. *A* **spread** *of one of the classical geometries is a family of generators (maximal totally isotropic respectively totally singular subspaces), which partition the set of isotropic (singular) points.*

We will speak of symplectic, hyperbolic, ... spreads. The construction is easiest in the symplectic case.

14.5 Theorem. *Each non-degenerate symplectic geometry has a spread.*

Proof. As all points of $PG(2m-1,q)$ are isotropic, a spread must consist of $(q^{2m-1}-1)/(q^m-1) = q^m+1$ generators. Use the extension field $L=\mathbb{F}_{q^m}$ and the trace $tr: L \rightarrow \mathbb{F}_q$. Let $V=V(2,q^m)$ be a 2-dimensional vector space over L with symplectic bilinear form $(,)$.

We can view V as a $2m$-dimensional vector space over $\mathbb{F}_q$. The idea is to equip it with a symplectic bilinear form using $(,)$ and the trace. We define

$$B(x,y) = tr(x,y)$$

As $(,)$ is bilinear and symplectic (over L), also $B(,)$ is bilinear and symplectic over $\mathbb{F}_q$ (for example $B(x,x) = tr(x,x) = tr(0) = 0$). It is also non-degenerate, for assume $x \neq 0$ is in the radical. As there is some y such that $(x,y) \neq 0$ ($(,)$ is non-degenerate) and $(,)$ is L-linear it follows that (x,y) takes on all values in L when y varies. As tr is not the 0-mapping we obtain a contradiction.

This means that we can work with this model of symplectic geometry whenever this is advantageous. In our situation it is. $V(2,q^m)$ trivially is partitioned into its q^m+1 points. Each point is a 1-dimensional L-vector space, hence an m-dimensional $\mathbb{F}_q$-vector space. It is also totally isotropic with respect to B. It follows that these points form a spread of the $2m$-dimensional vector space. □

Symplectic $V(4,2)$

As an example consider $V(4,2)$. The extension field is $L=\mathbb{F}_4=\{0,1,\omega,\overline{\omega}\}$ which the reader certainly remembers. The trace is $tr: \mathbb{F}_4 \longrightarrow \mathbb{F}_2$ and we recall that the elements of $\mathbb{F}_2$ have trace 0, whereas $\omega, \overline{\omega}$ have trace 1. The first step is to consider a 2-dimensional space $V = V(2,4)$ over L and equip it with the symplectic form $(,)$. Let $e_1 = (1,0),\ e_2 = (0,1)$ be a symplectic basis. Now we see V as a 4-dimensional space $V(4,2)$ over $\mathbb{F}_2$. A basis is

$$v_1 = e_1,\ v_2 = \omega e_1,\ v_3 = e_2,\ v_4 = \omega e_2.$$

The symplectic form B is defined as $B(x,y) = tr(x,y)$. The Gram matrix with respect to basis v_1, v_2, v_3, v_4 is $A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$. For example, $(e_1, e_1) = (e_1, \omega e_1) = 0$, which shows that in the top left corner of A we have a 0-matrix. The same is true of the Southeast corner, and in general of any 1-dimensional L-space. As another example, the Northeast entry of A is $tr(v_1, v_4) = tr(e_1, \omega e_2) = tr(\omega) = 1$. Our spread consists of the $4+1=5$ 1-dimensional $\mathbb{F}_4$-subspaces of $V(2,4)$. These are the points of the projective line $PG(1,4)$. As vectors generating them over $\mathbb{F}_4$ we can choose

$$e_2,\ e_1,\ e_1+e_2,\ e_1+\omega e_2, e_1+\overline{\omega}e_2.$$

The first of those contains the nonzero vectors $e_2 = v_3, \omega e_2 = v_4, v_3+v_4$. The last of those contains $e_1+\overline{\omega}e_2 = v_1+v_3+v_4, \omega e_1+e_2 = v_2+v_3$ and the sum $v_1+v_2+v_4$. It is left to the reader to write out a complete list.

We know that the symplectic 4-dimensional geometry and its generators simply form the points and lines of the symplectic GQ $W(q)$. Theorem 14.5 states that we can find q^2+1 lines which partition the $(q+1)(q^2+1)$ points of $W(q)$. What does that mean in the binary case, the by now familiar S_6-GQ? Start from the underlying 6-set S. Each point of the GQ is a pair of elements from S. A graph-theoretic way of expressing this is: the points of the S_6-GQ are the edges of the complete graph on 6 vertices. The lines are then the 1-factors of this complete graph (three edges which partition the vertices) and a spread is a family of 1-factors partitioning the edges of the complete graph. Such a structure is known as 1-**factorization** in graph theory. The special case $q=2$ of Theorem 14.5 states that the complete graph on 6 vertices has a 1-factorization, and we constructed it.

The hyperbolic case

This is harder than the symplectic case, but it is particularly interesting, as we will see later.

14.6 Theorem. *Let q be a power of 2. Each non-degenerate hyperbolic geometry on $V(4m,q)$ possesses a spread.*

Proof. Let Q be the corresponding hyperbolic quadric. A spread would have $q^{2m-1}+1$ elements (see Theorem 6.13), just as many as a symplectic spread on $V(4m-2,q)$. This makes us suspect that we could use the existence of symplectic spreads to construct our hyperbolic spreads.

Let x such that $Q(x)\neq 0$. Consider the hyperplane $x^\perp$ and the space $W = x^\perp/\langle x\rangle$ of dimension $4m-2$. It carries a non-degenerate symplectic structure (as $\langle x\rangle$ is of course the radical of the restriction of the symplectic form to $x^\perp$), but it does not have an orthogonal structure as different elements of the same coset $y+\langle x\rangle$ have different values under Q. It is therefore natural to start from a symplectic spread on W. Let $\overline{X}$ be an element of this symplectic spread. Then $dim(\overline{X}) = 2m-1$. Its preimage X in $x^\perp$ has dimension $2m$ and is totally isotropic but is of course not totally singular (as $Q(x)\neq 0$). The restriction of Q to X is semi-linear. The kernel is an $(2m-1)$-dimensional totally singular subspace Y. We have $Y^\perp = Y \perp H$, where H is a quadratic hyperbolic plane. It follows that Y is contained in precisely two generators.

Now we need some of the structure mentioned earlier (we were too lazy to prove that). The generators in hyperbolic space fall into two equivalence classes, say type 1 and type 2. In our case (dimension $4m$) two generators have same type if and only if their intersection has even dimension. In particular each $(2m-1)$-dimensional totally singular subspace is contained in one generator of each type.

We define a mapping ϕ from generators on symplectic space W to generators of our hyperbolic geometry by letting $\phi(\overline{X}) = Z$ be the generator of type 1 containing Y. The claim is that ϕ maps symplectic spreads to hyperbolic spreads. As observed in the beginning the numbers are right. It suffices to show that if $\overline{X_1}\cap\overline{X_2} = \{0\}$ (we see these as vector spaces), then $Z_1\cap Z_2=\{0\}$. We have $Z_i\cap x^\perp = Y_i$ and $Y_1\cap Y_2=\{0\}$. As Z_1, Z_2 have the same type their intersection dimension is even. As Y_i is a hyperplane in Z_i this forces $Z_1\cap Z_2=\{0\}$. $\square$

Chapter 15

Reed-Muller codes and Kerdock codes

There are close links between finite geometry and coding theory. We saw an important one in Chapter 3. In the present chapter we want to study another way how to obtain binary codes from finite geometries.

15.1 Definition. *Let $V = V(m, 2)$ the space of binary m-tuples. The set $\mathcal{F}_m$ of all mappings $f : V \to \mathbb{F}_2$ is a binary vector space of dimension 2^m. Here addition is defined by $(f + g)(x) = f(x) + g(x)$. A basis consists of the characteristic functions χ_a for $a \in V$ defined by $\chi_a(a) = 1, \chi_a(b) = 0$ for $b \neq a$.*

Each polynomial in m variables $x_1, x_2, \dots, x_m$ with coefficients in $\mathbb{F}_2$ describes an element of $\mathcal{F}_m$. This presupposes the choice of a fixed basis for V, so we see V as $\mathbb{F}_2^m$. In each monomial exponents higher than 1 of x_i are not needed as x_i and x_i^2 describe the same mapping on $\mathbb{F}_2$. So we can restrict to monomials, which are products of different x_i. Such a monomial is described by a subset of the index set $\{1, 2, \dots, m\}$. For example, the subset $\{1, 2, 4\}$ describes the monomial $x_1x_2x_4$. As simple as that. How many different such polynomials are there? The set $\{1, 2, \dots, m\}$ has 2^m subsets, so this is the number of different monomials of that type. In order to show that these monomials form a basis of $\mathcal{F}_m$ it suffices to show that the characteristic functions χ_a can be represented by polynomials. This is easy to see. In fact,

$$\chi_a = \prod_{i=1}^{m}(1 + x_i + a_i).$$

We have proved the following:

15.2 Theorem. *Let $V = \mathbb{F}_2^m$. Each mapping $f : V \longrightarrow \mathbb{F}_2$ can be written in a unique way as a polynomial in m variables $x_1, x_2, \ldots, x_m$ with coefficients in $\mathbb{F}_2$ such that in each monomial each variable occurs with exponent ≤ 1. This representation of f is known as the* **algebraic normal form** *(ANF).*

For example, the ANF of the mapping $00 \mapsto 0,\ 10 \mapsto 0,\ 01 \mapsto 1,\ 11 \mapsto 1$ simply is x_2, and the ANF of $00 \mapsto 1,\ 10 \mapsto 0,\ 01 \mapsto 1,\ 11 \mapsto 0$ is $x_1 + 1$. The algebraic normal form plays an important role in information transmission and cryptology.

If we want to study the behaviour of functions $f \in \mathcal{F}_m$ it is natural to construct a list of all function values. That is we order the elements of V in some way and consider the 2^m-tuple $(f(a))_{a \in V}$. A way to think of this is that f defines a word in a binary code of length 2^m.

15.3 Definition. *Let $V = \mathbb{F}_2^m$ and $f : V \to \mathbb{F}_2$. Order the elements of V in some way: $V = \{a_1, \ldots, a_n\}$, where $n = 2^m$. Then*

$$L_f = (f(a_1), f(a_2), \ldots, f(a_n)).$$

Observe that L_f is a binary 2^m-tuple.

Here is an example for $m = 3$. Let $f \in \mathcal{F}_3$ be as follows:

000	$\longrightarrow$	1
001	$\longrightarrow$	0
010	$\longrightarrow$	1
011	$\longrightarrow$	0
100	$\longrightarrow$	1
101	$\longrightarrow$	1
110	$\longrightarrow$	1
111	$\longrightarrow$	0

If we order the tuples from left to right in the same way that they appear vertically from 000 to 111 in the table, then clearly

$$L_f = (1, 0, 1, 0, 1, 1, 1, 0).$$

What is the ANF of f? Evaluate at 000. If we do this for the ANF, the result is the constant term. On the other hand $f(000) = 1$. It follows that the constant term is 1. Now write

$$f = 1 + ax_1 + bx_2 + cx_3 + dx_1x_2 + ex_1x_3 + fx_2x_3 + gx_1x_2x_3.$$

We can determine the coefficients $a, b, \dots, g$ one by one, starting from the terms of low degree. Evaluating at 100 yields $1 = f(100) = 1 + a$, so $a = 0$. In the same way, evaluating at triples of weight 1, we obtain $b = 0,\ c = 1:$

$$f = 1 + x_3 + dx_1x_2 + ex_1x_3 + fx_2x_3 + gx_1x_2x_3.$$

Now evaluate at 110. This shows $1 = f(110) = 1 + d$, so $d = 0$. Using all weight 2 triples we obtain

$$f = 1 + x_3 + x_1x_3 + gx_1x_2x_3.$$

Finally $0 = f(111) = 1 + 1 + 1 + g$, so $g = 1$. The ANF of f is

$$f = 1 + x_3 + x_1x_3 + x_1x_2x_3.$$

Each $f \in \mathcal{F}_m$ has an algebraic normal form. We consider the degree of this ANF (in the example above we had degree 3). For low degrees this leads us to familiar objects. If f is homogeneous of degree 2, then f is a binary quadratic form. This motivates the definition of a famous class of codes:

15.4 Definition. *The* **Reed-Muller code** $\mathcal{R}(r, m)$ *has as codewords the* L_f, *where* $f \in \mathcal{F}_m$ *varies over the mappings whose algebraic normal form has degree* $\leq r$.

As addition of polynomials does not increase the degree we see that $\mathcal{R}(r, m)$ is a linear code. Its length is by definition $n = 2^m$. The number of monomials of degree i is $\binom{m}{i}$. This shows that $\mathcal{R}(r, m)$ has dimension $\sum_{i=0}^{r} \binom{m}{i}$. Clearly $\mathcal{R}(0, m)$ consists only of the 0-tuple and of the 1-tuple $\mathbf{1}$. It has dimension 1 and minimum distance $n = 2^m$. Let us concentrate on small r. We have $dim(\mathcal{R}(1, m)) = m + 1$, and this code consists of L_f, where f is a linear mapping $: V \longrightarrow \mathbb{F}_2$ (these are the 2^m linear combinations of $x_1, x_2, \dots, x_m$) and the complementary tuples. Each non-trivial linear function f has 2^{m-1} zeroes. This shows $wt(L_f) = 2^m - 2^{m-1} = 2^{m-1}$. As $wt(L_f + \mathbf{1}) = n - wt(L_f)$, we see that each element of $\mathcal{R}(1, m) \setminus \mathcal{R}(0, m)$ has weight 2^{m-1}. This is the beginning of an inductive argument, which can be used to determine the minimum distance.

15.5 Theorem. *$\mathcal{R}(r,m)$ is a binary linear code of length 2^m, dimension $\sum_{i=0}^{r}\binom{m}{i}$ and minimum distance 2^{m-r}.*

Proof. Only the minimum distance is still in doubt. Let $f \in \mathcal{F}_m$ with ANF of degree $\leq r$. We can assume by induction that the degree is precisely r. Choose a variable which occurs in the ANF, without restriction x_1. We can write $f(x_1, \dots, x_m) = x_1 g(x_2, \dots, x_m) + h(x_2, \dots, x_m)$, where g is not identically zero. We can view g, h as elements of $\mathcal{F}_{m-1}$, mappings $W = \mathbb{F}_2^{m-1} \longrightarrow \mathbb{F}_2$. The elements of V are $(0, b)$ and $(1, b)$, where $b \in W$. We have $f(0, b) = h(b)$ and $f(1, b) = g(b) + h(b)$. Observe that g has degree $\leq r - 1$. By induction there are at least $2^{(m-1)-(r-1)} = 2^{m-r}$ points $b \in W$ such that $g(b) = 1$. For each such b we have that either $f(0, b) = 1$ or $f(1, b) = 1$. □

Consider $\mathcal{R}(2,m)$. Let $f \in \mathcal{F}_m$ of degree 2. If the constant term is 0, then f is a sum of terms of the form $x_i x_j$ for $i \neq j$ and of linear terms x_i. We can write x_i^2 instead of x_i and obtain the same mapping f. This shows that f simply is a binary quadratic form in dimension m. We studied quadratic forms in Chapter 8. The weight $wt(L_f)$ equals the number of vectors v where $f(v) = 1$. These representation numbers have been determined in earlier chapters.

As we know, the underlying symplectic form is determined by the terms $x_i x_j$ for $i \neq j$. Different choices of quadratic terms x_i^2 lead to different quadratic forms belonging to the same symplectic bilinear form. Consider the case when the symplectic form is non-degenerate. Then the quadratic form will automatically be non-degenerate, for every choice of the diagonal terms. This can happen only when $m = 2l$ is even. We know the representation numbers of these quadratic forms. They are

$$e_l(0) = 2^{2l-1} - 2^{l-1}, e_l(1) = 2^{2l-1} + 2^{l-1}$$

$$h_l(0) = 2^{2l-1} + 2^{l-1}, h_l(1) = 2^{2l-1} - 2^{l-1}.$$

Whenever f is elliptic we have $wt(L_f)) = e_l(0), wt(L_f + \mathbf{1} = e_l(1)$. If f is hyperbolic, then $wt(L_f)) = h_l(0), wt(L_f + \mathbf{1} = h_l(1)$. We see that only two different weights occur. Fixing the symplectic form means fixing a coset of $\mathcal{R}(1, 2l)$ in $\mathcal{R}(2, 2l)$. Whenever that coset is described by a non-degenerate symplectic form, all code-words for f in that coset will have weights $2^{2l-1} \pm 2^{l-1}$.

In order to construct good linear codes we would have to find a large family of symplectic forms all of whose nonzero linear combinations are non-degenerate. It turns out that such families are hard come by. It is in fact profitable to abandon the aim of constructing linear codes. Let us be content with nonlinear codes. The following general definition of a binary error-correcting code generalizes the definition of a linear code as given in Chapter 3.

15.6 Definition. *An $(n, M, d)_2$-code is a family of M binary tuples of length n whose pairwise Hamming distance is at least d.*

The reason why we considered linear codes exclusively so far is that linearity helps a great deal in constructing and applying codes. The geometrical description given in Chapter 3 is available only for linear codes. Our considerations of binary quadratic forms and Reed-Muller codes finally lead to a situation where highly structured nonlinear codes come into play. Observe that each (linear) $[n, k, d]_2$-code also is an $(n, 2^k, d)_2$-code.

We use the following strategy to construct good binary codes: find a family of symplectic bilinear forms on $V(2l, 2)$ such that for any two different of these forms their difference is non-degenerate. The union of the corresponding cosets of $\mathcal{R}(1, 2l)$ in $\mathcal{R}(2, 2l)$ will then be a code of minimum distance $2^{2l-1} - 2^{l-1}$. How many cosets can we expect in such a set? As we know each symplectic form on $V(2l, 2)$ is determined by a symmetric $(2l, 2l)$-matrix with zeroes on the diagonal. How many of those matrices can we find such that the difference between any two is non-degenerate? Certainly the first rows of these matrices must be different (if they are equal, the difference starts with the 0-row and therefore is degenerate). As the first row starts with a 0 there are only 2^{2l-1} possible first rows. So this is a bound on the size of such a family of matrices.

15.7 Definition. *A* **Kerdock set** *of $(2l, 2l)$-matrices is a family of 2^{2l-1} binary symmetric matrices with zeroes on the main diagonal (symplectic matrices), such that the difference (=sum) of any two different of those matrices is non-degenerate (equivalently: has nonzero determinant).*

15.8 Theorem. *If a Kerdock set of binary symplectic forms exist, we can construct a binary code (the corresponding* **Kerdock code***) of length 2^{2l} with 2^{4l} codewords and minimum distance $2^{2l-1} - 2^{l-1}$.*

By construction the Kerdock code is a union of cosets of $\mathcal{R}(1, 2l)$ in $\mathcal{R}(2, 2l)$.

Kerdock sets can be constructed using spreads. This provides an application of the construction of spreads and also gives us the opportunity to review the relevant facts. Start from a hyperbolic space on $V(4l,2)$. Why we work in double the dimension we are interested in? Just wait for a short while.

Choose a standard basis $v_1,\dots,v_{2l},w_1,\dots,w_{2l}$ such that $Q(\sum_i x_iv_i + \sum_i y_iw_i) = \sum_{i=1}^{2l} x_iy_i$ (we called $\langle v_i,w_i\rangle$ a quadratic hyperbolic plane in Chapter 8). It was proved in Chapter 14 that the hyperbolic quadric $Q+(4l-1,2)$ has a spread. We can choose notation such that the generators $E = \langle v_1,\dots,v_{2l}\rangle$ and $F = \langle w_1,\dots,w_{2l}\rangle$ belong to this spread. Let

$$\Sigma = \{E\} \cup \{F_1,\dots,F_{2^{2l-1}}\}$$

be a spread of $Q^+(4l-1,2)$, where $F_1 = F$.

We did not study the orthogonal groups (symmetry groups of quadratic forms) in detail. The definition is clear from the discussion in Chapter 13. The orthogonal group $O^+(4l,2)$ consists of the binary $(4l,4l)$-matrices mapping our standard basis $v_1,\dots,v_{2l},w_1,\dots,w_{2l}$ to a standard basis (and therefore leaving the quadratic form invariant). Restrict attention to the matrices in $O^+(4l,2)$ which act trivially on E, that is which map $v_i \mapsto v_i$ for all i. These matrices form a subgroup H of $O^+(4l,2)$. The scalar products show that elements of H have the form $A(M) = \begin{pmatrix} I & 0 \\ M & I \end{pmatrix}$. We have $(x|y)A(M) = (x+yM|y)$. It follows that $A(M) \in O^+(4l,2)$ if and only if $\sum_i x_iy_i = \sum_i (x_i + (yM)_i)y_i$, equivalently $yMy^T = 0$ for every y. This means that the bilinear form mapping the pair x,y to xMy^T is symplectic, in other words M is a binary symplectic matrix: symmetric with zeroes on the diagonal. These symplectic matrices form an additive group (with the 0-matrix as neutral element), clearly of order $2^{l(2l-1)}$. Let P be this group of symplectic binary $(2l,2l)$-matrices. The mapping $M \mapsto A(M)$ is an isomorphism : $P \to H \subset O^+(4l,2)$. Observe that for every generator L with zero intersection with E there is precisely one matrix M such that $A(M) : E \to L$.

15.9 Definition. *$K(\Sigma)$ consists of the symplectic $(2l,2l)$-matrices M such that $A(M)$ maps F_1 to some $F_j \in \Sigma$.*

As any two generators in the spread Σ have 0 intersection, we have that for every $j = 1,\dots,2^{2l-1}$ there is precisely one matrix M_j such that $A(M_j) : F \to F_j$. In particular $K(\Sigma)$ consists of 2^{2l-1} symplectic $(2l,2l)$-matrices.

We claim that $K(\Sigma)$ is a Kerdock set. Assume $M_j - M_k$ is singular for some $j \neq k$. There is some vector $y \neq 0$ such that $yM_j = yM_k$. Consider the nonzero vector $(0|y) \in F$. We have

$$(0|y)A(M_j) = (yM_j|y) = (yM_k|y) = (0|y)A(M_k) \in F_j \cap F_k,$$

a contradiction as F_j, F_k are different generators from a spread. We have seen that Kerdock sets exist. This also establishes the existence of Kerdock codes.

The Nordstrom-Robinson code

The Kerdock code of length 16 is known as the Nordstrom-Robinson code **NR**. We are in case $l = 2$. The number of codewords is $2^8 = 256$, the minimum distance is 6. Observe that there is no reason why **NR** should be linear. In fact it is non-linear. What is more, it can be shown that linear codes with parameters $[16, 8, 6]_2$ cannot exist.

The construction starts with a (trivial) spread in the symplectic $V(2, 8)$. Let a, b be a symplectic basis of this space. The spread simply consists of the points of the projective line. Let $A(s) = \langle sa + b\rangle$ for $s \in \mathbb{F}_8, A(\infty) = \langle a\rangle$. A $V(2, 8)$ is also a $V(6, 2)$. We obtain a spread in the symplectic $V(6, 2)$ (Theorem 14.5). Its elements are $A(s)$ and $A(\infty)$ as 3-dimensional binary spaces. We use the representation $\mathbb{F}_8 = \mathbb{F}_2(\epsilon)$, where $\epsilon^3 + \epsilon^2 + 1 = 0$ and $1, \epsilon, \epsilon^2$ as basis over $\mathbb{F}_2$. A symplectic basis of $V(2, 8)$ over $\mathbb{F}_2$ is

$$v_1 = a,\ v_2 = \epsilon a,\ v_3 = \epsilon^2 a,\ w_1 = \epsilon^4 b,\ w_2 = \epsilon^3 b,\ w_3 = \epsilon^5 b.$$

Observe that $1, \epsilon, \epsilon^2, \epsilon^4$ have trace $= 1$, the remaining elements of $\mathbb{F}_8$ have trace $= 0$. For example, $(v_1, w_1) = tr(\epsilon^4) = 1, (v_1, w_2) = tr(\epsilon^3) = 0$. In terms of the symplectic basis we have

$A(\infty)$:	$\langle v_1\rangle$		
$A(0)$	$\langle w_1\rangle$	$A(\epsilon^3)$:	$\langle v_1 + v_3 + w_1 + w_2 + w_3\rangle$
$A(1)$:	$\langle v_1 + w_1 + w_2 + w_3\rangle$	$A(\epsilon^4)$:	$\langle v_1 + v_2 + v_3 + w_1 + w_2 + w_3\rangle$
$A(\epsilon)$:	$\langle v_2 + w_1 + w_2 + w_3\rangle$	$A(\epsilon^5)$:	$\langle v_1 + v_2 + w_1 + w_2 + w_3\rangle$
$A(\epsilon^2)$:	$\langle v_3 + w_1 + w_2 + w_3\rangle$	$A(\epsilon^6)$:	$\langle v_2 + v_3 + w_1 + w_2 + w_3\rangle$

Here $\langle\rangle$ denotes the 1-dimensional space generated over $\mathbb{F}_8$. Next we apply the procedure in the proof of Theorem 14.6 to obtain a spread in the hyperbolic geometry on $V(8, 2)$. We use a standard basis $v_0, v_1, v_2, v_3, w_0, w_1, w_2, w_3$

and $Q(x_0, x_1, x_2, x_3, y_0, y_1, y_2, y_3) = \sum_{i=0}^{3} x_i y_i$. As point outside the quadric we choose $x = v_0 + w_0 = 10001000$. Identify $x^\perp/\langle x\rangle = \langle x, v_1, v_2, v_3, w_1, w_2, w_3\rangle/\langle x\rangle$ with our symplectic $V(6, 2)$.

Start from the $A(\infty) = \overline{X(\infty)} = \langle \overline{v_1}, \overline{v_2}, \overline{v_3}\rangle$, where $\langle\rangle$ now refers to the space generated over $\mathbb{F}_2$. Its preimage is the 4-dimensional space $X(\infty)$, the intersection with $x^\perp$ is $Y(\infty) = \langle v_1, v_2, v_3\rangle$. One of the two totally singular 4-dimensional spaces containing $Y(\infty)$ is $E = \langle v_0, v_1, v_2, v_3\rangle$, which we choose as the first element in our hyperbolic spread.

Next $X(0) = \langle v_0+w_0, w_1, w_2, w_3\rangle, Y(0) = \langle w_1, w_2, w_3\rangle$. The 4-dimensional totally singular space, which contains $Y(0)$ and intersects E trivially is $F(0) = \langle w_0, w_1, w_2, w_3\rangle$. The first matrix $M(0)$ of our Kerdock set is the 0-matrix

$A(1)$ yields $X(1) = \langle v_0+w_0, v_1+w_1+w_2+w_3, v_2+w_1+w_2, v_3+w_1+w_3\rangle$, $Y(1) = X(1)\cap x^\perp = \langle v_0+v_1+w_0+w_1+w_2+w_3, v_0+v_2+w_0+w_1+w_2, v_0+v_3+w_0+w_1+w_3\rangle$. We know from the theory that $Y(1)$ is contained in two totally singular 4-dimensional subspaces. One of them will intersect E (and then also $F(0)$) in dimension 0. This is our generator $F(1)$, the next element in the spread. $v_0 + w_1 + w_2 + w_3$ can be chosen as fourth generator:

$$F(1) = \langle 11001111,\ 10101110,\ 10011101,\ 10000111\rangle$$

After reorganization we obtain

$$F(1) = \langle 1000|0111,\ 0100|1000,\ 0010|1001,\ 0001|1010\rangle$$

The second element of the Kerdock set is $M(1) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$.

The same procedure yields $X(\epsilon) = \langle v_0 + w_0, v_2 + w_1 + w_2 + w_3, v_3 + w_1 + w_2, v_1 + v_3 + w_1 + w_3\rangle$, $Y(\epsilon) = \langle v_0 + v_2 + w_0 + w_1 + w_2 + w_3, v_3 + w_1 + w_2, v_1 + v_3 + w_1 + w_3\rangle$. The unique vector with E-coordinate v_0, which extends $Y(\epsilon)$ to a generator $F(\epsilon)$ meeting E in the 0-space is $v_0 + w_2$. After change of basis we obtain

$$F(\epsilon) = \langle 1000|0010,\ 0100|0011,\ 0010|1101,\ 0001|0110\rangle \text{ and}$$

$$M(\epsilon) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

We understand the procedure now. $X(\epsilon^2) = \langle v_0 + w_0, v_3 + w_1 + w_2 + w_3, v_1 + v_3 + w_1 + w_2, v_1 + v_2 + v_3 + w_1 + w_3\rangle$, $Y(\epsilon^2) = \langle v_0 + v_3 + w_0 + w_1 + w_2 + w_3, v_0 + v_1 + v_3 + w_0 + w_1 + w_2, v_1 + v_2 + v_3 + w_1 + w_3\rangle$,

$$F(\epsilon^2) = \langle 1000|0011,\ 0100|0001,\ 0010|1000,\ 0001|1100\rangle$$

$$M(\epsilon^2) = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

$X(\epsilon^3) = \langle v_0+w_0, v_1+v_3+w_1+w_2+w_3, v_1+v_2+v_3+w_1+w_2, v_1+v_2+w_1+w_3\rangle$, $Y(\epsilon^3) = \langle v_1+v_3+w_1+w_2+w_3, v_1+v_2+v_3+w_1+w_2, v_0+v_1+v_2+w_0+w_1+w_3\rangle$,

$$F(\epsilon^3) = \langle 1000|0101,\ 0100|1001,\ 0010|0001,\ 0001|1110\rangle$$

$$M(\epsilon^3) = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

$X(\epsilon^4) = \langle v_0+w_0, v_1+v_2+v_3+w_1+w_2+w_3, v_1+v_2+w_1+w_2, v_2+v_3+w_1+w_3\rangle$, $Y(\epsilon^4) = \langle v_0 + v_1 + v_2 + v_3 + w_0 + w_1 + w_2 + w_3, v_1 + v_2 + w_1 + w_2, v_0 + v_2 + v_3 + w_0 + w_1 + w_3\rangle$,

$$F(\epsilon^4) = \langle 1000|0001,\ 0100|0010,\ 0010|0100,\ 0001|1000\rangle$$

$$M(\epsilon^4) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

$X(\epsilon^5) = \langle v_0 + w_0, v_1 + v_2 + w_1 + w_2 + w_3, v_2 + v_3 + w_1 + w_2, v_1 + w_1 + w_3\rangle$, $Y(\epsilon^5) = \langle v_1+v_2+w_1+w_2+w_3, v_0+v_2+v_3+w_0+w_1+w_2, v_0+v_1+w_0+w_1+w_3\rangle$,

$$F(\epsilon^5) = \langle 1000|0110,\ 0100|1011,\ 0010|1100,\ 0001|0100\rangle$$

$$M(\epsilon^5) = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

$X(\epsilon^6) = \langle v_0 + w_0, v_2 + v_3 + w_1 + w_2 + w_3, v_1 + w_1 + w_2, v_2 + w_1 + w_3\rangle,$
$Y(\epsilon^6) = \langle v_2 + v_3 + w_1 + w_2 + w_3, v_0 + v_1 + w_0 + w_1 + w_2, v_2 + w_1 + w_3\rangle,$

$$F(\epsilon^6) = \langle 1000|0100,\ 0100|1010,\ 0010|0101,\ 0001|0010\rangle$$

$$M(\epsilon^6) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

15.10 Theorem. *The symplectic matrices $M(\alpha), \alpha \in \mathbb{F}_8$, form a Kerdock set. The Nordstrom-Robinson code* **NR** *consists of the cosets $R(1,4) + L_f$, where $f = f_\alpha$ varies over the quadratic forms described by the matrices $M(\alpha)$.*

Here $f_0 = 0$, so $R(1,4) + L_0 = R(1,4)$. We have $f_1 = x_0x_1 + x_0x_2 + x_0x_3 + x_2x_3,\ f_\epsilon = x_0x_2 + x_1x_2 + x_1x_3 + x_2x_3, \dots$ We could have saved some energy by using symmetries. For example, the 2-dimensional symplectic form is unchanged if we multiply by ϵ in the first coordinate and by $1/\epsilon = \epsilon^6$ in the second coordinate. This yields an automorphism σ of order 7 which fixes v_0, w_0 and maps

$$\sigma = (v_1, v_2, v_3, v_1 + v_3, v_1 + v_2 + v_3, v_1 + v_2, v_2 + v_3)$$

$$(w_1, w_2, w_1 + w_3, w_1 + w_2, w_1 + w_2 + w_3, w_2 + w_3, w_3).$$

This automorphism maps $M(\alpha) \mapsto M(\epsilon^2\alpha)$. It suffices therefore to know $M(1)$. The remaining nonzero matrices of the Kerdock set can be obtained by applying σ. We leave the details to the reader.

Problems

1. Describe the automorphism σ of the Nordstrom-Robinson code in matrix form.

Chapter 16

Projective planes

We defined a projective plane of order n as a Steiner 2-design $S(2, n+1, n^2+n+1)$, see Definition 11.5. The most important open problem in this area is the question of the existence of projective planes of composite order. No projective plane of composite (non prime-power) order is known. We want to prove a famous non-existence theorem, the **Bruck-Ryser theorem.**

16.1 Theorem. *If $n \equiv 1 \pmod 4$ or $\equiv 2 \pmod 4$ is the order of a projective plane, then n is sum of two integer squares.*

Proof. The proof can be reduced to the Hasse-Minkowski theorem. We present the proof from [9], which does not make explicit use of the Hasse-Minkowski theorem. Some facts from number theorem are used:

16.2 Lemma. *Every positive integer is sum of four integer squares.*

16.3 Lemma. *If a positive integer is sum of two rational squares, then it is sum of two integer squares.*

Let n as above be the order of a projective plane. Then $v = n^2+n+1 \equiv 3 \pmod 4$. Number the points $P_1, \dots, P_v$ and the lines $l_1, \dots, l_v$. We work in the polynomial ring with indeterminates $x_1, \dots, x_v$ and real (or rational) coefficients. Observe that there is one indeterminate x_i for each point P_i. Let $L_k = \sum_i x_i$, where the sum is over all i such that $P_i \in l_k$.

Consider the quadratic form (homogeneous quadratic polynomial)

$$\sum_{k=1}^{v} L_k^2 = 2\sum_{i \neq j} x_i x_j + (n+1)\sum_{j=1}^{v} x_j^2.$$

This identity follows from the basic properties of a projective plane. The term $x_i x_j$ can occur only in L_k^2, where l_k is the uniquely determined line containing P_i and P_j. The last term follows from the fact that each point is on precisely $n+1$ lines. We can rewrite this basic equation in the more convenient form

$$\sum_{k=1}^{v} L_k^2 = n \sum_{i=1}^{v} x_i^2 + (\sum_{i=1}^{v} x_i)^2.$$

Introduce an additional indeterminate x_{v+1} and abbreviate $S = \sum_{i=1}^{v} x_i$.

$$\sum_{k=1}^{v} L_k^2 + n x_{v+1}^2 = n \sum_{i=1}^{v+1} x_i^2 + S^2.$$

The following important identity is easy to check directly:

$$(a^2+b^2+c^2+d^2)(x^2+y^2+z^2+w^2) =$$

$$(ax-by-cz-dw)^2 + (bx+ay-dz+cw)^2 +$$

$$+(cx+dy+az-bw)^2 + (dx-cy+bz+aw)^2.$$

By Lemma 16.2 we can find integers a, b, c, d such that $n = a^2+b^2+c^2+d^2$. We can interpret the identity using the matrix

$$A = \begin{pmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{pmatrix}.$$

The identity says

$$n(x_1^2+x_2^2+x_3^2+x_4^2) = y_1^2+y_2^2+y_3^2+y_4^2,$$

where $(y_1, y_2, y_3, y_4) = (x_1, x_2, x_3, x_4)A$. Also, $det(A) = n^2$. Applying A^{-1} (a matrix with rational entries) we see that we can express each of x_1, x_2, x_3, x_4 as a linear combination of the y_1, y_2, y_3, y_4 with rational coefficients. Use this to eleminate x_1, x_2, x_3, x_4 from the basis equation. The right side is then $y_1^2+y_2^2+y_3^2+y_4^2+n\sum_{i=5}^{v+1} x_i^2 + S^2$. The new basic equation is an identity in the polynomial ring with indeterminates $y_1, y_2, y_3, y_4, x_5, \ldots, x_{v+1}$. As $v+1$

is a multiple of 4 we can continue this process, using 4 of the x_i each time, and end up with an identity

$$\sum_{k=1}^{v} L_k^2 + n x_{v+1}^2 = \sum_{i=1}^{v+1} y_i^2 + S^2$$

in the indeterminates $y_1, \dots, y_{v+1}$.

Consider point P_1 and line l_1. We have $L_1 = \sum_i \alpha_i y_i$. We aim at substituting for y_1 a linear combination of the remaining indeterminates such that $L_1^2 = y_1^2$. Assume at first $\alpha_1 \neq 1$. Use the substitution $y_1 = \frac{1}{1-\alpha_1} \sum_{i \neq 1} \alpha_i y_i$. Then $L_1 = y_1$. If $\alpha_1 = 1$ we find a similar substitution such that $L_1 = -y_1$. In both cases $L_1^2 = y_1^2$. It follows that we can substitute for y_1 a linear combination of the $y_i, i > 1$ such that

$$\sum_{k=2}^{v} L_k^2 + n x_{v+1}^2 = \sum_{i=2}^{v+1} y_i^2 + S^2$$

and this is a homogeneous quadratic identity in the polynomial ring with indeterminates $y_2, \dots, y_{v+1}$. Continuing in this fashion we will finally arrive at an identity

$$n x_{v+1}^2 = y_{v+1}^2 + S^2$$

in the polynomial ring with indeterminate y_{v+1} where $x_{v+1} = a y_{v+1}$ and $S = b y_{v+1}$. Comparing coefficients we obtain $a^2 n = 1 + b^2$. If $a = 0$, then $y_{v+1}^2(1 + b^2) = 0$, contradiction. It follows $n = \frac{1}{a^2} + (b/a)^2$, a sum of two rational squares. Lemma 16.3 implies that n is sum of two integer squares. □

It is an important theme of number theory to study which values are attained by quadratic forms and how often. We studied this question over finite fields. The question which integers can be written as sums of a given number of squares is of this type. It is a classical result that n can be written as a sum of two squares if and only if the square-free part of n is not divisible by primes $\equiv 3 \pmod 4$. This is a computationally more convenient form of the Bruck-Ryser theorem. The smallest orders excluded by the Bruck-Ryser theorem are $n = 6, 14, 21, 22, 30$. The only order which has been excluded by different means is $n = 10$. This was done by exhaustive computer search, based on a coding-theoretic approach.

16.4 Definition. *Let $V = V(2m, q)$. A* **spread** *of V is a family of $q^m + 1$ subspaces of dimension m, which partition the nonzero vectors of V.*

We encountered the concept of a spread in Chapter 14, where we partitioned the isotropic or singular points of a bilinear, sesquilinear or quadratic form into totally isotropic (totally singular) subspaces. Definition 16.4 is therefore a natural concept. Symplectic spreads are special cases. In particular we know that each even-dimensional vector space does have a spread. Next we want to see how spreads can be used to construct projective planes.

16.5 Lemma. *Let W_1, W_2 be different m-dimensional subspaces of $V = V(2m, q)$ such that $W_1 \cap W_2 = \{0\}$ and $x, y \in V$. Then $|(W_1 + x) \cap (W_2 + y)| = 1$.*

Proof. Assume $w_1 + x = w_2 + y$ and $w_1' + x = w_2' + y$, with obvious notation. By subtraction $w_1 - w_1' = w_2 - w_2' \in W_1 \cap W_2$. It follows $w_1 = w_1', w_2 = w_2'$. The intersection cardinality is therefore at most 1. As the q^m cosets of W_2 partition the vectors of V the intersection cardinality must be $= 1$. □

16.6 Definition. *Let $S = \{W_1, \ldots, W_{q^m+1}\}$ be a spread of $V = V(2m, q)$. Define an incidence structure $\Pi(S)$ as follows: The points are $P_1, \ldots, P_{q^m+1}$ (one point for each element of the spread) and the vectors of V.*

The lines are $l_\infty = \{P_1, \ldots, P_{q^m+1}\}$ (the line at infinity) and for every coset $W_i + x$ of some W_i a line containing the points of $W_i + x$ and P_i.

16.7 Theorem. *Let $S = \{W_1, \ldots, W_{q^m+1}\}$ be a spread of $V = V(2m, q)$. Then $\Pi(S)$ is a projective plane of order q^m.*

Proof. Let $n = q^m$. We have $q^{2m} + q^m + 1 = n^2 + n + 1$ points and $1 + (q^m + 1)q^m = n^2 + n + 1$ lines. Each line has $q^m + 1$ points, each point is on $q^m + 1$ lines. It follows from Lemma 16.5 that any two lines intersect in precisely one point. It follows that any two points are on precisely one line. □

The discussion in Chapter 13 shows what an automorphism of a projective plane is and when two planes are isomorphic (have the same structure). A bijective mapping between the point sets is an isomorphism if the image of each line in the first plane is a line in the second plane. Two planes are isomorphic if such an isomorphism exists. An isomorphism from a plane onto itself is an automorphism. The automorphisms form a group, the automorphism group or symmetry group of the plane. The classical plane $PG(2, q)$, where $q = p^f$, admits $P\Gamma L(3, q)$ of order $fq^3(q^3 - 1)(q^2 - 1)$ as a symmetry group. It is a theorem that $P\Gamma L(3, q)$ is indeed the full automorphism group of $PG(2, q)$. Recall the case of the Fano plane $PG(2, 2)$ whose automorphism group has order $2^3 7 \cdot 3 = 168$. Next we observe that the projective planes $\Pi(S)$ have rich automorphism groups as well.

16.8 Theorem. *Let $S = \{W_1, \ldots, W_{q^m+1}\}$ be a spread of $V = V(2m, q)$. For each $x \in V$ let τ_x be the permutation of the points of $\Pi(S)$, which maps $y \mapsto y + x$ for $y \in V$ and fixes all P_i. Then τ_x is an automorphism of $\Pi(S)$. The τ_x for $x \in V$ form an abelian group of automorphisms, which is transitive on the points outside l_∞ (the* **affine points***).*

Proof. By definition τ_x fixes the line at infinity. The image of the line defined by $W_i + y$ is the line given by $W_i + x + y$. Clearly the composition of τ_x and τ_y is τ_{x+y}. If x, y are affine points, then $y = \tau_{y-x}(x)$. □

Problems

1. Show that prime-powers p^f do satisfy the conditions of the Bruck-Ryser theorem.
2. Show that numbers $n \equiv 6 \pmod 8$ cannot be orders of projective planes.

Chapter 17

Generalized polygons

Geometric concepts can also be expressed in terms of graphs. We assume a certain familiarity with this concept and will therefore be brief in our description of basic terminology. A **graph** consists of a ground set V (the **vertices**) and a family E of unordered pairs (**edges**) from V. In particular we exclude here more general concepts involving directed edges and edges with more or less than 2 vertices. A **path** from vertex x to vertex y is a tuple $(x = x_0, x_1, \ldots, x_n = y)$, where $\{x_i, x_{i+1}\}$ is an edge for all i. The length of this path is n. A graph is connected if for any two vertices there is a path connecting them. A connected graph defines a metric on the vertices, where $n = d(x, y)$ is the length of a shortest path from x to y.

17.1 Definition. *The* **diameter** $d = d(\Gamma)$ *of a connected graph is the maximum distance between two vertices.*
The **girth** *of a connected graph* Γ *is the length of the shortest cycle in* Γ. *Here a* **cycle** *of length* n *is a closed path* $(x_0, x_1, \ldots, x_{n-1}, x_0)$ *where the* $x_i, i = 0, \ldots, n-1$ *are different.*

A graph is **bipartite** if the vertices can be partitioned into two non-empty subsets $V = L \cup R$ such that all edges are from L to R. It is easy to see that Γ is bipartite if and only if it does not contain cycles of odd length.

Assume Γ is connected. If it does not contain any cycles it is called an **acyclic** graph or a **tree.** In particular the girth is not even defined for trees. Trees should be excluded from the present discussion. Let us study the relationship between diameter and girth. Fix g and consider a cycle of length g. The distance of any two points of the cycle in our graph Γ is the same as the distance in the cycle: if there was a shorter path we could combine it with a path along the cycle to obtain a cycle of length $< g$. This shows that $d \leq g/2$ if g is even and $d \leq (g-1)/2$ if g is odd.

17.2 Proposition. *Let Γ be a connected graph, which is not a tree. The diameter d and girth g satisfy $d \geq \lfloor g/2 \rfloor$.*

If we restrict attention to bipartite graphs, then g is even and d is therefore at least $g/2$. In the case of equality an interesting situation arises:

17.3 Definition. *Let Γ be a connected bipartite graph with vertex classes L and R (recall $V = L \cup R$ and there are no edges inside L or R).*
We call Γ a **generalized n-gon** *if*

$$d(\Gamma) = n \text{ and } g(\Gamma) = 2n.$$

Call Γ **non-degenerate** *if for each $x \in V$ there is a vertex y such that $d(x, y) = n$.*

We have seen that generalized n-gons have extremal graph-theoretic properties. The concept could have been expressed just as well in geometric terminology. In fact, define vertices from L to be points and vertices from R to be lines. Point P is a point of line l if $\{P, l\}$ is an edge of Γ (in this terminology it can happen that different lines have all their points in common). We will use both graph-theoretic and geometric terminology. For example, the number of points on line l is the degree of vertex l. Dually, the number of lines through point P is the degree of vertex P.

17.4 Definition. *A generalized n-gon has* **order** (s, t) *if each line has $s + 1$ points and each point is on $t + 1$ lines. It is* **thick** *if $s > 1$ and $t > 1$.*

A generalized n-gon of order $(1, 1)$ simply is an ordinary n-gon (a cycle of length n). In case $n = 2$ we obtain a **complete bipartite graph:** each pair (P, l) is an edge.

17.5 Proposition. *Let Γ be a non-degenerate generalized triangle (3-gon) of order (s, t). Then $s = t$. If $s > 1$, then Γ is a projective plane of order s.*

Proof. Compare Definition 11.5 and our earlier discussion of projective planes. As $d = 3$ any two points are on a line, as $g = 6$ this line is uniquely determined. The dual statements also holds. By non-degeneracy for each line l there is a point $P \notin l$ and the dual statement. Considering the lines through P and the points on l, where $P \notin l$, shows $s = t$. $\square$

A generalized 4-gon of order (s, t) is equivalent to a generalized quadrangle of order (s, t), compare Chapter 11. The absence of 4-cycles implies that any two points have at most one line in common and the dual. Let $P \notin l$. As $d(P, l) = 3$ in Γ there is a line l' such that $P \in l'$ and l' intersects l in a point P. The absence of cycles of length ≤ 6 shows that (P, l', Q, l) is the only path of length 3 from P to l, thus verifying the main axiom of Definition 11.7.

We see that the notion of a generalized polygon is a generalization of the notions of projective planes and generalized quadrangles. The natural question arises for which n there exist (non-degenerate, thick) generalized n-gons. The basic theorem on this subject is the Feit-Higman theorem [5].

17.6 Theorem (Feit-Higman). *Let Γ be a non-degenerate generalized n-gon of order (s, t), where $st > 1$. Then*

$$n \in \{2, 3, 4, 6, 8, 12\}$$

If Γ is thick, then $n \neq 12$. If Γ is thick and $n = 6$, then st is a square. If Γ is thick and $n = 8$, then $2st$ is a square.

Most of the examples are related to classical groups. We want to describe the family of generalized hexagons related to the groups $G_2(q)$. These hexagons exist for every prime-power q and have order (q, q). They can be derived from the 7-dimensional orthogonal geometry.

The $G_2(q)$-hexagons and orthogonal geometry

Consider $V = V(7, q)$ with the quadratic form

$$Q(x_1, \ldots, x_7) = x_1x_2 + x_3x_4 + x_5x_6 + x_7^2$$

(three quadratic hyperbolic planes and an anisotropic 1-dimensional space). We have $V = H_1 \perp H_2 \perp H_3 \perp \langle v_0 \rangle$, where the H_i are (quadratic) hyperbolic planes with basis v_i, w_i, and $Q(sv_i + tw_i) = st, Q(v_0) = 1$. Let $L = \langle v_1, v_2, v_3 \rangle, R = \langle w_1, w_2, w_3 \rangle$. Observe that L and R are totally singular subspaces (planes).

The singular (isotropic) points of Q will be the points of the hexagon. We know that $|Q(6, q)| = (q^6 - 1)/(q - 1)$ (see Chapter 6 for the odd characteristic case. In characteristic 2 the formulas were the same).

construction idea: for each singular point P find a totally singular plane $W(P)$ containing P. We may call these the **tangent planes.** The lines of the hexagon through P will be all the lines t through P contained in $W(P)$.

Let H be the hyperplane generated by L and R. Observe that the restriction of Q to H is a hyperbolic space, with $(q^2+q+1)(q^2+1)$ points (see Theorem 6.13). We imagine $Q(6,q)$ partitioned into four parts, the q^2+q+1 points of $\langle L\rangle$, the q^2+q+1 points of $\langle R\rangle$, the remaining $(q^2+q+1)(q^2-1)$ singular points in H (**inner points**) and the $q^2(q^3-1)$ singular points outside H (the **outer points,** in fact, $(q^2+q+1)(q^2+1)+q^2(q^3-1) = (q^2+q+1)(q^3+1) = |Q(6,q)|$).

Call the lines $\langle x,y\rangle$, where $x\in L, y\in R, (x,y)=0$ **LR-lines.** There are $(q^2+q+1)(q+1)$ LR-lines. They partition the inner points: each inner point is on precisely one LR-line (simply count: each LR-line contains $q-1$ inner points, and the number of LR-lines multiplied by $q-1$ equals the number of inner points).

We need to define the planes $W(P)$ in an appropriate way. If $P=\langle v\rangle\in L$, define $W(P)=\langle v, v^\perp\cap R\rangle$. Analogously for $P=\langle w\rangle\in R$ define $W(P)=\langle w, w^\perp\cap L\rangle$. In order to define the tangent planes to inner points we use a subgroup of the orthogonal group.

17.7 Lemma. *Let K be the group consisting of the matrices*

$$K(A)=\begin{pmatrix} A & 0 & 0\\ 0 & (A^{-1})^t & 0\\ 0 & 0 & 1\end{pmatrix}$$

(with respect to our standard basis), where $A\in SL(3,q)$. Then $K\subset O(7,q)$ is transitive on the LR-lines and on inner points.

Proof. $K(A)$ maps $v_i\mapsto v_iA,\ w_j\mapsto w_j(A^{-1})^t$. The scalar product of the images is $v_iAA^{-1}w_j^t=v_iw_j^t=(v_i,w_j)$. As the v_i,w_j are singular vectors we conclude that $K\subset O(7,q)$.

It is easy to see that K is transitive on inner points. In fact, in order for $K(A)$ to stabilize the inner point $\langle v_1+w_2\rangle$ matrix A must have the form

$$A=\begin{pmatrix} \lambda & 0 & 0\\ a_{21} & 1/\lambda & a_{23}\\ a_{31} & 0 & 1\end{pmatrix}$$

where a_{21},a_{23},a_{31} are arbitrary and $\lambda\neq 0$. It follows that the stabilizer has order $(q-1)q^3$. As $|K|=q^3(q^2-1)(q^3-1)$ the length of the orbit of

$\langle v_1 + w_2\rangle$ under K is $(q+1)(q^3-1) = (q^2-1)(q^2+q+1)$. This is precisely the number of inner points. Transitivity on inner points implies transitivity on LR-lines. $\square$

Observe that the stabilizer of $\langle v_1, w_2\rangle$ consists of the $K(A)$ where

$$A = \begin{pmatrix} \lambda & 0 & 0 \\ a_{21} & \mu & a_{23} \\ a_{31} & 0 & \frac{1}{\lambda\mu} \end{pmatrix}$$

For the inner point $P = \langle v_1 + w_2\rangle$ we want $W(P)$ to contain the LR-line $\langle v_1, w_2\rangle$ containing P. As

$$\langle v_1, w_2\rangle^\perp = \langle v_1, v_3, w_2, w_3, v_0\rangle$$

we see that $W(\langle v_1\rangle)$ and $W(\langle w_2\rangle)$ are the only candidates contained in H. As we want tangential planes to different points to be different we must have $W(P) \notin H$.

Define $W(P)$ to be a totally singular plane containing $\langle v_1, w_2\rangle$, which is different from $W(\langle v_1\rangle)$ and from $W(\langle w_2\rangle)$. We choose $W(P) = \langle v_1, w_2, v_3 - w_3 + v_0\rangle$. The stabilizer of $W(P)$ equals the stabilizer of P. We can therefore choose $W(P)K(A)$ to be the tangential plane of $PK(A)$. The tangential planes to inner points therefore form an orbit under K.

Let $x = v_3 - w_3 + v_0$. Then $Q = \langle x\rangle$ is an outer point. The stabilizer of Q consists of the $K(A)$, where

$$A = \begin{pmatrix} X & 0 \\ 0 & 1 \end{pmatrix}.$$

The index of $SL(2,q)$ in $SL(3,q)$ is $q^2(q^3-1)$, the number of singular points outside H. It follows that K is transitive on these points. Let T denote the set of lines through inner points P contained in $W(P)$, which are not contained in H. Call T the set of **outer tangents.** As $W(P) \cap H$ is an LR-line, each inner point is on q outer tangents. Count the incidences of outer tangents and outer points in two ways:

$$(q^2+q+1)(q^2-1)q^2 = q^2(q^3-1)\alpha,$$

where α is the number of outer tangents through a given outer point. We conclude $\alpha = q+1$: each outer point is on $q+1$ outer tangents. One of the outer tangents through the outer point Q is $\langle x, v_1 + w_2\rangle$. The images of $P = \langle v_1 + w_2\rangle$ under the stabilizer of Q are the points on the line $m = \langle v_1 + w_2, v_2 - w_1\rangle$. It follows that the $q+1$ outer tangents through Q are on a plane, the plane generated by Q and the line m. Define $W(Q) = \langle Q, m\rangle$, analogously for the remaining outer points. Observe that $m = W(Q) \cap H$. We have defined a tangent plane $W(P)$, a totally singular plane containing P, for every point $P \in Q(6, q)$.

17.8 Definition. *We define an incidence structure $G(q)$ as follows:*

- *The points of $G(q)$ are the points $P \in Q(6, q)$.*
- *For each point P, the lines of $G(q)$ through P are the lines of the plane $W(P)$ through P. Call such a line a* **tangent to P.**

The number of points clearly is $(q^6 - 1)/(q - 1)$. Lines are the LR-lines and the outer tangents. The total number of lines is

$$(q^2 + q + 1)(q + 1) + (q^2 + q + 1)(q^2 - 1)q = (q^6 - 1)/(q - 1).$$

Each line has $q+1$ points, each point is on $q+1$ lines. The tangents to points from L or R are precisely the LR-lines. Each inner point is on one LR-line and on q outer tangents. Each inner point is on $q + 1$ outer tangents.

In order to prove that $G(q)$ is a generalized hexagon it is convenient to use the graph-theoretic terminology of Definition 17.3. We view points and lines as vertices of a graph Γ where P, l form an edge if $P \in l$, and we need to show $d(\Gamma) = 6,\ g(\Gamma) = 12$.

As Γ is bipartite cycles have even length. Assume there is a cycle of length 10. It consists of 5 points and 5 tangents:

$$(P_1, l_1, P_2, l_2, P_3, l_3, P_4, l_4, P_5, l_5, P_1).$$

We have that $P_2, P_3 \in P_1^\perp$. As $P_1, P_3 \in W(P_2)$ we also have $P_3 \in P_1^\perp$, by symmetry $P_4 \in P_1^\perp$. It follows that the P_i are pairwise orthogonal. Each three contiguous P_i generate a totally singular plane. As our quadratic form has Witt index 3, $\langle P_1, P_2, P_3, P_4, P_5\rangle = E$ is a plane. By definition of the tangents we have $E = W(P_1) = W(P_2) = \cdots = W(P_5)$, contradiction.

Clearly cycles of even shorter length are impossible. We have shown $g(\Gamma) \geq 12$. It remains to show $d(\Gamma) \leq 6$. Let d be the distance in Γ. Fix a point P. There are $(q + 1)q$ points at distance 2 and $(q + 1)q^3$ points at distance 4 from P. All these points are contained in $P^\perp$. On the other hand the number of singular vectors in $P^\perp$ is q^5, so the number of points in $P^\perp$ different from P is $(q^5-1)/(q-1)-1 = q^4+q^4+q^2+q$. This equals $(q+1)q+(q+1)q^3 = (q+1)q(q^2+1)$, the number of points at distance 2 or 4 from P.

17.9 Lemma. *For every singular point P the points at distance 2 or 4 from P in Γ are precisely the points in $P^\perp$.*

Fix a point P again. Consider paths (P, l_1, Q, l_2, R) to one of the points at distance 4. Continue this path by (l_3, S), where $l_3 \neq l_2, S \in l_3, s \neq R$. The absence of cycles of length ≤ 10 implies that $d(P,S)=6$. As there are $(q+1)q^3$ points at distance 4 from P we count $(q+1)q^5$ such paths of length 6. Each endpoint can be obtained at most $q+1$ times, so the number of points $\notin P^\perp$ at distance 6 from P is $\geq q^5$. As this is the total number of points $\notin P^\perp$ we have shown $d(\Gamma)=6$. This proves that $G(q)$ is indeed a generalized hexagon.

Chapter 18

Diagram geometries

A very general notion of a finite geometry would be based on objects of several types (points, lines, planes, ...) and their **incidences.** The **rank** of a finite geometry is the number of different types of objects it is composed of. For example, a projective plane is defined by points, lines and their incidences. The rank is 2. Of course, the names assigned to the different types of objects are a matter of convention. In $PG(3,q)$ we have points, lines and planes. This is a rank 3 geometry. Clearly $PG(r,q)$ has rank r.

We encountered several important types of rank 2 geometries. Let us assign diagrams to these geometries.

Projective planes

The diagram assigned to projective planes consists of two nodes and a simple line joining them. Here the left node, say, represents points, the right node represents lines and we use a single straight connecting line to denote projective planes. Observe that the diagram has two nodes because we have a rank 2 geometry. In general the number of nodes will equal the rank of the geometry.

Generalized quadrangles

The diagram consists of a double line joining the two nodes.

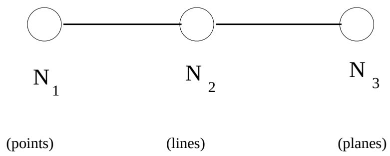

Figure 18.1: The diagram of $PG(3)$

Generalized hexagons

The nodes are joined by a triple line.

Complete bipartite graphs

This geometry is so trivial that it is easy to miss. Every point is incident with every line. The geometry looks less artificial if written in graph-theoretic notation. The left vertices represent the points of the geometry, the right vertices represent the lines. A point and a line form an edge of the graph if the point is incident with the line. In the **complete** bipartite graph all such pairs form edges. The corresponding diagram consists of two nodes, which are not joined at all. Complete bipartite graphs can also be considered as generalized 2-gons, see Chapter 17.

In order to illustrate the idea behind diagram geometries consider the rank 3 geometry $PG(3, q)$. The diagram has 3 nodes: N_1 representing points, N_2 representing lines and N_3 representing planes. In order to decide which connection to draw between N_1 and N_2 fix an arbitrary plane E and consider the rank 2 geometry (a **residual geometry**) consisting of the point and lines incident with E. This residual geometry is of course always a projective plane, so we draw a single line from $N_1 \leftrightarrow N_2$. Next fix a point P and consider the corresponding residual geometry consisting of the lines and planes through P. Again, this is a projective plane and so we draw a line $N_2 \leftrightarrow N_3$. Finally, fix a line l. All points on l will be incident with all planes through l, so the residual geometry is a complete bipartite graph. Consequently we draw no line from N_1 to N_3. The diagram of $PG(3, q)$ is a path, where each line is a single line.

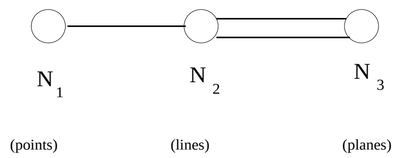

Figure 18.2: The diagram of 7-dim orthogonal geometry

Proceeding in an analogous fashion we see that the diagram of $PG(r)$ is a path, where each line is a single line.

As another example consider the 7-dimensional orthogonal geometry (we used it in Chapter 17 to construct a generalized hexagon). The objects of the geometry are the totally singular points, lines and planes (recall that the Witt index is 3, so there are no higher-dimensional totally singular subspaces). We have a rank 3 geometry. As before denote the nodes of the diagram by N_1, N_2, N_3, corresponding to points, lines, planes, respectively. It is clear that the residual geometry of a plane is a projective plane and that the residual geometry of a line is a complete bipartite graph. Fix a singular point. The geometry defined by the totally singular lines through P and the totally singular planes through P is the geometry formed by the singular points and totally singular lines in 5-dimensional orthogonal geometry. As the Witt index is 2 this is a generalized quadrangle.

The origins of this business are in group theory. The classical finite simple groups can be derived from the simple complex **Lie algebras** (see [3]). These Lie algebras have been completely classified. They can be described by certain highly symmetric finite sets of vectors in Euclidean r-space, the **root systems.** The description of the root systems can be reduced to certain diagram on r nodes, the **Dynkin diagrams.** It turns out that the corresponding classical simple groups are groups of automorphisms of rank r diagram geometries, whose diagram is exactly the Dynkin diagram. These geometries associated with classical groups are known as **buildings.** Diagram geometries generalize this notion. The objective behind the generalization is to obtain geometric descriptions not only for the **classical** finite simple groups but also for the **sporadic** groups. There are only 26 such sporadic groups, but they cause as much trouble as all the infinite series of classical groups taken together.

Chapter 19

The sporadic A_7-geometry

Let Ω be a 7-element set, $\Omega = \{x, 1, 2, 3, 4, 5, 6\}$. We start from the following elementary question: How many different Fano planes exist on Ω?
One way to approach this is to use a basic fact from group theory, the **orbit formula:** the length of an orbit under the action of a permutation group equals the index of the stabilizer group.
As all Fano planes are isomorphic, they form one orbit under the action of the symmetric group S_7. The stabilizer of a Fano plane E in S_7 is by definition the automorphism group of E. This is the linear group $GL(3, 2)$, the simple group of order 168, see the Problems section. It follows that the number of different Fano planes is $7!/168 = 30$. It is easy to confirm this result by a purely combinatorial argument. Now restrict to the alternating group A_7. As $GL(3, 2)$ is a simple group, it is contained in A_7 (the commutator group of $GL(3, 2)$ is contained in the commutator group of S_7, which is A_7, and the commutator group of the simple group $GL(3, 2)$ is $GL(3, 2)$ itself).

As $GL(3, 2)$ has index 15 in A_7, the group A_7 has two orbits $\mathcal{E}_1$ and $\mathcal{E}_2$ of Fano planes, each of length 15.

19.1 Lemma. *There are* 30 *different Fano planes on a given* 7*-element set. They form one orbit under* S_7*, two orbits of length* 15 *each under* A_7*.*

Upon counting incident pairs of triples (3-subsets of Ω) and Fano planes, we see that every triple is a line of exactly 6 Fano planes, 3 from each A_7-orbit. Let $E \in \mathcal{E}_1$ be a Fano plane and $H \cong GL(3, 2)$ its stabilizer. How does H act on the remaining 14 elements (Fano planes) of $\mathcal{E}_1$? The element of order 7 shows that H either is transitive or has two orbits of length 7. As $\binom{15}{2} = 105$ is odd, A_7 certainly is transitive on the unordered pairs from

$\mathcal{E}_1$ (A_7 is **2-homogeneous**). By counting triples (E, E', T), where E and E' are different Fano planes from $\mathcal{E}_1$ and T a triple, which is a line in both, we see that different Fano planes from the same A_7-orbit have exactly one line in common. This shows in particular that no more than two linewise disjoint Fano planes can be found.

Consider the action of H on $\mathcal{E}_2$. An element of order 7 has orbits of lengths $1, 7, 7$. Assume H stabilizes a (unique) Fano plane $E' \in \mathcal{E}_2$. This yields an orbit of length 7 under the action A_7 of pairs (E_1, E_2) such that $E_i \in \mathcal{E}_i$ and E_1, E_2 have the same stabilizer. Let E_1 and E_2 have x lines in common. and y the number of pairs (E_1, E_2) from the orbit having a given line in common. The usual counting argument shows $7 \times x = 35y$, hence $x = 5$. This is a contradiction as no two Fano planes have as many as 5 lines in common. We conclude that the stabilizer H of $E \in \mathcal{E}_1$ has two orbits of lengths 7 and 8, respectively, on $\mathcal{E}_2$. It follows that we have two orbits of pairs (E_1, E_2), where $E_i \in \mathcal{E}_i$ under the action of A_7, one of length $15 \times 7 = 105$, the other of length $15 \times 8 = 120$. Let pairs in the shorter orbit have x lines in common and y the number of lines which pairs in the longer orbit have in common. Count triples (E_1, E_2, T), where T is a common line. This yields $105x + 120y = 35 \times 9$. It follows that 7 divides y, consequently $y = 0$ and $x = 3$.

19.2 Theorem. *Let $\mathcal{E}_1$, $\mathcal{E}_2$ be the orbits of Fano planes under the action of A_7. Then A_7 is doubly transitive on $\mathcal{E}_1$ and on $\mathcal{E}_2$. There are two orbits of pairs (E_1, E_2) where $E_i \in \mathcal{E}_i$, of lengths* 105 *and* 120. *The number of common lines is* 3 *for the shorter, it is* 0 *for the longer orbit.*

Proof. It remains to prove the 2-transitivity of A_7 on $\mathcal{E}_1$. Choose

$$E = \{x12, x34, x56, 135, 146, 236, 245\}$$

$$E' = \{x12, x35, x46, 136, 145, 234, 256\}$$

$$E'' = \{x12, x36, x45, 235, 246, 134, 156\}$$

Then E, E', E'' are Fano planes. They belong to the same A_7-orbit as they pairwise have exactly one line 012 in common. The permutation $\sigma = (1,2)(3,4) \in A_7$ fixes E and maps $E' \to E''$. □

It follows from Theorem 19.2 that the maximum number of linewise disjoint Fano planes which can be constructed on a given 7-set is 2. This result is attributed to Cayley [4].

Next we show how the action of A_7 on the orbit $\mathcal{E}_1$ (analogously on $\mathcal{E}_2$) can be used to obtain an embedding of A_7 in $GL(4, 2)$.

19.3 Definition. *Let* $V = \{0\} \cup \mathcal{E}_1$. *Define an addition on* V *with* 0 *as neutral element such that* $v + v = 0$ *for all* $v \in V$ *and the sum of any two different Fano planes from* $\mathcal{E}_1$ *is the third plane containing the line which the first two have in common.*

Observe that Definition 19.3 makes sense as we know that two Fano planes from the same A_7-orbit have precisely one common line and there is exactly one further such Fano plane containing that line. As an example, the Fano planes E, E', E'' from the proof of Theorem 19.2 sum to 0.

19.4 Lemma. A_7 *is regular on the triples* (E, F, G) *of Fano planes from* $\mathcal{E}_1$ *where* E, F, G *are different and* $G \neq E + F$. *For each such triple the lines of pairwise intersection form a triangle. There is a uniquely determined point of* Ω *which is outside this triangle.*

Proof. The number of these triples is $15 \times 14 \times 12 = |A_7|$. It suffices to show that no nontrivial element of A_7 fixes such a triple. The stabilizer of E and F has order 12 and clearly is isomorphic to A_4. Let $1 \neq \sigma$ in the stabilizer of E, F and G. Consider the three triples, which are the lines that two of our Fano planes have in common. As any two of them belong to a common Fano plane, they pairwise intersect in one point. Assume they form a triangle. Then σ is the identity, contradiction. On the other hand these triples cannot form a concurrent bundle as in this case E, F, G would have three lines in common. $\square$

19.5 Proposition. V *is an elementary abelian group of order* 16.

Proof. Addition is certainly commutative, and each element has order 2, by definition. Only associativity needs to be shown:

$$(E + F) + G = E + (F + G).$$

It can be assumed that E, F, G are different and nonzero, $G \neq E + F$. By Lemma 19.4 it suffices to check associativity for one special triple E, F, G of Fano planes. We are without restriction in the following situation:

$$E = \{x12, x34, x56, 135, 146, 236, 245\},$$

$$F = \{x12, x35, x46, 136, 145, 234, 256\},$$

$$G = \{x34, x15, x26, 136, 124, 235, 456\}.$$

The triangle of pairwise intersection is $\{x12, x34, 136\}$ and 5 is the unique element of Ω not on this triangle. The Fano plane $E + F = \{x12, x36, x45, 235, 246, 134, 156\}$ shares 136 with G. It follows

$$(E + F) + G = \{235, 2x4, 216, x56, x13, 346, 145\}.$$

On the other hand F and G share 136, $F + G = \{136, 1x4, 125, x23, x56, 246, 345\}$ and finally

$$E + (F + G) = \{x56, x13, x24, 145, 126, 235, 346\}$$

which happens to coincide with $(E + F) + G$. $\square$

As the addition of V is respected by the action of A_7 we conclude that we have an embedding of A_7 in the linear group $GL(4, 2)$, the group of automorphisms of an elementary abelian group of order 16, equivalently of a 4-dimensional vector space over $\mathbb{F}_2$.

19.6 Theorem. *A_7 possesses a 2-transitive permutation representation on the 15 nonzero vectors of a 4-dimensional vector space over $\mathbb{F}_2$. The stabilizer of a vector is $\cong GL(3, 2)$. Moreover A_7 is regular (sharply transitive) on ordered triples (a, b, c) of linearly independent vectors.*

It follows from Theorem 19.6 that A_7 is a subgroup of $GL(4, 2)$. The simple group $GL(4, 2)$ acts on the 8 cosets of A_7. This yields an embedding of $GL(4, 2)$ in S_8. As $GL(4, 2)$ is simple it is embedded in A_8. These groups have the same order. It follows that these simple groups are isomorphic.

19.7 Theorem. *The simple groups $GL(4, 2)$ and A_8 are isomorphic.*

How do we compute with the embedding of A_7 in $GL(4, 2)$? It is combinatorially obvious that each bundle of 3 concurrent lines is contained in precisely 2 Fano planes, one from each A_7-orbit. Let E, F be Fano planes from the same orbit and $x12$ their common line. Assume E contains the bundle $\{x12, x34, x56\}$ and F contains $\{x12, x35, x46\}$. Then $E + F$ contains $\{x12, x36, x45\}$ and $E + F$ is the uniquely determined Fano plane from the same orbit that contains this bundle.

Let x be an element which is not on the line that E and F have in common. The bundle of lines of E through x defines a 1-factor of $\Omega \setminus \{0\}$, likewise for F. These two 1-factors together form a cycle of length 6 on $\Omega \setminus \{x\}$. The 1-factor determined by $E + F$ consists of the diagonals of this cycle. We have seen the following:

19.8 Proposition. *Let E, F be two Fano planes in the same A_7-orbit, defined on the ground set Ω. Let l be the line that E and F have in common and $a \in \Omega$.*

If $a \in l$ and $\{a, b, c\}$, $\{a, d, e\}$ are the lines $\neq l$ of E containing a, and $\{a, b, d\}$, $\{a, c, e\}$ the corresponding lines of F, then $E + F$ contains the lines $l, \{a, b, e\}$ and $\{a, c, d\}$.

If $a \notin l$, then the pairs collinear with a on a line of $E+F$ are the diagonals of the 6-cycle which the bundles of lines through a in E and in F define on $\Omega \setminus \{a\}$.

As an example consider the Fano planes E, F from the proof of Proposition 19.5. The first rule when applied to x shows $\{x12, x36, x45\} \subset E+F$. Apply the second rule to the element 6. The 1-factors on $\Omega \setminus \{6\}$ are $\{x5, 14, 23\}$ and $\{x4, 13, 25\}$. Together they define the cycle $(x, 5, 2, 3, 1, 4)$ whose diagonals are $x3, 15, 24$. This shows that $E + F$ contains $\{x36, 156, 246\}$. Proposition 19.8 also suggests how the pairs and the 1-factors on a 6-set can be given an algebraic structure. Let $X = \{1, 2, 3, 4, 5, 6\}$. Let W consist of 0 and of the unordered pairs from X. Here we write ij for $\{i, j\}$. Define an addition on W by $0 + w = w,\ w + w = 0$ and

$$12 + 13 = 23,\ 12 + 34 = 56$$

(the sum of two intersecting pairs is the third pair contained in the union, the sum of two disjoint pairs is the pair which is disjoint from the union). Then W is an elementary abelian group of order 16. This defines an embedding of S_6 in $GL(4, 2)$. Moreover a symplectic bilinear form is defined by $(w_1, w_2) = 1$ if and only if the w_i are different pairs which intersect in a point of X. This symplectic form is respected by the action of S_6, defining an embedding of S_6 in the symplectic group $Sp(4, 2)$. As these groups have the same order one obtains a second exceptional isomorphism: $S_6 \cong Sp(4, 2)$.

Alternatively we can use the 15 1-factors on X to define an elementary abelian group W', where

$$(12)(34)(56) + (12)(35)(46) = (12)(36)(45),$$

$$(12)(34)(56) + (13)(25)(46) = (16)(24)(35)$$

and a symplectic form where two 1-factors have symplectic product $= 1$ if and only if they do not have a pair in common. This of course leads us back to our old friend, the S_6-GQ.

We can define the famous A_7-geometry:

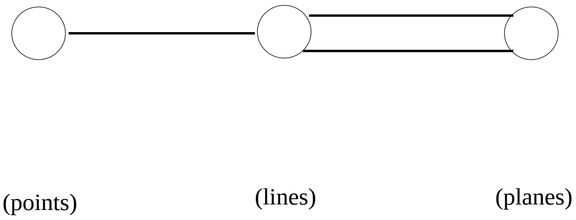

Figure 19.1: The A_7-geometry

19.9 Definition. *The points of our geometry are the* 7 *elements of the ground set* Ω, *the lines are the* $\binom{7}{3} = 35$ *triples from* Ω, *the* 15 *planes are the Fano planes from one* A_7*-orbit* $\mathcal{E}_1$. *Incidence is defined in the natural way.*

19.10 Theorem. *The* A_7*-geometry is a diagram geometry, with diagram as given in Figure 19.1.*

Proof. Fix a plane E. This really is a Fano plane. The residual geometry consists of the points and lines of E. We conclude that the residual geometry of each plane is a projective plane. This is represented by drawing a single line between the nodes representing points and lines.

Fix a line of our geometry. This is a triple T of Ω. The corresponding residual geometry consists of the three elements of Ω on T and of the Fano planes from our orbit having T as a line. All of those objects are incident. The residual geometry of T is a complete bipartite graphs. These are represented by not joining the corresponding nodes of the diagram. In our diagram the nodes representing points and planes are not joined.

Most interesting is the residual geometry corresponding to a fixed element $a \in \Omega$. The points of the residual geometry are the 15 triples containing a, its lines are the 15 Fano planes from orbit $\mathcal{E}_1$. This is our S_6-generalized quadrangle. The graphical representation of a generalized quadrangle is a double line connecting the corresponding nodes. $\square$

The diagram of the A_7-geometry is the same as that for the 7-dimensional orthogonal geometry given in Chapter 18.

Another famous structure is related to the A_7-geometry, the Hoffman-Singleton graph which we denote by Γ. Recall that we constructed a model of $PG(3,2)$ whose 35 lines are the triples from the ground set Ω of size 7.

There are at least 3 types of pairs of lines, according to the size of intersection of the triples (0, 1 or 2) in Ω. Under the full $GL(4,2)$ there are of course only two types of pairs of lines.

The 50 vertices of Γ are the 15 points and the 35 lines of $PG(3,2)$. Two points are never neighbours. A point and a line form an edge if in $PG(3,2)$ the point is on the line. Two lines are neighbours if the corresponding triples are disjoint. Observe that this is **not** equivalent with the corresponding lines of $PG(3,2)$ being parallel.

Each vertex of point type has valency 7 as a point of $PG(3,2)$ is on 7 lines. A vertex of line type has 3 neighbours of point type and $\binom{4}{3} = 4$ neighbours of line type. It follows that Γ is regular of valency 7.

19.11 Theorem. *The Hoffman-Singleton graph* Γ *is a* **Moore graph.** *This means*

- Γ *is regular (of valency* 7*).*
- *There are no triangles in* Γ
- *Any two non-adjacent vertices have precisely one common neighbour.*

Proof. We know the valency, and the absence of triangles is obvious: two disjoint triples cannot be on a common Fano plane. In order to check the final axiom several cases have to be distinguished. Two vertices of point type are on precisely 1 line of $PG(3,2)$. Let T_1, T_2 be triples intersecting in cardinality 1. Then there is precisely one point of $PG(3,2)$ (Fano plane from $\mathcal{E}_1$) containing both as lines. If T_1, T_2 intersect in cardinality 2 there is no such point (Fano plane), but there is exactly one triple disjoint from T_1 and T_2. Finally, let a triple T and a point $P \in PG(3,2)$ be given, where $P \notin T$. In our A_7-language this means we are given a Fano plane E (the projective point) and three points of X forming a triangle in E. The complementary set of 4 points from X contains exactly one line of E. □

We remark that the Moore graphs have more or less been classified. Each Moore graph either is the Petersen graph or the Hoffman-Singleton graph or possibly a certain graph on 3250 vertices, of valency 56 whose existence is in doubt.

From the A_7-geometry to $\mathcal{NR}$.

In the preceding section we constructed a semidirect product $G = VA$, where V is elementary abelian of order 16 and $A \cong A_7$ in its 2-transitive action on V. We write V additively, as a 4-dimensional vector space over $\mathbb{F}_2$, with basis v_1, v_2, v_3, v_4. In terms of Fano planes we can make the following choice:

$$v_1 = x12, x34, x56, 135, 146, 236, 245.$$
$$v_2 = x12, x35, x46, 136, 145, 234, 256.$$
$$v_3 = x34, x15, x26, 124, 136, 235, 456.$$
$$v_4 = 135, 1x4, 126, x25, x36, 234, 456.$$

The normal subgroup V acts by translation on the vectors from V. The action of V is transitive. Let $\tau(v)$ denote translation by $v \in V$. The stabilizer of vector $0 \in V$ is $A \cong A_7$, which is 2-transitive on the 15 nonzero vectors (see Theorem 19.6). It follows that G acts 3-transitively. In particular each orbit of G in its action on subsets of V defines a 3-design. Define the **dimension** of a subset $S \subseteq V$ as the dimension of the affine subspace generated by S. One point has dimension 0, two points have dimension 2 and a 3-point set has dimension 3. A k-set is **in general position** if it has dimension $k-1$.

19.12 Lemma. *G is regular on the ordered 4-tuples in general position.*

Proof. This follows from Theorem 19.6. □

A 4-set in general position is $V_1 = \{0, v_1, v_2, v_3\})$. This defines the orbit $\mathcal{V}_1$ of length $16 \times 15 \times 14 \times 12/24 = 16 \times 15 \times 7$. As G is 3-transitive and each triple is contained in precisely one 4-set of dimension 2 (its elements sum to 0), the 4-sets of dimension 2 form an orbit $\mathcal{V}_2$ of length 140, with representative $V_2 = \{0, v_1, v_2, v_1 + v_2\}$. As $|\mathcal{V}_1| + |\mathcal{V}_2| = \binom{16}{3}$ these are all the orbits on 4-sets. The elements of $\mathcal{V}_2$ form a Steiner quadruple system $S(3, 4, 16)$. It follows from Lemma 19.12 that the stabilizer $G(V_1)$ of V_1 under the action of G is the symmetric group S_4. We have

$$G(V_1) = \langle r, z, \tau(v_1)h\rangle$$

where $r = (x, 1, 3)(2, 6, 4),\ z = (x, 1)(4, 6),\ h = (1, 2)(3, 4)$. The action of these linear operations on V is described by

$$r : v_1 \mapsto v_2 \mapsto v_3 \mapsto v_1,\ v_4 \mapsto v_2 + v_3 + v_4,$$

$$z : v_1 \leftrightarrow v_2,\ v_3 \mapsto v_3,\ v_4 \leftrightarrow v_3 + v_4,$$

$$h : v_1 \mapsto v_1,\ v_2 \leftrightarrow v_1 + v_2,\ v_3 \leftrightarrow v_1 + v_3,\ v_4 \leftrightarrow \sum v_i.$$

The group $G(V_1)$ has two orbits on the set of 8 points complementing V_1 to a 5-set in general position. These orbits are $\{v_2 + v_4, v_1 + v_3 + v_4\}$ and the remaining 6 points. Let

$$H = \{0, v_1, v_2, v_3, v_2 + v_4, v_1 + v_3 + v_4\}$$

and denote by $\mathcal{H}$ the orbit containing H. The 6-sets of this orbit are the **hexads.** Observe that $G(V_1)$ is contained in the stabilizer of H. Another element of $G(H)$ is $\tau(v_2 + v_4)l$, where $l = (x, 6, 1, 2, 4)$. The action of l on V is described by

$$l : v_1 \mapsto v_1 + v_2,\ v_2 \mapsto v_1 + v_3,\ v_3 \mapsto v_1 + v_2 + v_4,\ v_4 \mapsto v_3$$

and $\tau(v_2+v_4)l$ acts on H as a 5-cycle, with fixed point $v_1+v_3+v_4$. The group generated by $G(V_1)$ and $\tau(v_2+v_4)l$ clearly is A_6. It follows $|\mathcal{H}| \leq |G|/|A_6| = 16 \times 7$. All 4-sets contained in H are from $\mathcal{V}_1$. By double counting it follows that we have equality and V_1 is contained in precisely one hexad.

19.13 Proposition. *Let $\mathcal{H}$ be the G-orbit of 6-sets containing $H = \{0, v_1, v_2, v_3, v_2 + v_4, v_1 + v_3 + v_4\}$ (the hexads). There are $112 = 16 \times 7$ hexads, and each 4-set in general position is contained in precisely one hexad.*

Let us describe the completion of a 4-set in general position to its uniquely determined hexad in a combinatorial way. Start from V_1. By Lemma 19.4 the lines of pairwise intersection of v_1, v_2, v_3 form a triangle. This is the triangle $\{x12, x34, 136\}$. Let v be one of the two elements of V (Fano planes) that complement V_1 to a hexad. The vertices of the intersection triangle form a line $x13$ of v. The line through x and 6 contains either 2 or 4, similarly for the other pairs of opposite points in the triangle. Choose $x62$ as a line of v. The same rule shows that 324 and 146 are lines of v, which by now is uniquely determined:

$$v = v_2 + v_4 = \{x13, x62, 324, 146, x45, 356, 125\}.$$

The choice of $x64$ as a line leads to

$$v = v_1 + v_3 + v_4 = \{x13, x64, 142, 326, x25, 156, 345\}$$

These two completion points are the Fano planes $\neq v_1 + v_2 + v_3$ from orbit $\mathcal{E}_1$ containing $x13$ as a line.

Another orbit we are going to need are the **affine hyperplanes.** These are the subgroups of order 8 of V and their complements. There are 30 affine hyperplanes and clearly they form an orbit. We obtain another representation of one of our favorite objects, the Nordstrom-Robinson code $\mathcal{NR}$:

19.14 Definition. *Use the elements of V as coordinates of the vector space $\mathbb{F}_2^{16}$. Identify each vector from $\mathbb{F}_2^{16}$ with its support, a subset of V. The code $\mathcal{NR}$ is the union of the following words:*

- *The 0-word and the 1-word,*
- *The affine hyperplanes,*
- *the hexads and their complements.*

The number of codewords of $\mathcal{NR}$ is $2+30+112+112 = 2^8$. By definition $\mathcal{NR}$ admits G as a group of automorphisms. The constant words and the affine hyperplanes together form a linear subcode. As any two different affine hyperplanes have either 0 or 4 points in common, this is a $[16, 5, 8]_2$-code, the Reed-Muller code $\mathcal{R}$.

19.15 Lemma. *The hexads define a design $3-(16, 6, 4)$. A hexad and an affine hyperplane intersect either in 4 or in 2 points. For each hexad there are 15 affine hyperplanes meeting it in 4 points.*

Proof. The parameters of the design follow from double counting. Let H be a hexad. Each of the 15 4-subsets of H is in exactly one affine hyperplane E, and E meets H in precisely those 4 points. The complements of those 15 affine hyperplanes intersect H in 2 points. □

19.16 Proposition. *$\mathcal{NR}$ is the union of 8 cosets of $\mathcal{R}$. If H is a hexad, then the coset $H + \mathcal{R}$ consists of the images of H under the translation subgroup V and their complements.*

In order to prove Proposition 19.16 it suffices to prove the last claim, which itself is a consequence of the following lemma.

19.17 Lemma. *Let H be a hexad and E an affine hyperplane.*

If $|E \cap H| = 4$, then $H + E = H + a + b$, where a, b are the points of H which are not in E.

If $|E \cap H| = 2$, then $H + E$ is the complement of $H + a + b$, where $E \cap H = \{a, b\}$.

Proof. Considering complements it suffices to prove the first rule. Let $h+a$ or $h=b$. Then $h+a+b \in H \setminus E \subset H+E$. If $h \in H \setminus \{a,b\}$, then $h+a+b \in E$ but $h+a+b \notin H$ as H does not contain a 4-set of dimension 2. It follows $h+a+b \in H+E$ also in this case. As the cardinalities are the same we are done. □

19.18 Definition. *Let $x \in \mathbb{F}_2^{16}$. Denote by $A_i(x)$ the number of codewords of $\mathcal{NR}$ at distance i from x.*

19.19 Theorem. *For each $x \in \mathcal{NR}$ we have*

$$A_0(x) = A_{16}(x) = 1,\ A_6(x) = A_{10}(x) = 112,\ A_8(x) = 30.$$

In particular $\mathcal{NR}$ is a $(16, 2^8, 6)$-code.

Proof. By definition this is true for $x = 0$. Proposition 19.16 shows that it holds for all $x \in \mathcal{R}$ and it suffices to prove the statement for a hexad of our choice. Let H be a hexad. It follows from Lemma 19.15 that the distances from H to the elements of $\mathcal{R}$ are 6 and 10, each occurring 16 times. We need to know how hexads intersect. As each 4-set in general position is in precisely one hexad, the intersection sizes are ≤ 3. As the hexads form a $3-(16,6,4)$, there are precisely $\binom{6}{3} \times 3 = 60$ hexads intersecting H in 3 points. The usual counting arguments show that 15 hexads meet H in 2 points and 36 meet it in precisely one point. As $1+60+15+36 = 112$ this exhausts all hexads. We conclude that precisely 15 hexads are at distance 8 from H, while all other hexads $\neq H$ are at distance 6 or 10. The presence of the all-1-word shows that our claim is true. □

Problems

1. Use the orbit formula from the theory of permutation groups to show that the group of automorphisms of the Fano plane is precisely the group $GL(3,2)$ of order 168.

Bibliography

[1] L.M.Batten: *Combinatorics of finite geometries,* Cambridge University Press 1997.

[2] P.J. Cameron and J.H. van Lint: *Designs, Graphs, Codes and their Links,* Cambridge University Press 1991.

[3] R.W. Carter: *Simple Groups of Lie type,* Wiley 1972.

[4] A. Cayley: *On the triadic arrangements of seven and fifteen things, London, Edinburgh and Dublin Philos. Mag and J.Sci* **37** (1850),50-53.

[5] W.Feit and G.Higman: *The nonexistence of generalized polygons, Journal of Algebra* **1**(1964), 114-131.

[6] J.W.P. Hirschfeld: *Projective geometries over finite fields, second edition,* Oxford Mathematical Monographs, Clarendon Press 1998.

[7] J.W.P. Hirschfeld: *Finite projective spaces of three dimensions,* Clarendon, Oxford 1985.

[8] J.W.P. Hirschfeld and J.A. Thas: *General Galois geometries,* Oxford University Press 1991.

[9] D.R.Hughes and F.C.Piper: *Projective planes,* Springer, New York 1973.

[10] S.E.Payne and J.A. Thas: *Finite generalized quadrangles,* Pitman Press, Boston 1984.

[11] J.P. Serre: *Cours d'arithmétique,* Presses Universitaires de France, 1970.