Preface

This course is an introduction to combinatorial design theory.
# Contents

**Preface** iii

1 Block Designs 1
   1.1 Definitions ........................................... 1
   1.2 Examples ............................................... 1
       1.2.1 Fano Plane ........................................ 1
       1.2.2 Trivial Cases ..................................... 2
       1.2.3 Complements ....................................... 2
       1.2.4 Another difference set construction .......... 2
       1.2.5 Affine Planes .................................... 2
   1.3 Relations between Parameters ........................ 2

2 Symmetric Designs 5
   2.1 Incidence Matrices of Designs ........................ 5
   2.2 Constructing Symmetric Designs ....................... 6
   2.3 Two Open Problems ................................... 8
   2.4 Bilinear Forms ....................................... 8
   2.5 Cancellation ......................................... 10
   2.6 The Bruck-Ryser-Chowla Theorem ..................... 11
   2.7 Applications ......................................... 14

3 Hadamard Matrices 17
   3.1 A Lower Bound ........................................ 17
   3.2 Equality .............................................. 18
   3.3 The Kronecker Product ................................ 19
   3.4 Symmetric and Regular Hadamard Matrices .......... 21
       3.4.1 Examples ........................................ 21
   3.5 Conference Matrices .................................. 22
   3.6 Type-II Matrices ..................................... 24

4 Planes 27
   4.1 Projective Planes ...................................... 27
   4.2 Affine Planes ......................................... 29
   4.3 Möbius Planes ......................................... 30
CONTENTS

5 Orthogonal Arrays 33
  5.1 Latin Squares ................................................. 33
  5.2 Examples ....................................................... 34
  5.3 Affine Planes .................................................. 34
  5.4 Partial Geometries ............................................ 35
  5.5 Strongly Regular Graphs .................................... 37

6 Block Intersections 39
  6.1 Quasi-Symmetric Designs .................................... 39
  6.2 Triangle-free Strongly Regular Graphs .................... 40
  6.3 Resolvable Designs ........................................... 42
  6.4 Designs with Maximal Width .................................. 44

7 $t$-Designs 47
  7.1 Basics .......................................................... 47
  7.2 Extending Fisher's Inequality ............................... 48
  7.3 Intersection Triangles ...................................... 49
  7.4 Complements and Incidence Matrices ....................... 50
  7.5 Polynomials .................................................... 51
  7.6 Gegenbauer Polynomials ...................................... 52
  7.7 A Positive Semidefinite Matrix ............................ 53

8 Witt Designs 55
  8.1 Codes ........................................................... 55
  8.2 Perfect Codes .................................................. 56
  8.3 The Binary Golay Code ....................................... 57
  8.4 The Plane of Order Four ...................................... 58
  8.5 The Code of $PG(2,4)$ .......................................... 59

9 Groups and Matrices 61
  9.1 Group Matrices ............................................... 61
  9.2 Eigenvalues and Eigenvectors ................................ 62
  9.3 Multipliers ..................................................... 63

10 Mutually Unbiased Bases 67
  10.1 Complex Lines and Angles .................................. 67
  10.2 Equiangular Lines ............................................ 68
  10.3 Reality .......................................................... 69
  10.4 Difference Sets and Equiangular Lines ................. 70
  10.5 MUB's .......................................................... 71
  10.6 Real MUB's ..................................................... 73
  10.7 Affine Planes .................................................. 74
  10.8 Products ........................................................ 75

11 Association Schemes 77
  11.1 Intersection Matrices ........................................ 77
Chapter 1

Block Designs

1.1 Definitions

We use the following definition for this lecture. (For the official definition, see Corollary ??.) A block design $D$ consists of a point set $V$ and a set $B$ of blocks where each block is a $k$-subset of $V$ such that:

(a) There is a constant $r$ such that each point lies in exactly $r$ blocks.

(b) There is a constant $\lambda$ such that each pair of distinct points lies in exactly $\lambda$ blocks.

In fact (b) implies (a), which we leave as an exercise. What we call a block design might also be called a 2-design.

1.2 Examples

We offer some simple examples of designs.

1.2.1 Fano Plane

Here we have $V = \mathbb{Z}_7$ and the blocks are as follows:

\[
\begin{align*}
\{0, 1, 3\} \\
\{1, 2, 4\} \\
\{2, 3, 5\} \\
\{3, 4, 6\} \\
\{4, 5, 0\} \\
\{5, 6, 1\} \\
\{6, 0, 2\}
\end{align*}
\]
This is a block design with parameters \((v, b, r, k, \lambda) = (7, 7, 3, 3, 1)\).

In general \(b\) is very large and it may be inconvenient or impossible to present the design by listing its blocks. The Fano plane is an example of a difference set construction. A difference set \(S\) in an abelian group \(G\) is a subset of \(G\) with the property that each non-zero element of \(G\) appears the same number of times as a difference of two elements of \(S\). Here \(\alpha = \{0, 1, 3\}\) is a difference set for \(G = \mathbb{Z}_7\). If \(G\) is an abelian group and \(S \subseteq G\) then the set

\[
S + g = \{x + g \mid x \in S\}
\]

is called a translate of \(S\). In our example, the design consists of all translates of \(\alpha\). The Fano plane is the projective plane of order two.

### 1.2.2 Trivial Cases

A design is trivial if \(k \in \{0, 1, v - 1, v\}\). These are valid designs, although not so interesting. For a non-trivial design, \(2 \leq k \leq v - 2\). The complete design consists of all \(k\)-subsets of a set of size \(v\).

### 1.2.3 Complements

If we take the complement in \(\mathbb{Z}_7\) of each block of the Fano plane we get a design on 7 points with block size 4. This holds in general, i.e. if we take the complement of each block in a design we obtain another design.

### 1.2.4 Another difference set construction

Let \(V = \mathbb{Z}_{11}\), then \(\alpha = \{0, 2, 3, 4, 8\}\) is a difference set and the set of all translates of \(\alpha\) is a 2-design with parameters \((v, b, r, k, \lambda) = (11, 11, 5, 5, 2)\). A design with \(b = v\) is called a symmetric design. We see later on that \(k = r\) in symmetric designs.

### 1.2.5 Affine Planes

Let \(V\) be a vector space and let \(B\) be the set of lines (a line is a coset of a 1-dimensional subspace). This is a 2-design with \(\lambda = 1\). If we take \(V\) to be the 2-dimensional vector space over \(\mathbb{Z}_3\) we get a 2-(9,3,1) design with \(b = 12\) and \(r = 4\). So

\[
V = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \mid a, b \in \mathbb{Z}_3 \right\}.
\]

This is an affine plane of order three.

### 1.3 Relations between Parameters

The parameters \(v, b, r, k, \lambda\) are not independent. Define a flag to be an ordered pair \((v, \alpha)\) where \(v \in V\) and \(\alpha \in B\). Then by counting with respect to the first
or second coordinate we obtain $r|V| r|V|$ and $k|B|$ respectively as the number of blocks. Thus

$$vr = kb \quad \text{or} \quad \frac{v}{k} = \frac{b}{r}.$$  

Now take ordered triples $(v, w, \alpha)$ where $v \neq w$ and $v, w \in \alpha$. If we count the number of such triples with respect to the first two coordinates or with respect to the third coordinate we obtain $v(v - 1)\lambda$ and $bk(k - 1)$ respectively. Thus

$$v(v - 1)\lambda = bk(k - 1)$$

or equivalently

$$\frac{v(v - 1)}{k(k - 1)} = \frac{b}{\lambda}.$$ 

Also, since $bk = vr$, we have $v(v - 1)\lambda = vr(k - 1)$. Therefore

$$\frac{v - 1}{k - 1} = \frac{r}{\lambda}.$$ 

We see that the ratios $b/\lambda$ and $r/\lambda$ are determined by $v$ and $k$—this is worth remembering.

**Example.** Suppose $\lambda = 1$ and $k = 3$. Then $b = v(v-1)/6$ and $r = (v-1)/2$. Thus

$$v \equiv 1, 3 \mod 6.$$ 

So the possible parameter sets with $\lambda = 1$, $k = 3$ and $v \leq 15$ are

$$(3, 3, 1), (7, 3, 1), (9, 3, 1), (13, 3, 1), (15, 3, 1)$$

and so on. Note that for a 2-(13,3,1) design we get $b = 26$ and so the number of blocks must be twice the number of points. This observation suggests trying to construct such a design by using a difference set with two initial blocks.

Designs with $\lambda = 1$ and $k = 3$ are called **Steiner triple systems**.
Chapter 2
Symmetric Designs

2.1 Incidence Matrices of Designs

For a \((v, b, r, k, \lambda)\) design there are \(\lambda = \lambda_2\) blocks on each point of points. This gives

\[
\frac{v(v - 1)}{k(k - 1)} = \frac{b}{\lambda}, \quad \frac{v - 1}{k - 1} = \frac{r}{\lambda}.
\]

Our design has an incidence matrix \(N\) with the following properties:

(a) \(N1 = r1\)
(b) \(1^TN = k1T\)
(c) \(NN^T = (r - \lambda)I + \lambda J\)

where \(J\) is the matrix with all entries equal to one. These equations hold if and only if the corresponding instance structure is a 2-design.

2.1.1 Theorem. If \(\mathcal{D}\) is a 2-design with parameters \((v, b, r, k, \lambda)\) and \(\mathcal{D}\) has at least 2 points and at least 2 blocks, then \(b \geq v\).

Proof. We aim to show that the rows of \(N\) are linearly independent over \(\mathbb{R}\) and, since we are working over the reals, can do this by proving that \(NN^T\) is invertible. Now

\[
NN^T = (r - \lambda)I + \lambda J
\]

and we can write down the inverse explicitly. The key is to note that

\[
(xI + J)(yI + J) = xyI + (x + y)J + vJ = xyI + (x + y + v)J,
\]

from which it follows that \(xI + J\) is invertible if \(x \neq 0\). \(\square\)

The inequality \(b \geq v\) is called Fisher’s inequality. Can we have \(b = v\)? We’ve already seen two examples: the 2-(7, 3, 1) and 2-(11, 5, 2) designs. A design with \(v = b\) is called a symmetric design. Note that \(b = v\) if and only if \(r = k\).
2.1.2 Theorem. If $D$ is a symmetric design with parameters $(v, k, \lambda)$, then any two distinct blocks have exactly $\lambda$ points in common.

Proof. Since $b = v$, the incidence matrix $N$ is invertible. We have

$$NN^T = (r - \lambda)I + \lambda J$$

The rows and columns of $N^TN$ are indexed by the blocks of the design; the $\alpha, \beta$ entry of $N^TN$ is the size of $|\alpha \cap \beta|$. We want to show that $|\alpha \cap \beta|$ is constant for $\alpha \neq \beta$ and is equal to $k$ when $\alpha = \beta$.

$$N^TN = N^{-1} (NN^T) N$$
$$= N^{-1} ((r - \lambda)I + \lambda J) N$$
$$= (r - \lambda)I + \lambda N^{-1}JN$$

Note that since $N1 = k1$, we have

$$\frac{1}{k}1 = N^{-1}1.$$ It follows that

$$\lambda N^{-1}JN = \lambda N^{-1}11^TN = \frac{1}{k}k11^T = \lambda J$$

and hence $N^TN = (r - \lambda)I + \lambda J$. \hfill $\Box$

Remarks: If $b = v$, we have shown that $N^TN = NN^T$, i.e., the incidence matrix of a symmetric design is normal. In terms of incidence structures, we see that the dual of a symmetric design is symmetric. (In general the dual of a block design is not a block design.)

2.2 Constructing Symmetric Designs

Suppose that we have a symmetric design with parameter set $(v, k, \lambda)$. We know that

$$\frac{v(v - 1)}{k(k - 1)} = \frac{b}{\lambda} = \frac{v}{\lambda}$$

so that $v = 1 + \frac{k^2 - k}{\lambda}$. For example if we let $\lambda = 1$, then the following table lists the possible parameter sets when $v \leq 9$.

| $k$ | $v$ | This is the trivial design
|-----|-----|---------------------------
| 2   | 3   | This is the trivial design
| 3   | 7   | Does not exist
| 4   | 13  | Does not exist
| 5   | 21  | Does not exist
| 6   | 31  | Does not exist
| *7  | 43  | Does not exist
| 8   | 57  | Does not exist
| 9   | 73  | Does not exist
2.2. CONSTRUCTING SYMMETRIC DESIGNS

The fact that there is no symmetric \((43, 7, 1)\)-design follows from the Bruck-Ryser-Chowla theorem, which we will meet before long. We now present a construction for a class of symmetric designs using subspaces of vector spaces. This class will include all the designs that do exist in the above table.

Let \(V\) be a vector space of dimension \(d\) over \(GF(q)\), where in all interesting cases \(d \geq 3\). We build an incidence structure as follows: The points are the 1-dimensional subspaces of \(V\) and the blocks are the subspaces with dimension \(d - 1\). Incidence is determined by inclusion (symmetrized). We call the \(d - 1\) dimensional subspaces hyperplanes. The number of elements of \(V\) is \(q^d\), and the number of elements in a 1-dimensional subspace is \(q\). The 1-dimensional subspaces partition \(V\) into sets of size \(q - 1\). So there are \(\frac{q^d - 1}{q - 1}\) points in our incidence structure.

Each hyperplane is the kernel of a \(1 \times d\) matrix \([a_1, ..., a_d] = a\). If \(a\) and \(b\) are two non-zero vectors of length \(d\), then \(\ker(a) = \ker(b)\) if and only if \(b\) is a non-zero, scalar multiple of \(a\). It follows that the number of hyperplanes is equal to the number of 1-dimensional subspaces, that is, \(\frac{q^d - 1}{q - 1}\).

The non-zero vectors \([a_1, ..., a_d]\) and \([b_1, ..., b_d]\) determine distinct hyperplanes if and only if the matrix

\[
\begin{pmatrix}
    a_1 & \cdots & a_d \\
    b_1 & \cdots & b_d \\
\end{pmatrix}
\]

has rank two. The kernel of this matrix is a subspace of dimension \(d - 2\) and the 1-dimensional subspaces are the 1-dimensional subspace that lie on both hyperplanes.

This allows us to conclude that the subspaces of dimension one and \(d - 1\) are the points and blocks of a symmetric design with parameters

\[
2 - \left( \frac{q^d - 1}{q - 1}, \frac{q^{d-1} - 1}{q - 1}, \frac{q^{d-2} - 1}{q - 1} \right).
\]

Now, suppose that \(d = 3\), then the designs constructed in this way have parameter sets as follows for the first few choices of \(q\).

<table>
<thead>
<tr>
<th>(q)</th>
<th>(v)</th>
<th>(k)</th>
<th>(\lambda)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>7</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>31</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>*6</td>
<td>43</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>57</td>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>

For \(d = 3\) we have infinitely many symmetric designs with \(\lambda = 1\) and block size \(q + 1\), where \(q\) is a prime power.
2.3 Two Open Problems

**Question:** Is there a symmetric design with $\lambda = 1$ where $k - 1$ is not a prime power?

A projective plane is a thick incidence structure such that

(a) each pair of distinct points lies on exactly one line.

(b) each pair of distinct lines has a common point.

A finite incidence structure is a projective plane if and only if it is a symmetric design with $\lambda = 1$. It is straightforward to show that the design is a projective plane, the converse is left as an exercise.

The theory of projective planes is rich and complex. We know comparatively little about the symmetric designs where $\lambda \geq 2$. Thus we do not know the answer to the following.

**Question:** Do there exist infinitely many symmetric designs for fixed $\lambda \geq 2$?

Note that the complement of a projective plane is a symmetric design with $\lambda > 1$.

2.4 Bilinear Forms

We introduce some of the theory of bilinear and quadratic forms; we will apply this to develop a very important necessary condition for the existence of symmetric designs.

Let $V$ be a vector space over a field $\mathbb{F}$. A bilinear form $\beta$ on $V$ is a function from $V \times V$ to $\mathbb{F}$ that is linear in both variables. If

$$\beta(u, v) = \beta(v, u)$$

for all $u$ and $v$ we say that $\beta$ is symmetric. The canonical example is take a symmetric matrix $B$ and define

$$\beta(u, v) = u^T B v$$

If $\beta$ is a bilinear from on $V$ and $v \in V$, then we define

$$v^\perp := \{x : \beta(v, x) = 0\}.$$  

If $U \subseteq V$, then

$$U^\perp := \cap_{u \in U} u^\perp.$$  

It is possible that $v \in v^\perp$ and $v \neq 0$, but we do have that

$$\dim(U^\perp) = \dim(V) - \dim(U)$$

and $(U^\perp)^\perp = U$. We see that if $v \neq 0$ then $v^\perp$ is a subspace of $V$ with codimension at most 1. We say $\beta$ is non-degenerate if $v^\perp = V$ imples $v = 0$.

A quadratic form $Q(v)$ over $V$ is function from $V$ to $\mathbb{F}$ such that
1. $Q(au) = a^2 Q(u)$ for all $a$ in $F$ and $u \in V$.

2. $Q(u + v) - Q(u) - Q(v)$ is a symmetric bilinear form on $V$.

For example, if $\beta$ is a symmetric bilinear form then
\[
\beta(x + y, x + y) = \beta(x, x) + \beta(x, y) + \beta(y, x) + \beta(y, y) = \beta(x, x) + \beta(y, y) + 2\beta(x, y)
\]
and so $Q_\beta(x) := \beta(x, x)$ is a quadratic form. If 2 is invertible in $F$, then the quadratic form determines the bilinear form.

We assume now that 2 is an invertible element of our underlying field. Hence any quadratic form $Q$ we discuss can be written in the form (sorry)
\[
Q(x) = x^T Ax
\]
where $A$ is symmetric.

Two quadratic forms $Q_1$ and $Q_2$ in $d$ variables over $F$ are equivalent if there is an invertible matrix $G$ such that, for all $x$ and $y$ in $V$
\[
Q_2(x) = Q_1(Gx).
\]
It is easy to check that this is, as the name suggests, an equivalence relation.

Two symmetric matrices $A$ and $B$ are congruent if there is an invertible matrix $G$ over $F$ such that
\[
B = G^T AG.
\]
We write $A \approx B$ to denote that $A$ and $B$ are congruent. If two quadratic forms are equivalent then their associated bilinear forms are congruent.

Note that if $P$ is a permutation matrix, then $A$ and $P^T AP$ are equivalent, hence
\[
\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \approx \begin{pmatrix} b & 0 \\ 0 & a \end{pmatrix}.
\]
If $c \neq 0$, then
\[
\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \approx \begin{pmatrix} c^2 a & 0 \\ 0 & b \end{pmatrix}.
\]

Our next result provides the connection with design theory: it shows that existence of a symmetric design implies the equivalence of two diagonal matrices.

### 2.4.1 Theorem.
If there is a symmetric $(v, k, \lambda)$ design, then
\[
\begin{pmatrix} I & 0 \\ 0 & -\lambda \end{pmatrix} \approx (k - \lambda) \begin{pmatrix} I & 0 \\ 0 & -\lambda \end{pmatrix}
\]

**Proof.** Assume $N$ is the incidence matrix of a symmetric $(v, k, \lambda)$ design. If
\[
\hat{N} := \begin{pmatrix} N \\ \lambda I^T \\ 1 \\ k \end{pmatrix}
\]
then a direct calculation shows that
\[
\hat{N} \begin{pmatrix} I & 0 \\ 0 & -\lambda \end{pmatrix} \hat{N}^T \approx (k - \lambda) \begin{pmatrix} I & 0 \\ 0 & -\lambda \end{pmatrix}.
\]
To deduce equivalence, we should verify that $\hat{N}$ is invertible. You do it. \qed
2.5 Cancellation

A quadratic space is a pair \((V, q)\), where \(V\) is a vector space and \(q\) is a quadratic form on \(V\). Two quadratic spaces \((V_1, q_1)\) and \((V_2, q_2)\) isometric if there is an invertible linear map \(L : V_1 \to V_2\) such that
\[
q_2(Lv) = q_1(v)
\]
for all \(v\). If \(U\) is a subspace of \(V\) then \(U\) along with the restriction \(q\mid_U\) is a quadratic space, a subspace. Next suppose that \(V\) is the direct sum of \(U_1\) and \(U_2\) and \(q_1\) and \(q_2\) are quadratic forms on \(U_1\) and \(U_2\) respectively. If we define
\[
q(u_1, u_2) := q_1(u_1) + q_2(u_2)
\]
then \(q\) is a quadratic form on \(V\). We say that the quadratic space \((V, q)\) is the sum of the spaces \((U_1, q_1)\) and \((U_2, q_2)\).

If \(U\) is a subspace of the quadratic space \(V\), then the radical of \(U\) is the subspace \(U \cap U^\perp\). If the radical of \(U\) is zero then \(U + U^\perp = V\) and \(V\) is the direct sum of \(U\) and \(U^\perp\).

2.5.1 Lemma. Let \((V, q)\) be a quadratic space and suppose \(u\) and \(v\) are elements of \(V\) such that \(q(u) = q(v) \neq 0\). Then there is an isometry of \(V\) that maps \(u\) to \(v\).

Proof. First we define a class of isometries. Let \(\beta\) be the symmetric bilinear form associated with \(q\). If \(a \in V\) and \(q(a) \neq 0\), define the map \(\tau_a\) on \(V\) by
\[
\tau_a(v) := v - 2\frac{\beta(a, v)}{q(a)} a.
\]
Then \(\tau_a\) is linear and \(\tau_a^2\) is the identity. You may also check that \(q(\tau_a(v)) = q(v)\) for all \(v\); whence \(\tau_a\) is an isometry.

Your next exercise is to show that if \(q(u - v) \neq 0\), then \(\tau_{u-v}\) swaps \(u\) and \(v\).

Now suppose that \(q(u) = q(v) \neq 0\). If \(q(u - v) \neq 0\) and \(a = u - v\), then \(\tau(a)\) swaps \(u\) and \(v\). If \(q(u + v) \neq 0\) and \(b = u + v\), then \(\tau(b)\) swaps \(u\) and \(v\), and therefore \(-\tau_{u+b}\) swaps \(u\) and \(v\). If the worst happens and
\[
q(u - v) = q(u + v) = 0
\]
then
\[
0 = q(u - v) + q(u + v) = 2q(u) + 2q(v) = 4q(u) \neq 0.
\]
Hence the lemma is proved.

2.5.2 Theorem. Suppose \(U_1\) and \(U_2\) are non-zero subspaces of the quadratic space \((V, q)\) and the radical of \(U_1\) is zero. Then if there is an isometry \(\rho : U_1 \to U_2\), there is an isometry from \(V\) to itself whose restriction to \(U_1\) is equal to \(\rho\).
2.6. THE BRUCK-RYSER-CHOWLA THEOREM

Proof. If \( q \) vanishes on \( V \) then the radical of \( U \) is \( U \), so there is a vector \( u \) in \( U_1 \) such that \( q(u) \neq 0 \). By the previous lemma, there is an isometry \( \sigma \) on \( V \) such that \( \sigma(\rho(u)) = u \). If \( \dim(U_1) = 1 \), we are done.

So \( \sigma \rho \) is an isometry from \( U_1 \) to \( \sigma(U_2) \) that fixes \( u \). If \( \sigma \rho \) extends to an isometry \( \tau \) (say) of \( V \), then \( \tau \) followed by \( \sigma^{-1} \) is an isometry of \( V \) that extends \( \rho \).

We proceed by induction on \( \dim(U_1) \). Now \( U_1 \) is the sum of the span of \( v \) and the space \( v^\perp \cap U_1 \), which is a complement to \( v \) in \( U_1 \). Since \( \sigma \rho \) is an isometry, \( \sigma \rho(v^\perp \cap U_1) \) is a complement to \( v \) in \( \sigma(U_2) \). By induction, there is an isometry \( \phi \) on \( v^\perp \) that coincides with \( \sigma \rho \) on \( v^\perp \cap U_1 \).

The linear map that fixes \( v \) and agrees with \( \phi \) on \( v^\perp \cap U_1 \) is an isometry of \( V \) that agrees with \( \sigma \rho \) on \( U_1 \). \( \square \)

The following corollary is a form of Witt cancellation.

2.5.3 Corollary. Let \((V_1, q_1)\) and \((V_2, q_2)\) be isometric quadratic spaces. Let \( U_1 \) and \( U_2 \) be subspaces of \( V_1 \) and \( V_2 \) respectively. If the radical of \( U_1 \) is zero and \( U_1 \) is isometric to \( U_2 \), then \( U_1^\perp \) and \( U_2^\perp \) are isometric.

Proof. If \( \sigma \) is an isometry from \( V_2 \) to \( V_1 \), then \( \sigma(U_2) \) is a subspace of \( V_1 \) isometric to \( U_1 \). So we may assume that \( V_2 = V_1 \) and \( q_1 = q_2 \). Now the theorem yields that there is an isometry \( \sigma \) such that \( \sigma(U_1) = U_2 \), and so \( \sigma(U_1)^\perp = U_2^\perp \). As \( \sigma(U_1)^\perp = \sigma(U_2)^\perp \), we are done. \( \square \)

If the radical of \((V, q)\) is zero, we say that \( q \) is non-singular. If \( 2 \) is invertible and \( q(x) = x^T A x \) for a symmetric matrix \( A \), then \( q \) is non-singular is \( A \) is.

Suppose \( f_1 \) and \( f_2 \) are equivalent non-singular quadratic forms in \( x_1, \ldots, x_m \) and \( g_1 \) and \( g_2 \) are quadratic forms in a disjoint set of variables \( y_1, \ldots, y_n \). Then if \( f_1 + g_1 \) and \( f_2 + g_2 \) are equivalent (as forms in \( m + n \) variables), the previous corollary implies that \( g_1 \) and \( g_2 \) are equivalent. This is the form in which we will use Witt cancellation.

2.6 The Bruck-Ryser-Chowla Theorem

We need one preliminary result, due to Lagrange.

2.6.1 Theorem. If \( n \) is a positive integer then \( I_4 \approx n I_4 \).

Proof. Define the matrix

\[
G := \begin{pmatrix}
a & -b & -c & -d \\
b & a & -d & c \\
c & d & a & -b \\
d & -c & b & a
\end{pmatrix}
\]

and verify that

\[
G^T G = (a^2 + b^2 + c^2 + d^2) I_4.
\]

By a famous theorem due to Lagrange, every positive integer is equal to the sum of four squares and so the theorem follows. \( \square \)
To place this result in some context, we observe that
\[
\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}
\]
This implies that if \( I_2 \approx nI_2 \), then \( n \) must be the sum of two squares. Thus \( I_2 \) is not equivalent to \( 3I_2 \). (It is not too hard to show that \( I_2 \) and \( nI_2 \) are equivalent if and only if \( n \) is the sum of two squares.)

The following necessary conditions for the existence of a symmetric design are due to Bruck, Rysler, and Chowla. We call \( r - \lambda \) the order of a \((v,k,\lambda)\)-design, and we denote it by \( n \).

2.6.2 Theorem. If there is a nontrivial symmetric \((v,k,\lambda)\) design, one of the following two cases hold:

(a) If \( v \) is even, \( k - \lambda \) is a square.

(b) If \( v \) is odd, then the equation \((k - \lambda)x^2 + (-1)^{\frac{v+1}{2}}\lambda y^2 = z^2\) has a nonzero integer solution.

Proof. First suppose \( v \) is even. Recall that
\[
\hat{N} \begin{pmatrix} I & 0 \\ 0 & -\lambda \end{pmatrix} \hat{N}^T = (k - \lambda) \begin{pmatrix} I & 0 \\ 0 & -\lambda \end{pmatrix}
\]
Take determinants of both sides of this to get
\[
(\det \hat{N})^2(-\lambda) = (k - \lambda)^{(v+1)}(-\lambda)
\]
From this we see that \((k - \lambda)^{(v+1)}\) is a square. This implies \((k - \lambda)\) is a square.

Now suppose that \( v \) is odd. There are two sub-cases to consider. If \( v \) is congruent to 1 modulo 4, we use Witt cancellation and the fact that \( I_4 \approx (k - \lambda)I_4 \) to cancel as many leading \( 4 \times 4 \) blocks as possible, deducing as a result that
\[
\begin{pmatrix} 1 & 0 \\ 0 & -\lambda \end{pmatrix} \approx \begin{pmatrix} n & 0 \\ 0 & -n\lambda \end{pmatrix}
\]
Since these forms are equivalent, they take the same integer values. Furthermore, since
\[
\begin{pmatrix} 1 & 0 \\ 0 & -\lambda \end{pmatrix} \begin{pmatrix} n & 0 \\ 0 & -n\lambda \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = n
\]
the right hand form takes the value \( n \). Thus there are \( u \) and \( v \) such that
\[
n = \begin{pmatrix} u \\ v \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -\lambda \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} \begin{pmatrix} u \\ -\lambda v \end{pmatrix} = u^2 - \lambda v^2
Here \( u \) and \( v \) are rational, and so by clearing denominators we get

\[
x_n^2 = y^2 - \lambda z^2, \quad x, y, z \in \mathbb{Z}.
\]

This gives us the desired equation.

The second subcase is when \( v \) is congruent to 3 modulo 4. We have

\[
x_1^2 + \cdots + x_v^2 - \lambda x_0^2 \approx ny_1^2 + \cdots + ny_v^2 - n\lambda y_0^2
\]

and therefore

\[
x_1^2 + \cdots + x_v^2 - \lambda x_0^2 + ua_1^2 + na_2^2 \approx ny_1^2 + \cdots + ny_v^2 - n\lambda y_0^2 + b_1^2 + nb_2^2.
\]

Consequently

\[
(\lambda x_0^2 + na_2^2) \approx (ny_v^2 + \cdots + ny_0^2 + b_1^2) - n\lambda y_0^2 + b_1^2
\]

Since \( v \equiv 3 \) modulo 4, we can cancel the terms in parentheses, and deduce that

\[
\begin{pmatrix}
\lambda & 0 \\
0 & n
\end{pmatrix} \approx \begin{pmatrix}
1 & 0 \\
0 & -\lambda n
\end{pmatrix}
\]

Thus we have

\[
n = (0 \ 1) \begin{pmatrix}
\lambda & 0 \\
0 & n
\end{pmatrix} (0 \ 1)
\]

and so there are \( n \) and \( v \) such that

\[
n = \begin{pmatrix}
u \\
v
\end{pmatrix} \begin{pmatrix}1 & 0 \\
0 & -\lambda n
\end{pmatrix} \begin{pmatrix}u \\
v
\end{pmatrix}
\]

\[
= (u \ v) \begin{pmatrix}u \\
-\lambda nv
\end{pmatrix}
\]

\[
= u^2 - \lambda nv^2.
\]

From this we see that

\[
n + n\lambda v^2 = u^2
\]

and multiplying both sides by \( n \) we have

\[
n^2 + \lambda (nv)^2 = nu^2.
\]

Setting \( v_1 = nv \) we get

\[
n^2 + \lambda v_1^2 = nu^2
\]

and clearing denominators gives

\[
n^2 z_1^2 = nx^2 - \lambda y^2.
\]

Finally by defining \( z_1^2 = n^2 z^2 \) we have

\[
z_1^2 = nx^2 - \lambda y^2.
\]
2.7 Applications

We apply the Bruck-Ryser-Chowla conditions to projective planes. These are 2-designs with \( \lambda = 1 \) and \( v = n^2 + n + 1 \). There are two cases.

If \( n \equiv 0, 3 \mod 4 \), then \( v \equiv 1 \mod 4 \) and \( nx^2 + y^2 = z^2 \). Here \((x, y, z) = (0, 1, 1)\) is a non-zero integer solution.

If \( n \equiv 1, 2 \mod 4 \), then \( v \equiv 1 \mod 4 \) and \( nx^2 = y^2 + z^2 \). Thus we have

\[
n = \left(\frac{y}{x}\right)^2 + \left(\frac{z}{x}\right)^2
\]

which implies that \( n = a^2 + b^2 \) for some integers \( a \) and \( b \). Therefore if \( n \equiv 1, 2 \mod 4 \), then \( n \) must be the sum of the two squares.

2.7.1 Corollary. There is no projective plane of order 6.

Due a difficult computation by Clement Lam from Concordia, we know that there is no projective plane of order 10, even though the conditions in the BRC Theorem are satisfied. However this is the only case we know where the BRC conditions for the existence of a symmetric design are satisfied, but the design does not exist.

Consider the problem of finding a non-zero solution to the equation

\[
Ax^2 + By^2 + Cz^2 = 0
\]

(where \( A, B, \) and \( C \) are integers). Assume that each pair from \( \{A, B, C\} \) is coprime. Then necessary conditions for the existence of a non-zero solution are:

(a) \( A, B, C \) do not all have the same sign

(b) If the odd prime \( p \) divides \( A \), then \(-BC\) is a square modulo \( p \).

(c) If the odd prime \( p \) divides \( B \), then \(-AC\) is a square modulo \( p \).

(d) If the odd prime \( p \) divides \( C \), then \(-AB\) is a square modulo \( p \).

For example iff \( p | A \), then

\[
By^2 + Cz^2 = 0 \mod p
\]

and therefore

\[
B^2y^2 + BCz^2 = 0 \mod p,
\]

from which it follows that \(-BC\) must be a square modulo \( p \). Legendre proved that the above four conditions are sufficient as well.

As examples, consider symmetric designs where \( \lambda = 2 \). Here

\[
v = \frac{k(k-1)}{2} + 1 = \binom{k}{2} + 1.
\]
Consider $k = 8$ (so $(v, k, \lambda) = (29, 8, 2)$). Then the BRC equation is

$$6x^2 + 2y^2 = z^2$$

If there is a non-zero solution, then $2 | z$. If $z = 2z_1$, then

$$6x^2 + 2y^2 - 4z_1^2 = 0$$

and so

$$3x^2 + y^2 - 2z_1^2 = 0.$$ 

Here $-BC = 2$, which is not a square modulo $A = 3$. 

\begin{tabular}{ccc}
\hline
k & v & n & Reason for nonexistence \\
\hline
4 & 7 & 2 & \\
5 & 11 & 3 & \\
6 & 16 & 4 & \\
7 & 22 & 5 & $n \neq \text{square}$ \\
8 & 29 & 6 & BRC \\
9 & 37 & 7 & $n \neq \text{square}$ \\
10 & 46 & 8 & \\
11 & 56 & 9 & \\
12 & 67 & 10 & \\
13 & 79 & 11 & \\
\hline
\end{tabular}
Chapter 3

Hadamard Matrices

A Hadamard matrix is an $n \times n$ matrix $H$ with entries $\pm 1$, such that

$$HH^T = nI.$$ 

For example

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}.$$ 

We will meet many examples as we continue.

The are two families of operations we can apply that take a Hadamard matrix to a Hadamard matrix:

(a) Permute the rows and/or columns.

(b) Multiply all entries in a row or column by $-1$.

An arbitrary combination of these operations is a monomial operation. we say two Hadamard matrices are monomially equivalent if we can one from the pother by monomial operations. A monomial matrix is a product of a permuting matrix and a diagonal matrix with diagonal entries equal to $\pm$; if $M_1$ and $M_2$ are monomial matrices and $H$ is Hadamard, then $M_1HM_2$ is Hadamard and is monomially equivalent to $H$.

A Hadamard matrix is normalized if all entries in its first row and column are equal to 1. Note that the equivalence class of $H$ will contain many different normalized Hadamard matrices.

3.1 A Lower Bound

If $D$ is a symmetric $(v, k, \lambda)$ design, we define the difference $k - \lambda$ to be the order of $D$, and we usually denote it by $n$. (This is consistent with the order of a finite projective plane.)
3.1.1 Theorem. If $\mathcal{D}$ is a symmetric $(v, k, \lambda)$ design, then
\[ 4n - 1 \leq v \leq n^2 + n + 1. \]

Proof. Our concern here is with the lower bound, so we leave the upper bound as an exercise.

Let $N$ be the incidence matrix of $\mathcal{D}$. Then each entry of $2N - J$ is equal to $\pm 1$ and $(2N - J)\mathbf{1} = (2k - v)\mathbf{1}$. Consequently
\[
1^T (2N - J)^T (2N - J) \mathbf{1} = v(v - 2k)^2.
\]

One the other hand
\[
1^T (2N - J)^T (2N - J) \mathbf{1} = 1^T (4NN^T - 2NJ - 2JN^T + J^2) \mathbf{1}
= 1^T (4nI + 4\lambda J - 2kJ - 2kJ + vJ) \mathbf{1}
= 1^T (4nI - 4nJ + vJ)
= 4nv + v^2(v - 4n).
\]

Consequently
\[
v(v - 4n) + 4n = (v - 2k)^2 \geq 0
\]
and therefore
\[
v^2 - 4nv + 4n^2 + 4n \geq 4n^2,
\]
which yields that
\[
(v - 2n)^2 \geq 4n^2 - 4n.
\]
If $n > 0$ then $n^2 - n$ is not square and thus we have slightly stronger inequality
\[
(v - 2n)^2 \geq (2n - 1)^2.
\]
This proves the lower bound. \(\square\)

3.2 Equality

We decide what happens when a symmetric design has $v = 4n - 1$. Let $\overline{N}$ denote the matrix $2N - J$. Then
\[
\overline{N} \overline{N}^T = 4nI - J
\]
and therefore
\[
(1 - \overline{N}) (1 - \overline{N})^T = 4nI.
\]
Since $\overline{N} \mathbf{1} = \mathbf{1}$, it follows that matrix
\[
\begin{pmatrix}
1 & \mathbf{1}^T \\
\mathbf{1} & -\overline{N}
\end{pmatrix}
\]
is a normalized Hadamard matrix of order $4n$. Conversely, it is not too hard to show that a normalized Hadamard matrix of order $4n$ gives rise to a symmetric
design with \( v = 4n - 1 \). We determine the parameters of this design in terms of \( n \).

From the equation

\[
v - 1 = \frac{k^2 - k}{\lambda}
\]

we find that

\[
(4n - 2)\lambda = (n + \lambda)(n + \lambda - 1) = n^2 + (2\lambda - 1)n + \lambda^2 - \lambda
\]

and hence

\[
0 = n^2 - (2\lambda + 1)n + \lambda(\lambda + 1) = (n - \lambda)(n - \lambda - 1).
\]

If \( \lambda = n \), then \( k = 2n \). If \( \lambda = n - 1 \), then \( k = 2n - 1 \). Thus the parameters of the design are one of the pair

\[
(4n - 1, 2n - 1, n - 1), \quad (4n - 1, 2n, n),
\]

where the second pair is complementary to the first. A design with these parameters is called a Hadamard design. For a Hadamard matrix \( H \), we get one Hadamard design for each possible way of normalizing \( H \). In general these designs are not all isomorphic.

3.3 The Kronecker Product

If \( A \) and \( B \) are matrices over the same field, we define their Kronecker product \( A \otimes B \) to be the block matrix we get when we replace the \( ij \)-entry of \( A \) by \( A_{i,j}B \), for all \( i \) and \( j \). To give a very simple example, if

\[
A = \begin{pmatrix} a \\ b \end{pmatrix}, \quad B = \begin{pmatrix} u & v \end{pmatrix},
\]

then

\[
A \otimes B = \begin{pmatrix} au & av \\ bu & bv \end{pmatrix}, \quad B \otimes A = \begin{pmatrix} au & av \\ bu & bv \end{pmatrix}.
\]

We see that in general the Kronecker product is not commutative, but this is one of its few failings. It is bilinear, thus

\[
A \otimes (xB + yC) = x(A \otimes B) + y(A \otimes C),
\]

\[
(xB + yC) \otimes A = x(B \otimes A) + y(C \otimes A).
\]

One consequence of these identities is that

\[
(xA) \otimes B = x(A \otimes B) = A \otimes (xB).
\]

It is also easy to see that \( (A \otimes B)^T = A^T \otimes B^T \).

The following provides one of the most important properties of the Kronecker product.
3.3.1 Lemma. If the matrix products $AC$ and $BD$ are defined, then

$$\text{(A} \otimes \text{B})(\text{C} \otimes \text{D}) = \text{AC} \otimes \text{BD}.$$ \hfill \Box

It follows that

$$A \otimes B = (A \otimes I)(I \otimes B).$$

If $x$ and $y$ are vectors of the right orders, then we have the following useful special case of the above:

$$\text{(A} \otimes \text{B})(\text{x} \otimes \text{y}) = \text{Ax} \otimes \text{By}.$$  

If $e_1, \ldots, e_m$ is a basis for a vector space $V$ over $\mathbb{F}$ and $f_1, \ldots, f_n$ is a basis for $W$, then the vectors

$$e_i \otimes f_j, \quad 1 \leq i \leq m, \ 1 \leq j \leq n$$

form a basis for a vector space of dimension $mn$, which we denote by $V \otimes W$. We call $V \otimes W$ the tensor product of the vector spaces $V$ and $W$. (Note that there will be elements of $V \otimes W$ that are not of the form $v \otimes w$, although the vectors of this form do span $V \otimes W$.

There is a unique linear mapping $P : V \otimes V \to V \otimes V$ such that

$$P(x \otimes y) = y \otimes x$$

for all $x$ and $y$. With respect to the basis

$$\{e_i \otimes e_j : 1 \leq i, j \leq \text{dim}(V)\},$$

this is clearly a permutation matrix and $P^2 = I$, whence $P = P^T$. We call $P$ the flip operator (or matrix).

3.3.2 Lemma. If $A$ and $B$ are $m \times m$ matrices and $P$ is the flip matrix, then $P(\text{A} \otimes \text{B}) = (\text{B} \otimes \text{A})P$.

Proof. We calculate:

$$P(\text{A} \otimes \text{B})(\text{x} \otimes \text{y}) = P(\text{Ax} \otimes \text{By}) = \text{By} \otimes \text{Ax} = (\text{B} \otimes \text{A})P(\text{y} \otimes \text{x}).$$ \hfill \Box

As a corollary it follows that if $A$ is square then, although $A \otimes A^T$ need not be symmetric, the product $P(A \otimes A^T)$ is symmetric.

If $A$ and $B$ are two matrices of the same order, we define their Schur product $A \circ B$ by

$$(A \circ B)_{i,j} = A_{i,j}B_{i,j}, \quad \forall i, j.$$  

(It has been referred to as the “bad student’s product”.) This product is commutative and bilinear, and the matrix $J$ is an identity for it. We find that

$$(A \otimes B) \circ (C \otimes D) = (A \circ C) \otimes (B \circ D).$$
3.4 Symmetric and Regular Hadamard Matrices

Suppose $H$ is a Hadamard matrix of order $n$. Then $H$ is a normal matrix and therefore there is a unitary matrix $U$ and a diagonal matrix $D$ such that

$$U^* H U = D.$$  

Consequently

$$D = U H U^*$$

and since $n^{-1/2} H$ is unitary, it follows that $n^{-1/2} D$ is a product of of three unitary matrices. Therefore it is unitary and so its diagonal entries must have absolute value 1. We conclude that all eigenvalues of $H$ have absolute value $\sqrt{n}$. (Note that a real normal matrix is symmetric if and only its eigenvalues are real.)

If you prefer a more elementary argument, suppose $H$ is Hadamard of order $n$ and $z$ is an eigenvector for $H$ with eigenvalue $\theta$. Then

$$nz = H^T H z = \theta H^T z$$

and thus

$$nz^* z = \theta z^* H^T z = \theta z^* H^* z = \theta (Hz)^* z = \theta \bar{\theta} z^* z.$$  

Hence all eigenvalues of $H$ have absolute value $\sqrt{n}$.

3.4.1 Lemma. If $H$ is a symmetric Hadamard matrix with constant diagonal of order $n$, then $n$ is square.

Proof. If all diagonal entries of $H$ are equal to $-1$, we can multiply it by $-1$ and get a symmetric Hadamard matrix with 1's on the diagonal. Hence, we can assume the diagonal of $H$ contains only 1's. From our remarks above, the eigenvalues of $H$ are $\pm \sqrt{n}$. Assume $\sqrt{n}$ has multiplicity $a$.

As $\text{tr}(H)$ is equal to the sum of its eigenvalues and as $\text{tr}(H) = n$ we have

$$n = a \sqrt{n} - (n-a) \sqrt{n}.$$  

If we divide this by $\sqrt{n}$, we see that $\sqrt{n} = 2a - n$, which is an integer. \hfill \square

3.4.1 Examples

(a) $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is symmetric and its order is not a square.

(b) $\begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$ is symmetric with constant diagonal.
(c) If $H_1$ and $H_2$ are symmetric Hadamard matrices with constant diagonal, so is $H_1 \otimes H_2$.

(d) If $H$ is Hadamard and $P$ is the flip, then $(H \otimes H^T)$ is symmetric. What is its diagonal?

A Hadamard matrix is regular if its row and column sum are all equal. The class of regular Hadamard matrices is closed under Kronecker product.

3.4.2 Lemma. If all rows of $H$ has the same sum then $H$ is regular and its order is square.

Proof. Suppose $H1 = k1$ for some $k$. Then

$$kH^T1 = H^TH1 = n1$$

and so

$$H^T1 = \frac{n}{k}1.$$  

This proves the regularity. Since all row sums are equal and all column sum are equal it follows that $\frac{n}{k} = k$ and $k = \pm\sqrt{n}$.

If $H$ is regular then the sum of the entries of $H$ is $\pm n\sqrt{n}$. It can be shown that if $K$ is an $n \times n$ Hadamard matrix then the absolute value of the sum of the entries of $K$ is at most $n\sqrt{n}$, and equality holds if and only if the matrix is regular.

We can construct another interesting class of designs from regular Hadamard matrices.

3.4.3 Lemma. Let $H$ be $n \times n$ matrix with entries $\pm 1$ and assume $N = \frac{1}{2}(J - H)$. Then $H$ is a regular Hadamard matrix with row sum $h$ if and only if $N$ is the incidence matrix of a symmetric design with parameters $(4h^2, 2h^2 - h, h^2 - h)$.

Design with these parameters, or their complements, are known as Menon designs.

3.4.4 Lemma. A non-trivial symmetric $(v, k, \lambda)$-design is a Menon design if and only if $v = 4n$.

In the exercises you will be offered the chance to prove that if $D$ is a symmetric design on $v$ and $v$ is a power of 2, then $D$ is a Menon design. Since the class of regular Hadamard matrices is closed under Kronecker product, we have an infinite supply of Menon designs.

3.5 Conference Matrices

An $n \times n$ matrix $C$ is a conference matrix if $c_{i,i} = 0$ for all $i$ and $c_{i,j} = \pm 1$ if $i \neq j$ and $CC^T = (n - 1)I$. 

3.5. CONFERENCE MATRICES

Examples:

\[
\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
\]

A conference matrix is normalized if all non-zero entries in the first row are equal and all non-zero entries in the first column are equal.

If \( n \geq 2 \) and there is a conference matrix of order \( n \), then \( n \) is even.

3.5.1 Lemma. If \( C \) is a skew symmetric conference matrix then \( I + C \) is a Hadamard matrix.

Proof. 

\[
(I + C)(I + C)^T = (I + C)(I - C) = I - C^2 = I + CC^T = nI
\]

\( \square \)

3.5.2 Lemma. If \( C \) is a symmetric conference matrix, then 

\[
\begin{pmatrix} C + I & C - I \\ I - C & C + I \end{pmatrix}
\]

is a Hadamard matrix.

Proof. Compute \( HH^T \).

\( \square \)

3.5.3 Theorem. If \( C \) is a normalized conference matrix of order \( n \) then either \( n \equiv 0 \pmod{4} \) and \( C^T = -C \) or \( n \equiv 2 \pmod{4} \) and \( C^T = C \).

Proof. Exercise.

\( \square \)

3.5.4 Theorem. If \( C \) is a conference matrix of order \( n \) and \( n \equiv 2 \pmod{4} \) then \( n - 1 \) is the sum of two squares.

Proof. We have 

\[
CC^T = (n - 1)I
\]

and so the symmetric bilinear forms \( I_n \) and \( (n - 1)I_n \) are equivalent. By Witt cancellation we deduce that

\[
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \approx \begin{pmatrix} n - 1 & 0 \\ 0 & n - 1 \end{pmatrix}.
\]

Hence there is an invertible matrix \( \begin{pmatrix} a & b \\ c & d \end{pmatrix} \) with

\[
\begin{pmatrix} n - 1 & 0 \\ 0 & n - 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & * \\ * & * \end{pmatrix},
\]

which implies \( a^2 + b^2 = n - 1 \).

\( \square \)
Before long we will see to how construct conference matrices of order \( q + 1 \), one for each odd prime power \( q \). These matrices will be skew symmetric when \( q \equiv 3 \pmod{4} \) and symmetric if \( q \equiv 1 \pmod{4} \). In the former case we will then get Hadamard matrices of order \( 2q+2 \). (Using these and the Kronecker product, we can construct Hadamard matrices of all order \( 4m \) between 8 and 96, except the case \( 4m = 92 \).)

Let \( q \) be an odd prime power and let \( \mathbb{F} \) be a finite field of order \( q \). (All finite fields of the same size are isomorphic). If \( a \in \mathbb{F} \) then

\[
\mathcal{X}(a) = \begin{cases} 
1 & a \text{ is a square, but not 0} \\
-1 & a \text{ is not a square} \\
0 & a = 0
\end{cases}
\]

We call \( \mathcal{X} \) the Legendre function or quadratic character.

The Paley matrix is the \( q \times q \) matrix with rows and columns indexed by \( \mathbb{F} \), such that \( M_{a,b} = \mathcal{X}(b-a) \) for \( a, b \in \mathbb{F} \).

Note that all the diagonal entries of \( M \) are zero. The off-diagonal entries are \( \pm 1 \). If \( q \equiv 1 \pmod{4} \), then \( M \) is symmetric. If \( q \equiv 3 \pmod{4} \), \( M \) is skew-symmetric \((M^T = -M)\).

If \( C \) is a normalized Conference matrix, then the matrix we get by deleting the first row and column is its core. A Paley matrix is the core of a conference matrix (The proof is an exercise).

### 3.6 Type-II Matrices

A square \( v \times v \) matrix \( W \) over \( \mathbb{C} \) is a type-II matrix if its Schur inverse \( W^{(-)} \) exists and \( WW^{(-)}T = vI \).

1. If \( H \) is a Hadamard matrix, then \( H^{(-)} = H \), and so

\[
H^{(-)}T H = H^{(-)}T H = vI
\]

So Hadamard matrices are type-II.

2. A matrix over \( \mathbb{C} \) is flat if all of its entries have the same absolute value.

A flat unitary matrix is type-II:

\[
\begin{pmatrix}
1 & 1 & 1 \\
1 & \omega & \omega^2 \\
1 & \omega^2 & \omega
\end{pmatrix}, \text{ where } \omega^3 = 1, \omega \neq 1
\]

### 3.6.1 Lemma

Any two of the following imply the third

(a) \( W \) is type-II.

(b) \( W \) is flat.
3.6. TYPE-II MATRICES

(c) \( W \) is a non-zero scalar multiple of a unitary matrix.

3.6.2 Lemma. If \( W_1, W_2 \) are type-II, so is \( W_1 \otimes W_2 \).

Two type-II matrices are equivalent if we can get one from the other by permuting rows and columns and by non-zero rescaling of columns.

If \( W \) is type-II, so are \( W^{-1} \) and \( W^T \). These are not in general equivalent to \( W \). If \( W \) is type-II, then \( W^{-1} \) exists and is equal to \( \frac{1}{\lambda}W^{-1} \).

3.6.3 Lemma. If \( N \) is the incidence matrix of a symmetric design, and \( W = (t - 1)N + J \), then \( W \) is type-II if \( 2 - t - t^{-1} = \frac{v}{n} \).

Proof. Note first that \( WW^T = ((t - 1)N + J)((t - 1)N^T + J) \)
\[ = (t - 1)(t - 1)NN^T + k(t + t^{-1} - 2)J + vJ \]
\[ = (t - 1)(t - 1)(k - \lambda I) + (k(t + t^{-1} - 2) + v)J \]
\[ = (2 - t - t^{-1})nI + (-n(2 - t - t^{-1}) + v)J \]
Here the coefficient of \( J \) is zero if \( 2 - t - t^{-1} = \frac{v}{n} \).

Physicists refer to flat type-II matrices as generalized Hadamard matrices.

An \( n \times n \) matrix \( H \) with entries \( \pm 1, \pm i \) is a complex Hadamard matrix if \( HH^* = nI \). (Note: these are type-II matrices). Every complex Hadamard matrix \( H \) can be written as \( H = H_0 + iH_1 \), where \( H_0, H_1 \) are real, and \( H_0 \circ H_1 = 0 \).

In fact \( H \) is complex Hadamard if and only if the following is Hadamard:
\[ \begin{pmatrix} H_0 + H_1 & H_0 - H_1 \\ H_1 - H_0 & H_0 + H_1 \end{pmatrix} \]
CHAPTER 3. HADAMARD MATRICES
Chapter 4

Planes

4.1 Projective Planes

A projective plane is an incidence structure such that:

(a) Any two distinct points lie on exactly one line.

(b) Any two distinct lines have exactly one point in common.

(c) There is a set of four points such that no three are collinear.

It is not difficult to construct examples where the points are the 1-dimensional subspaces of the 3-dimensional vector space $V(3,F)$ over the field $F$, and the lines are the 2-dimensional subspaces.

To go further we introduce some more terminology. An incidence structure is a partial linear space if any two distinct points lie on at most one line. It follows that in a partial linear space, any two distinct lines have at most one point in common. We conclude that the dual of a partial linear space is a partial linear space. A linear space is an incidence structure where each pair of distinct points lie on exactly one line. A linear space is an incidence structure where each pair of distinct points lie on exactly one line. The dual of a linear space will be a partial linear space, but in general it is not a linear space. Clearly projective planes are both linear and dual linear spaces.

An $s$-arc in an incidence structure is a set of $s$ points such that no three are incident with the same line. The third axiom for a projective plane assert that a projective plane must contain a 4-arc. The dual of a projective plane is a projective plane but this is not an immediate consequence of the axioms: we must show that the dual does not contain a 4-arc.

An incidence structure that is linear and dual linear but does not contain a 4-arc is sometimes called a degenerate projective plane. The incidence structure with point set $V$ and with $V$ itself as the only line provides one example. The dual structure has one point together with a bunch of lines incident with this point. A second example arises by taking

$$V = \{0, 1, \ldots, n\}$$
and defining the lines to be $V \setminus 0$ and the pairs \{0,i\} for $i = 1, \ldots, n$. This example is self-dual.

An incidence structure is thick if each point is one at least three lines and each line is on at least three points. The non-degenerate projective planes we introduced above are not thick, but the projective planes we constructed from $V(3,F)$ are thick—each point is on $|F| + 1$ lines and each line is incident with $|F| + 1$ points.

4.1.1 Lemma. Suppose the incidence structure $P$ is linear and dual linear. Then $P$ is a projective plane if and only if it is thick.

Proof. Assume first that $P$ contains a 4-arc $S$ and let $x$ be a point in $P$. If $x \in S$, then there are three distinct lines joining $x$ to each of the remaining points in $S$. Any line of $P$ that does not contain $x$ must meet these three lines in three distinct points. This holds for each point in $S$; since no line meets all points of $S$ we conclude that each line has at least three points on it.

If $\ell$ is a line and $y$ is a point not on it then there are at least three lines on $y$ joining it to the points on $\ell$. So we done unless there is a point, $x$ say, that lies on every line. Since we have a 4-arc there are at least two lines, $\ell_1$ and $\ell_2$ say, on $x$. Since each line contains at least three points there points $z_1$ and $z_2$ on $\ell_1$ and $\ell_2$ respectively that are distinct from $x$. If the unique line on $z_1$ and $z_2$ contained $x$, it would met $\ell_1$ in two distinct points. We conclude that $P$ is thick.

Now assume $P$ is thick. We construct a 4-arc. Choose lines $\ell_1$ and $\ell_2$ that lie on the point $x$. Choose distinct points $y_1$ and $z_1$ on $\ell_1$ distinct from $x$, and then two points $y_2$ and $z_2$ on $\ell_2$, again distinct from $x$. Now $S = \{y_1, z_1, y_2, z_2\}$ is a 4-arc, because any line that meets $S$ in three points must contain two points on $\ell_1$ or two points on $\ell_2$. \qed

Since the dual of a thick incidence structure is thick, it follows that the dual of a projective plane is a projective plane.

Finally we make the connection with symmetric designs.

4.1.2 Theorem. If $P$ is a finite projective plane there is an integer $n$ such that each point is on exactly $n + 1$ lines and each line is incident with exactly $n + 1$ points.

Proof. Suppose $\ell$ is a line that is incident with exactly $n + 1$ points. If $x$ is a point not on $\ell$ and $y$ is a point on $\ell$, then there is a unique line on $x$ that meets $\ell$ in $y$. Since each line on $x$ must intersect $\ell$, we infer that each point off $\ell$ is on exactly $n + 1$ lines.

Now let $\ell'$ be a second line of $P$. Then there is a point $y$ not on $\ell$ or $\ell'$ and the map that takes the point $z$ on $\ell$ to the point $(y \lor z) \land \ell$ is a bijection. Therefore all lines in $P$ are incident with exactly $n + 1$ points.

In the proof of the previous lemma saw that each point is missed by some line, and from this it follows that all points lie on exactly $n + 1$ lines. \qed

The integer $n$ in this theorem is the order of the plane By counting edges of the incidence graph of the finite plane we see that the number of points in
4.2. **AFFINE PLANES**

A finite plane is equal to the number of lines. There a finite projective plane is a symmetric design. For a symmetric design with parameters \((v, k, \lambda)\) we recall that

\[ v = 1 + \frac{k^2 - k}{\lambda} \]

we have the following:

**4.1.3 Corollary.** If \(\mathcal{P}\) is a finite projective plane, then it is a symmetric design with parameters \((n^2 + n + 1, n + 1, 1)\) for some \(n\).

In the exercises you have the chance to prove the following:

**4.1.4 Theorem.** Let \(\mathcal{P}\) be a finite thick incidence structure. The \(\mathcal{P}\) is a projective plane if and only if its incidence graph has diameter three and girth six.

### 4.2 Affine Planes

An incidence structure \(\mathcal{A}\) is an affine plane if

(a) Each two distinct points lie on exactly one line.

(b) If \(\ell\) is a line and \(x\) is a point not on \(\ell\), there is a unique line through \(x\) and disjoint from \(\ell\).

(c) There are three points not all on one line.

In (b), we say that the line on \(x\) is **parallel** to \(\ell\).

The vector and the cosets of the 1-dimensional subspaces of a 2-dimensional vector space \(V\) over \(F\) form an affine plane.

We can also construct affine planes from projective planes. Suppose \(\mathcal{D}\) is an incidence structure with point set \(V\). Let \(B\) be a block in \(\mathcal{D}\). We form a new incidence structure with point-set \(V \setminus B\); its blocks are the blocks of \(\mathcal{D}\) distinct from \(B\) and a block is incident with a point in the new structure if it was incident with it in \(\mathcal{D}\). We call the new structure the **residual** of \(\mathcal{D}\) relative to \(B\).

**4.2.1 Theorem.** If \(\mathcal{P}\) is a projective plane and \(\ell\) is a line it, the residual of \(\mathcal{P}\) relative to \(\ell\) is a an affine plane.

**Proof.** Let \(\mathcal{A}\) denote the residual structure. It is a partial linear space because \(\mathcal{P}\) is.

We find a 3-arc. Let \(x\) be a point of \(\mathcal{P}\) not on \(\ell\). Choose two lines on \(x\) and, on both of these choose a point distinct from \(x\) and not on \(\ell\). If these points are \(y\) and \(z\), then \(x, y\) and \(z\) are not collinear.

Now let \(m\) be a line of \(\mathcal{P}\) not equal to \(\ell\) and let \(x\) be a point not on \(m\) (and not on \(\ell\)). Let \(z\) be the intersection of \(\ell\) and \(m\). Then in \(\mathcal{A}\), the unique line joining \(x\) and \(z\) is disjoint from \(\ell\). So there is a line through \(x\) parallel to \(\ell\) and we have to show it is unique. But a second parallel would be a second line in \(\mathcal{P}\) that contains \(x\) and \(z\). Since this is impossible, \(\mathcal{A}\) is an affine plane.
4.2.2 Lemma. Parallelism is an equivalence relation on the lines of an affine plane.

Proof. Let \( \ell, m \) and \( n \) be lines in the affine plane \( A \) such that \( m \) is disjoint from \( \ell \) and \( n \) is disjoint from \( m \). Assume by way of contradiction that \( \ell \) and \( n \) meet in the point \( x \). Then we have found two lines on \( x \) that are disjoint from \( n \), which contradicts our axioms.

The equivalence classes with respect to parallelism are called parallel classes.

You may check that if \( P \) is a plane of order \( n \), then the affine plane we obtain as the residual relative to a line \( \ell \) is a 2-(\( n^2, n, 1 \)) design.

4.2.3 Theorem. A 2-(\( n^2, n, 1 \)) design where \( n \geq 2 \) is an affine plane.

Proof. Let \( D \) be a 2-(\( n^2, n, 1 \)) design. Since \( \lambda = 1 \), it is a partial linear space. If \( x \) is a point not in the block \( \beta \) then \( x \) together with any two points from \( \beta \) is a 3-arc.

We see that \( r = n + 1 \). If \( x \) is a point not on the block \( \beta \), there are exactly \( n \) lines that join \( x \) to points in \( \beta \) and therefore there is a unique block on \( x \) disjoint from \( \ell \).

The process of going from a projective plane to an affine plane is reversible.

4.2.4 Theorem. Let \( A \) be an affine plane and let \( E \) be the set of parallel classes of \( A \). Let \( P \) be the incidence structure defined as follows.

(a) The points of \( P \) are the points of \( A \) and the elements of \( E \).

(b) The lines of \( P \) are the lines of \( A \), and one new line \( \ell_\infty \).

(c) The line \( \ell_\infty \) is incident only with the elements of \( E \).

(d) A line of \( A \) is incident with the element of \( E \) that contains it.

(e) The incidence between points and lines of \( A \) are unchanged.

Then \( P \) is a projective plane and \( A \) is its residual relative to \( \ell_\infty \).

4.3 Möbius Planes

We introduce our first class of \( t \)-designs where \( t > 2 \). We start with an infinite construct. Let \( V \) be the unit sphere in \( \mathbb{R}^3 \). Define a circle to be a subset of \( C \) of \( V \) such \( |C| > 1 \) and \( C \) is the intersection of \( V \) with a plane in \( \mathbb{R}^3 \) (not necessarily containing the origin). Then any three distinct points on the sphere lie in a unique circle, and we have a 3-design.

The problem is we want finite examples. Let \( F \) be a field, and let \( E \) be a quadratic extension of \( F \). In other words \( E \) is a vector space of dimension two over \( F \). Let \( \sigma \) be an automorphism of \( E \) with order two that fixes each element
of $\mathbb{F}$. One examples comes from $\mathbb{R}$ and $\mathbb{C}$, with complex conjugation as $\sigma$. A second arises if we take $\mathbb{F}$ to be a finite field of order $q$ and define

$$x^\sigma := x^q.$$ 

Then

$$(x^\sigma)^\sigma = x^{q^2} = x$$

and so $\sigma$ has order two. If $q$ is a power of the prime $p$, then

$$(x + y)^p = x^p + y^p$$

from which it follows that $\sigma$ is an automorphism. Since the multiplicative group of $\mathbb{F}$ has order $q$, it follows that $\sigma$ fixes each element of $\mathbb{F}$.

Now we work over $\mathbb{E}$. If

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

is an invertible matrix over $\mathbb{E}$, define the map $\tau_A$ by

$$\tau_A(x) = \frac{ax + b}{cx + d}.$$ 

We view this as a map from $\mathbb{E} \cup \infty$ to itself, where $\infty$ satisfies all the obvious rules that you were not permitted to use in Calculus. In particular

$$\tau_A(\infty) = \frac{a}{c}.$$ 

(Since $A$ is invertible, $a$ and $c$ are not both zero and therefore $\frac{a}{c}$ is a well-defined element of $\mathbb{E} \cup \infty$.) We note that if $B$ is a non-zero scalar multiple of $A$, then $\tau_A = \tau_B$.

The set of maps $\tau_A$, where $A$ runs over the invertible $2 \times 2$ matrices with elements from $\mathbb{E}$ is a group: you may check that

$$\tau_A \circ \tau_B = \tau_{AB}$$

and the multiplication is associative because it is composition of functions. It is denoted by $PGL(2, \mathbb{E})$. It has a subgroup consisting of the maps $\tau_A$ where the entries of $A$ are in $\mathbb{F}$. This is denoted by $PGL(2, \mathbb{F})$ and it fixes $\mathbb{F}$ as a set. Since $PGFL(2, \mathbb{F})$ is isomorphic to the group of $2 \times 2$ matrices modulo its center, its order is

$$\frac{(q^2 - 1)(q^2 - q)}{q - 1} = q^3 - q.$$ 

The index of $PGL(2, \mathbb{F})$ in $PGL(2, \mathbb{E})$ is

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

4.3.1 Theorem. Let $\mathbb{F}$ be a a field of order $q$ and let $\mathbb{E}$ be a quadratic extension of $\mathbb{F}$. Then the images of $\mathbb{F} \cup \infty$ under the action of $PGL(2, \mathbb{E})$ form a $3$-$(q^2 + 1, q + 1, 1)$ design.

Proof. Exercise.
Chapter 5

Orthogonal Arrays

5.1 Latin Squares

Traditionally a Latin square of order $n$ is an $n \times n$ array with entries from \{1, \ldots, n\} such that each integer occurs exactly once in each row and column. If we allow ourselves to use any set of size $n$ for the entries, we see that the multiplication table of a group is a Latin square. Therefore there is a Latin square of order $n$ for each non-negative integer $n$.

There is an alternative definition which works better. Suppose $A$ is matrix that represents a latin square of order $n$. Then we can have $n^2$ triples

$$(i, j, A_{i,j}).$$

We can now write down an $n^2 \times 3$ matrix with these triple as rows. (Although for typographical reasons we might write down the transpose instead.) This matrix has the property that each ordered pair of columns contains each ordered pair from \{1, \ldots, n\} exactly once. Conversely any $n^3 \times 3$ matrix over \{1, \ldots, n\} with this property comes from a latin square.

We can now introduce a key concept: an orthogonal array $OA(n,k)$ over \{1, \ldots, n\} is a matrix with $k$ columns and entries from \{1, \ldots, n\}, such that each ordered pair of columns contains each ordered pair of elements from \{1, \ldots, n\} exactly once. It follows that an orthogonal array has exactly $n^2$ rows. An orthogonal array $OA(n,2)$ is more or less the edge set of the complete bipartite graph $K_{n,n}$. An $OA(n,3)$ is a latin square. Two orthogonal arrays are equivalent if we can get one from the other by a combination of permuting rows, columns or symbols.

We can generalize the concept of orthogonal array to that of an orthogonal array with index $\lambda$, where each ordered pair occurs exactly $\lambda$ times in each pair of columns. An orthogonal array has strength at least $t$ if, for each $s$ with $1 \leq s \leq t$, each ordered $s$-tuple occurs exactly $\lambda_s$ times in each set of $s$ columns.
5.2 Examples

Let $F$ be a finite field of order $q$. If $a_1, \ldots, a_{q-1}$ are the non-zero elements of $F$, form the array with $q^2$ rows of the form

$$(x, y, x + a_1 y, \ldots, x + a_{q-1} y)$$

This is an orthogonal array. If $|F| = 3$, the rows are

$$
\begin{array}{cccc}
0 & 0 & 0 & 0 \\
0 & 1 & 1 & 2 \\
0 & 2 & 2 & 1 \\
1 & 0 & 1 & 1 \\
1 & 1 & 2 & 0 \\
1 & 2 & 0 & 2 \\
2 & 0 & 2 & 2 \\
2 & 1 & 0 & 1 \\
2 & 2 & 1 & 1 \\
\end{array}
$$

Suppose $A$ and $B$ are orthogonal arrays with $k$ columns with elements from $M$ and $N$ respectively. If $(i, j, \alpha)$ and $(k, \ell, \beta)$ are rows of $A$ and $B$ respectively, define their product to be

$$
((i, k), (j, \ell), (\alpha_1, \beta_1), \ldots, (\alpha_{k-2}, \beta_{k-2}))
$$

The set of all products forms an array over $M \times N$, and you may verify that this is an orthogonal array.

We present an application. Suppose we have a 2-$(v, k, 1)$ design $\mathcal{D}$ with point set $V$ and we want to construct a 2-$(vk, k, 1)$ design. (The parameters work.) We begin by taking $k$ disjoint copies of $\mathcal{D}$, which means we have $kb$ blocks and so we are short by

$$
\frac{vk(vk - 1)}{k(k - 1)} - k \frac{v(v - 1)}{k(k - 1)} = v^2.
$$

So we can finish our construction if we can find a collection of $v^2$ $k$-sets, consisting of one point from each copy of $V$. It is not hard to verify that this set must be an $OA(v, k)$, and that any $OA(v, k)$ will work.

5.3 Affine Planes

We have the following bound.

5.3.1 Lemma. If an $OA(n, k)$ exists, then $k \leq n + 1$.

Proof. The graph of an orthogonal array $OA(n, k)$ has the $n^2$ rows of the array as its vertices, two rows are adjacent if they agree on some coordinate. The rows that take the same value in the same column form a clique in the graph, and thus each column determines a set of $n$ vertex-disjoint cliques of size $n$. We call
this a parallel class. Since two rows agree on at most one coordinate, we can color each edge of the graph by the index of the column where the corresponding rows agree. The subgraph formed by the edges of a given color form a parallel class, and different parallel classes are edge disjoint. Now the number of edges in the graph is at most
\[ \binom{n^2}{2} \]
and the number of edges in a parallel class is
\[ n \binom{n}{2} \]
and therefore the number of parallel classes is at most
\[ \frac{n^2(n^2 - 1)}{n^2(n - 1)} = n + 1. \]

We can shorten this proof by omitting the graph, but the graph will be needed again. You may prove that a coclique in the graph of an OA\((n, k)\) has size at most \(n\); if the chromatic number of the graph is \(n\) then the OA\((n, k)\) can be extended to an OA\((n, k + 1)\) by adding a column.

**5.3.2 Theorem.** An OA\((n, n + 1)\) is an affine plane of order \(n\).

**Proof.** We take the \(n^2\) rows as point set and the \(n\)-cliques of the graph as the lines. By our argument above, any two distinct points lie in exactly one line. Thus our points and lines form a 2-(\(n^2, n, 1\)) design and any such design is an affine plane.

In fact we see that an OA\((n, n + 1)\) is, as an incidence structure, the dual of an affine plane. Any OA\((n, k)\) can be viewed as an incidence structure: the rows are the lines, the points are ordered pairs consisting of a column and a symbol and the point \((i, a)\) is incident with the rows that have \(a\) in the \(i\)-th position. Our graph above is the block graph of the incidence structure.

The graph of an orthogonal array does not determine the array in general—all affine planes of order \(n\) give rise to the complete graph \(K_{n^2}\). Of course if we keep track of the coloring, then we can reconstruct the array.

### 5.4 Partial Geometries

A partial geometry is a point and line-regular partial linear space with the property that there is a positive integer \(\alpha\) such that each point not on a line is collinear with exactly \(\alpha\) points on the line. We write PG\((s, t, \alpha)\) to denote a partial geometry with lines of size \(s + 1\) and \(t + 1\) lines on each point, where a point not on a line is collinear with exactly \(\alpha\) points on the line. So an OA\((n, k)\) is a PG\((k - 1, n - 1, k - 1)\). A 2-design with \(\lambda = 1\) is a PG\((k - 1, r - 1, k)\). (I am sorry about all the \(-1\)'s, but the geometers got here first.)

You can prove the following:
(a) A \( PG(s, t, s + 1) \) is equivalent to a 2-design with parameters \((st + s + 1, s + 1, 1)\).

(b) A \( PG(s, t, s) \) is equivalent to an orthogonal array \( OA(t + 1, s + 1) \).

(c) The dual of a partial geometry \( PG(s, t, \alpha) \) is a \( PG(t, s, \alpha) \).

(d) The edges and 1-factors of \( K_6 \) form a \( PG(2, 2, 1) \).

(e) An incidence structure is a \( PG(s, t, 1) \) if it is point and block regular and its incidence graph has diameter four and girth eight.

There are partial geometries where
\[ 1 < \alpha < \min\{s, t\} \]
but they are not easy to find. (They are sometimes said to be proper.)

**5.4.1 Lemma.** Let \( N \) be a 01-matrix. Then \( N \) is the incidence matrix of a partial geometry if and only if there are positive integers \( s, t \) and \( \alpha \) such that

(a) \( N1 = (t + 1)1 \).

(b) \( N^T1 = (s + 1)1 \).

(c) \( NN^T N = (t + s + 1)N + \alpha(J - N) = (s + t + 1 - \alpha)N + \alpha J \). \( \square \)

We leave the proof of this as an exercise. If our partial geometry has \( v \) points and \( b \) lines, then we have
\[ (t + 1)^2(s + 1) = (t + s + 1 - \alpha)(t + 1) + \alpha b \]
whence we obtain
\[ b = (t + 1) \frac{st + \alpha}{\alpha}, \quad (s + 1) = \frac{st + \alpha}{\alpha} \]

(The expression for \( v \) is a consequence of the expression for \( b \), in more than one way.)

A partial geometry is a partial linear space and therefore the matrix \( NN^T - (t + 1)I \) is a 01-matrix. Hence it is the adjacency matrix of the point graph of the geometry and \( N^T N - (s + 1)I \) is the adjacency matrix of its line graph. If \( A \) and \( B \) are matrices such that both products \( AB \) and \( BA \) are defined (i.e., \( A \) and \( B^T \) have the same order), then
\[ \det(I + xAB) = \det(I + xBA) \]
From this it follows that \( NN^T \) and \( N^T N \) have the same non-zero eigenvalues with the same multiplicities. Now from Lemma 5.4.1(c) we get that
\[ (NN^T)^2 = (s + t + 1 - \alpha)N + \alpha(s + 1)J \]
and from this it follows that the eigenvalues of $NN^T$ are

$$0, \quad s + t + 1 - \alpha, \quad (s + 1)(t + 1)$$

with respective multiplicities $v - \text{rk}(N)$, $\text{rk}(N) - 1$ and 1. Consequently the eigenvalues of the point graph are

$$-t - 1, \quad s - \alpha, \quad s(t + 1)$$

and those of the line graph are

$$-s - 1, \quad t - \alpha, \quad (s + 1)t.$$ 

In both these case the third eigenvalue is the valency of the graph, and is simple, while the multiplicity of the second eigenvalue is equal to $\text{rk}(N) - 1$.

### 5.5 Strongly Regular Graphs

A graph $X$ is **strongly regular** if it is neither complete nor empty and there are integers $k$, $a$ and $c$ such that

(a) $X$ is regular with valency $k$.

(b) Any two adjacent points have exactly $\lambda$ common neighbors.

(c) Any two points that distinct and not adjacent have exactly $c$ common neighbors.

If $X$ is a strongly regular graph on $v$ vertices its parameter vector is $(v,k;a,c)$.

If $m,n > 1$ then the disjoint union $mK_n$ of $m$ copies of $K_n$ is strongly regular with $c = 0$; its complement is strongly regular. The complement of a strongly regular is strongly regular. A strongly regular graph is **trivial** if it is not connected (in which case it is $mK_n$), or if its complement is not connected (in which case it is $\overline{mK_n}$).

#### 5.5.1 Lemma. Let $A$ be the adjacency matrix of the graph $X$. The $X$ is strongly regular with parameters $(v,k;a,c)$ if and only if

$$A^2 - (a - c)A - (k - c)I = cJ.$$ 

The block graph of a $2-(v,k,1)$ design is strongly regular or complete. The simplest way to prove this is to show to compute $k$, $a$ and $c$. However this fact is a corollary of the following:

#### 5.5.2 Lemma. The points and block graphs of a partial geometry are strongly regular.

This lemma is an easy consequence of the eigenvalue computations above, Lemma 5.5.1 and the following:

#### 5.5.3 Lemma. A connected graph is strongly regular if and only if its adjacency matrix has exactly three distinct eigenvalues.
A connected regular graph whose adjacency matrix has exactly two eigenvalues must be a complete graph. (The number of distinct eigenvalues, less one, is an upper bound on the diameter of a graph.)
Chapter 6

Block Intersections

6.1 Quasi-Symmetric Designs

A design is quasi-symmetric if there are distinct integers $x$ and $y$ such that any two distinct blocks have either $x$ or $y$ points in common. Any $2-(v,k,1)$ provides an example where $x = 0$ and $y = 1$. We usually assume that $x < y$. We define the block graph to be the graph with the blocks as its vertices, where two blocks are adjacent if and only if they intersect in exactly $y$ points.

The first class of examples is the least interesting: if $m > 1$, take $m$ copies of each of the blocks of a symmetric design. We call this a multiple of a symmetric design, it is quasi-symmetric with $x = \lambda$ and $y = k$. It easy to see that any quasi-symmetric design with $y = k$ must be a multiple of a symmetric design.

The second class consists of the $2-(v,k,1)$ designs, here we have $x = 0$ and $y = 1$. Conversely, any quasi-symmetric design with $x = 0$ and $y = 1$ is a $2-(v,k,1)$ design.

Thirdly we have the so-called strongly resolvable designs. A design is strongly resolvable if there is a partition of its blocks into classes and constants $\rho$ and $\mu$ such that any two distinct blocks in the same class intersect in $\rho$ points, while two blocks in disjoint classes meet in $\mu$ points. These are quasi-symmetric with $x = \rho$ and $y = \mu$, and can be characterized as the quasi-symmetric designs with $x = k + \lambda - r$. (The claims in this last sentence are not obvious, but will proved in Section ?? and Section ??.) The simplest examples of strongly resolvable designs are the affine planes. The block graph of a strongly resolvable design is complete multipartite graph.

The fourth class may be finite. Let $D$ be a symmetric design and let $\beta$ be a fixed block in it. The residual design has the points of $D$ not in $\beta$ as its point set, and the intersection of this set with the blocks of $D$ (other than $\beta$) as its blocks. Any residual design has $r = k + \lambda$; designs for which this condition holds are sometimes known as quasi-residual. The residual design of a symmetric design where $\lambda = 2$ can be shown to be quasi-symmetric with degree set \{1,2\}. A symmetric design such that $\lambda = 2$ is ‘dignified’ by the name biplane. It is
an open question as to whether there are infinitely many biplanes. (Or, more
generally, whether there are infinitely many symmetric 2-(v, k, λ) designs for any
value of λ greater than one.) Any 2-design with the parameters of a residual
biplane, that is with λ = 2 and r = k + 2, must be the residual design of a
biplane (see ???).

The fifth class is finite, with cardinality four. It consists of the Witt designs
on 22 and 23 points, and their complements.

Suppose B is the incidence matrix of a quasi-symmetric 2-(v, k, λ) with de-
gree set \{x, y\}. Then

\[ BB^T = (r - \lambda)I + \lambda J, \]
\[ B^T B = (k - x)I + (y - x)A + xJ. \]

As \( BB^T \) and \( B^T B \) have the same non-zero eigenvalues with the same multiplic-
ties, we can use these identities to determine the spectrum of A. Because B
has constant row and column sum, \( B^T B \) must commute with J, hence A must
commute with J.

We have

\[ kr^1 = B^T B^1 = (k - x)^1 + (y - x)A^1 + xv^1, \]

whence

\[ A^1 = (y - x)^{-1}(kr - k + x - xv)^1. \]

This gives one eigenvalue and eigenvector for A.

### 6.2 Triangle-free Strongly Regular Graphs

A graph is strongly regular with parameters \((n, k; a, c)\) if it has \(n\) vertices, va-
leny \(k\), any two adjacent vertices have exactly \(a\) common neighbours and any
two distinct non-adjacent vertices have exactly \(c\) common neighbours. The com-
plement of a strongly regular graph is strongly regular; in fact we could define
strongly regular graphs to be the graphs arising as colour classes in association
schemes with two classes. The disjoint union \(mK_n\) of \(m\) copies of \(K_n\) is strongly
regular provided \(m > 1\) and \(n > 1\)—the complete and empty graphs are not usu-
ally considered to be strongly regular. A strongly regular graph \(G\) is primitive if
both \(G\) and its complement are connected; the only imprimitive strongly regular
graphs are the graphs \(mK_n\) and their complements, the complete multipartite
graphs.

The smallest non-trivial strongly regular graph is the pentagon \(C_5\). The line
graphs of the complete graphs \(K_n\) and the complete bipartite graphs \(K_{n,n}\) are
as well. Hence the Petersen graph is strongly regular. A strongly regular graph
is triangle-free if it has no triangles, which is the same as requiring that \(a = 0\).
Only seven primitive triangle-free strongly regular graphs are known (and you
have just met two of them). One can be constructed from the Witt design on
22 points as follows.

This design has parameters 3-(22, 6, 1); hence it has 77 blocks. We construct
a graph \(HS\) with vertex set consisting of all points and blocks of this design,
and one extra point which we denote by $\infty$. The adjacencies are as follows. The vertex $\infty$ is adjacent to each of the 22 vertices corresponding to the points of the design. Each of these 22 vertices is in turn adjacent to the vertices representing the 21 blocks which lie on it. Each ‘block vertex’ is adjacent to the vertices representing the blocks disjoint from it. Although it is not obvious, this construction produces a vertex-transitive graph which is strongly regular with parameters $(100, 22; 0, 6)$. It is known as the Higman-Sims graph. (It contains, as induced subgraphs, strongly regular graphs on 16, 50 and 56 vertices—the Clebsch, Hoffman-Singleton and Gewirtz graphs respectively.)

This construction can be reversed in part. Suppose $X$ is an $(n, k; 0, c)$ strongly regular graph, and let $V$ denote the set of vertices adjacent to some fixed vertex $u$ in $X$. We define an incidence structure with point set $V$ and blocks consisting of the subsets of $V$ with size $c$ which have a common neighbour at distance two from $u$. It is not hard to show that this forms a $2-(k, c, c-1)$ design with $k(k-1)/c$ blocks, possibly repeated, and $r = k-1$. Two adjacent vertices at distance two from $u$ have no common neighbour adjacent to $u$; hence they determine disjoint blocks.

6.2.1 Lemma. Let $X$ be an $(n, k; 0, c)$ strongly regular graph, and let $D$ denote the design on the neighbours of some fixed vertex, formed as just described above. Then the following conditions are equivalent:

(a) $D$ is a 3-design,
(b) $D$ is quasi-symmetric with $x = 0$,
(c) $k = \frac{1}{2}[(3c + 1) + (c - 1)\sqrt{4c + 1}]$,
(d) The graph induced by the vertices at distance two from a fixed vertex is strongly regular.

Proof. The design on the neighbours of the vertex $u$ of $X$ has parameters $2-(k, c, c-1)$. As it has $k(k-1)/c$ blocks, it follows from ?? that (a) and (b) are equivalent.

If $v$ is at distance two from $u$ then it has exactly $k-c$ neighbours at distance two from $u$; therefore there are at least $k-c$ blocks disjoint from the block corresponding to $v$. By ??, we then have

$$\frac{k(k-1)}{c} \geq 1 + \frac{c(k-2)^2}{c^2 - 3c + k} + k - c$$

After some calculation (in Maple, preferably) we find that the difference between the two sides of this inequality is

$$(k-c)\frac{k^2 - 3kc - k + c + 4c^2 - c^3}{c(c^2 - 3c + k)};$$

from this we deduce that

$$k \geq \frac{1}{2}[(3c + 1) + (c - 1)\sqrt{4c + 1}],$$
with equality if and only if \( D \) is quasi-symmetric (with \( x = 0 \)). Hence (b) and (c) are equivalent.

The argument we just used shows that if equality holds in the previous equation, then there are exactly \( k - c \) blocks disjoint from a given block in \( D \). It follows that two vertices at distance two from \( v \) are adjacent if and only if the corresponding blocks are disjoint, and therefore the graph \( X_2(v) \) is the complement of the block graph of \( D \). As \( D \) is quasi-symmetric, it must be strongly regular. Thus (c) implies (d).

The size of the intersection of two distinct blocks of \( D \) is determined by the number of common neighbours of the corresponding vertices in \( X_2(v) \). Therefore \( D \) is quasi-symmetric if and only if \( X_2(v) \) is strongly regular. \( \square \)

### 6.3 Resolvable Designs

A parallel class in a design is a collection of blocks that partitions the point set. A design is resolvable if its block set can be partitioned into parallel classes. The canonical example is the partition of the lines of an affine plane into parallel classes. A resolvable design where any two blocks in distinct classes meet in the same number of points is known as an affine resolvable design. Examples are provided by the points and lines of an affine geometry, and by the Hadamard 3-designs.

We have the following strengthening of Fisher’s inequality.

**6.3.1 Lemma.** Let \( D \) be a 2-design with \( b \) blocks. If there is a partition \( \pi \) of the blocks of \( D \) into 1-designs, then \( b \geq v + |\pi| - 1 \).

**Proof.** Let \( B \) be the incidence matrix of \( D \) and let \( R_1, \ldots, R_c \) be the classes of a partition of the blocks of \( D \) into 1-designs. As \( R_i \) is a 1-design, the sum of the columns of \( B \) corresponding to the blocks in \( R_i \) is a positive multiple of \( 1 \). Let \( B' \) be the \( v \times (b - c + 1) \) matrix we get from \( B \) by deleting one column from each class, and then adding a column with each entry equal to one. By what we have just proved, \( B \) and \( B' \) have the same column space and therefore they have the same rank. Because \( \text{rk} B = v \), it follows that \( b - c + 1 \geq v \) as required. \( \square \)

It is natural now to ask what happens if equality holds here; we will prove that this can happen if and only if the cells of the partition are dual 2-designs. For this we will need the following result from linear algebra.

**6.3.2 Lemma.** Let \( B \) be a matrix with linearly independent rows. Then the matrix representing orthogonal projection onto the column space of \( B^T \) is \( B^T(BB^T)^{-1}B \). \( \square \)

We will need some consequences of this fact, the proofs of which are left as exercises. If \( B \) is the incidence matrix of a 2-design then the projection \( P \) onto the column space of \( B^T \) is given by

\[
P = \frac{1}{r - \lambda} \left( B^T B - \frac{\lambda k}{r} J \right).
\]
Next, suppose that $R_1, \ldots, R_c$ is a partition of the blocks of $\mathcal{D}$ and let $Q$ be the $b \times c$ matrix whose $i$-th column is $|R_i|^{-1/2}$ times the characteristic vector of the class $R_i$. The column space of $Q$ is the space of functions on the blocks of $\mathcal{D}$ that are constant on the classes of the resolution and orthogonal projection onto this space is represented by the matrix $QQ^T$. The projection onto the orthogonal complement of $1$ in this space is given by

$$M := QQ^T - \frac{1}{b} J.$$  

**6.3.3 Lemma.** If the $P$ are are just defined, then $PM = MP = 0$.

**Proof.** As $MP = (PM)^2$, we need only show that $PM = 0$. For this we show that $JM = 0$ and $BM = 0$. The former is left entirely up to you. For the latter, observe that $QQ^T$ is block diagonal and the $i$-th block is equal to $1_mJ$ (where $m$ is the size of the $i$-th cell of $|\pi|$). The sum of the columns of $B$ in the $i$-th cell of $\pi$ is equal to

$$\frac{mk}{v} 1$$

and therefore

$$BQQ^T = \frac{k}{v} J = \frac{r}{b} J.$$

Since $BJ = rJ$, we conclude that $BM = 0$. 

**6.3.4 Theorem.** Suppose $\mathcal{D}$ be a 2-design whose block set can be partitioned into $c$ 1-designs. If $b = v + c - 1$, two distinct blocks in the same class meet in exactly $r - \lambda - k$ points while blocks in distinct classes meet in $k^2/v$ points.

**Proof.** Let $B$ be the incidence matrix of $\mathcal{D}$. Suppose $R_1, \ldots, R_c$ is a partition of the blocks of $\mathcal{D}$ into 1-designs and $c = b + 1 - v$. Let $P$ be the matrix representing orthogonal projection onto the column space of $B^T$, and let $M$ be as above. By the lemma, $PM = 0$ whence it follows that $P + M$ is also a projection matrix. Because its rank is $v + c - 1$, we find that $P + M$ is the identity matrix when $b = v + c - 1$. Accordingly $P + M = I$.

If $\alpha$ and $\beta$ lie in different classes of our partition then

$$(M)_{\alpha,\beta} = -1/b.$$  

On the other hand, if $|\alpha \cap \beta| = y$ then

$$(P)_{\alpha,\beta} = \frac{1}{r - \lambda} \left(y - \frac{\lambda k}{r}\right).$$

Since $P + M = I$ we see that $(P)_{\alpha,\beta} + (M)_{\alpha,\beta} = 0$, whence

$$y = \frac{\lambda k}{r} + \frac{r - \lambda}{b} = \frac{\lambda vr + r(r - \lambda)}{br} = \frac{\lambda v + r - \lambda}{b}$$

Because $\lambda(v - 1) = r(k - 1)$, this proves that $y = k^2/v$. 

Assume next that \( \alpha \) and \( \beta \) are distinct blocks in the same class of the partition, and that \( |\alpha \cap \beta| = x \). Suppose that there are exactly \( m \) blocks in the class containing \( \alpha \) and \( \beta \). Then

\[
(M)_{\alpha,\beta} = \frac{1}{m} - \frac{1}{b}, \quad (P)_{\alpha,\beta} = \frac{1}{r - \lambda} \left(x - \frac{\lambda k}{r}\right),
\]

from which it follows that

\[
m \left(\frac{k^2}{v} - x\right) = r - \lambda.
\]

Now count the pairs \((i, \gamma)\) where \( \gamma \in D \) and \( i \in \gamma \cap \alpha \); this yields

\[
k + (m - 1)x + (b - m)\frac{k^2}{m} = kr
\]

and from this we find that

\[
m \left(\frac{k^2}{v} - x\right) = k - x.
\]

This implies that \( x = k + \lambda - r \), and also that each cell of \( \pi \) has size \( m \).

This theorem shows that a 2-design which admits a partition into \( b - v + 1 \) 1-designs must be quasisymmetric. Further each 1-design is a dual 2-design (which might have repeated blocks).

In the next section we will see that any two distinct blocks in a 2-design meet in at least \( k + \lambda - r \) points.

### 6.4 Designs with Maximal Width

The \textit{distance} between blocks \( \alpha \) and \( \beta \) is \( |\alpha \setminus \beta| \). The \textit{width} of a design is the maximum distance between two blocks in it. Majumdar has shown that two distinct blocks in 2-design must have at least \( k + \lambda - r \) points in common which provides an upper bound of \( r - \lambda \) on the width.

We prove Majumdar’s result.

#### 6.4.1 Theorem.

Any two blocks in a 2-design have at least \( k + \lambda - r \) points in common. If equality holds than the relation “meet in \( k \) or \( k + \lambda - r \) points” is an equivalence relation on the blocks of the design.

**Proof.** Let \( D \) be a 2-design with incidence matrix \( B \), and let \( P \) be the orthogonal projection onto the column space of \( B^T \). We saw that in the last section that

\[
P = \frac{1}{r - \lambda} \left(B^T B - \frac{\lambda k}{r} J\right).
\]
Let $\alpha$ and $\beta$ be distinct blocks of $D$ and suppose $|\alpha \cap \beta| = x$. Consider the $2 \times 2$ submatrix $M$ of $(r - \lambda)(I - P)$ formed by the intersections of the rows and columns corresponding to $\alpha$ and $\beta$; it is equal to

\[
\begin{pmatrix}
  r - \lambda - k + \frac{\lambda k}{r} & -x + \frac{\lambda k}{r} \\
  -x + \frac{\lambda k}{r} & r - \lambda - k + \frac{\lambda k}{r}
\end{pmatrix}.
\]

Because $I - P$ is a projection it is positive semi-definite, and thus $M$ is positive semi-definite as well. Hence $\det M \geq 0$, which implies that $r - \lambda - k + \frac{\lambda k}{r} \geq -x + \frac{\lambda k}{r}$.

Thus we have proved that any two distinct blocks of $D$ meet in at least $k + \lambda - r$ points.

We now focus on the situation when equality holds in this bound. As $I - P$ is positive semi-definite, it is the Gram matrix of a set of vectors in $\mathbb{R}^m$ (where $m = \text{rk}(I - P)$). If $x = k + \lambda - r$ then all entries of $M$ are equal, from which it follows that the $\alpha$- and $\beta$-rows of $I - P$ must be equal. This shows that $\gamma$ is a block in $D$ other than $\alpha$ and $\beta$ then

$$|\gamma \cap \alpha| = |\gamma \cap \beta|.$$  

Thus we have proved that “meeting in $k$ or $k + \lambda - r$ points” is an equivalence relation on the blocks of $D$.

Let $D$ be a 2-design with degree set $\{x, y, z\}$, where $x < y < z$ and $x = k + \lambda - r$. By an exercise???, the number of blocks which meet a given block $\alpha$ of $D$ in exactly $i$ points is independent of $\alpha$. Hence the number of blocks that meet $\alpha$ in exactly $x$ points does not depend on our choice of $\alpha$. Thus there is a partition of the blocks of $D$ with $c$ cells, all of size $b/c$, such that the size of the intersection of two distinct blocks is determined by the cells in which they lie. In particular, two distinct blocks lie in the same cell if and only if they intersect in $x$ points.

The class graph of $D$ is defined to be the graph with the equivalence classes of $D$ as its vertices, and with two vertices adjacent if and only a block in one class meets a block in the other class in $y$ points. The class graph of $D$ must be regular, but more is true. The degree of a design is the number of different values taken by $|\alpha \cap \beta|$, where $\alpha$ and $\beta$ run over the pairs of distinct blocks.

**6.4.2 Lemma.** Suppose $D$ is a 2-design with degree three. If $k + \lambda - r$ lies in its degree set of $D$, then the class graph of $D$ is strongly regular.

**Proof.** Let $\{x, y, z\}$ be the degree set of $D$. Assume $x < y < z$ and $x = k + \lambda - r$; let $c$ denote the number of classes of $D$. If $B$ is the incidence matrix of $D$ then, from our discussion above, we may write

$$B^T B = (k - x)I + M \otimes J_{b/c}$$
where
\[ M = xI + yA + z(J - I - A) = (x - z)I + (y - z)A + zJ. \]

We will show that \( M \) has exactly three distinct eigenvalues. It follows that the same is true for \( A \); since the class graph is regular this implies it must be strongly regular. (See, e.g., [] for this.)

The key observation is that the non-zero eigenvalues of \( B^T B \) are the non-zero eigenvalues of \( BB^T \). As
\[ BB^T = (r - \lambda)I + \lambda J, \]
the eigenvalues of \( BB^T \) are \( kr \) (with multiplicity one) and \( r - \lambda \) (with multiplicity \( v - 1 \)). The eigenvalue of \( M \otimes J_{b/c} \) are the products of the eigenvalue of \( M \) with the eigenvalues \( (b/c \) and 0) of \( J_{b/c} \). Accordingly the eigenvalues of \( M \) are
\[ 0, -(k - x)^c_b, \lambda v^c_b, \]
(with multiplicities \( v - 1 - b + c, b - v \) and 1 respectively).

The above proof shows that \( v - 1 + c - b \geq 0 \). It follows that if classes of \( D \) form a partition into 1-designs, the 1-designs are dual 2-designs.
Chapter 7

$t$-Designs

A $t$-design is a block-regular incidence structure such that each subset of $t$ points is incident with the same number of blocks. We denote the number of blocks that contain a given set of $i$ points by $\lambda_i$ (if it is defined). With this convention we have $\lambda_0 = b$ and $\lambda_1 = r$. A $t$-design where $\lambda_t = 1$ is called a Steiner system.

Although $t$-designs where $t \geq 3$ are not easy to find, they include some very interesting structures.

Hadamard matrices provide the most accessible class of examples. Suppose $H$ is a $v \times v$ Hadamard matrix, with the first row equal to $1$. Then each row other than the first has an equal number of 1’s and $-1$’s, and hence determines two complementary subsets of $\{1, \ldots, v\}$ of size $v/2$. The combined set of $2v - 2$ blocks forms a $3$-$(v, v/2, \lambda_3)$ design. Counting pairs consisting of an ordered triple of distinct points and a block which contains them, we find that

$$(2v - 2)v(v - 2)(v - 4)/8 = v(v - 1)(v - 2)\lambda_3$$

whence $\lambda_3 = \frac{1}{4}(v - 1)$.

7.1 Basics

If $D$ is an incidence structure with point set $V$ and $t \geq 1$, we can form the incidence structure $D_{(t)}$ whose points are the $t$-subsets of $V$ and whose blocks are the blocks of $D$, where are $t$-subset is incident with the blocks that contain it. (If we are greedy and choose $t$ too large, our incidence structure will not be very interesting.) We use $N_t$ to denote the incidence matrix of $D_{(t)}$. If $D$ is the complete design on $v$ points with blocks of size $k$, we denote its incidence matrix by $W_{t,k}(v)$ or (usually) just $W_{t,k}$.

7.1.1 Lemma. Suppose $D$ is a $t$-design and $t \geq 2$. If $x$ is a point in $D$, the points not equal to $x$ and the blocks on $x$ form a $(t - 1)$-design. Also the points not equal to $x$ and the blocks not incident with $x$ form a $(t - 1)$-design.

47
The blocks on $x$ gives form the derived design of $D$ relative to $x$. The blocks off $x$ form the complement to the derived design (relative to $x$) of the complement of $D$.

If $D'$ is the derived design of $D$, we say that $D$ is an extension of $D'$.

7.1.2 Lemma. If $D$ is a $t$-design and $s \leq t$, then $D$ is an $s$-design.

Proof. First we note that $D$ is an $s$-design if and only if the row sums of $N_s$ are all equal, that is, if $N_s \mathbf{1} = c \mathbf{1}$ for some $c$. If $D$ is a $t$-design, then $N_t \mathbf{1} = \lambda_t \mathbf{1}$ and it will be enough to show that each row of $N_s$ lies in row($N_t$).

If $\sigma$ and $\beta$ are respectively $s$-subsets and $k$-subsets of $V$, then $(W_{s,t} N_t)_{\sigma,\beta}$ is equal to the number of $t$-subsets of $\beta$ that contain $\sigma$—hence it is $\binom{k-s}{t-s}$ if $\sigma \subseteq \beta$ and zero otherwise. Consequently

$$W_{s,t} N_t = \binom{k-s}{t-s} N_s.$$  

We will prove the following important result as Corollary 7.1.3.

7.1.3 Theorem. If $t \leq k \leq v - t$, then the rows of $W_{t,k}$ are linearly independent.

7.2 Extending Fisher’s Inequality

Let $\varphi(t, v)$ denote the minimum number of blocks in a $t$-design on $v$ points. From Lemma 7.1.1 we have the inequality

$$\varphi(t + 1, v) \geq 2\varphi(t, v - 1).$$

As $\varphi(2, v) = v$, we see that a 3-design on $v$ must have at least $2v - 2$ blocks. This shows that our inequality is tight for one value of $t$, at least (and that Hadamard matrices again give rise to an exceptional class of designs).

The bound we have just found is not tight for $t > 3$. However we have an important extension of Fisher’s inequality, due to Ray-Chaudhuri and Wilson.

7.2.1 Theorem. If $D$ is a $t$-design with $b$ blocks and $v$ points, then

$$b \geq \binom{v}{\frac{t}{2}}.$$

Proof. If $N_i$ is the incidence matrices for $i$-subsets versus blocks and $2i \leq t$, then the $(\alpha, \beta)$-entry of $N_i N_i^T$ is equal to $\lambda_{\alpha \cup \beta}$. As

$$\frac{\lambda_i}{\lambda_j} = \frac{\binom{v-i}{k-i}}{\binom{k-j}{k-j}},$$

it follows that

$$\frac{1}{N_i} N_i N_i^T = \frac{1}{\binom{v-i}{k-i}} W_{i,k} W_{i,k}^T.$$

Since the rows of $W_{i,k}$ are linearly independent, the rows of $N_i$ must be linearly independent too, and this yields the bound.
If equality holds in this bound then we say that \( D \) is tight. Symmetric designs are thus tight \( 2 \)-designs.

For a tight \( 4 \)-design with \( \lambda_4 = 1 \), we have

\[
\binom{v}{2} k(k-1)(k-2)(k-3) = v(v-1)(v-2)(v-3)
\]

and so

\[
\binom{k}{2} \binom{k-2}{2} = \binom{v}{2}.
\]

### 7.3 Intersection Triangles

We work with \( t \)-designs on the set \( \{1, \ldots, v\} \). If \( i \) and \( j \) are non-negative integers and \( \alpha = \{a_1, \ldots, a_{i+j}\} \) is a sequence of distinct points, then \( \lambda_{i,j}(D) \) denotes the number of blocks \( \beta \) of \( D \) such that

\[
\beta \cap \{a_1, \ldots, a_{i+j}\} = \{a_1, \ldots, a_i\}.
\]

We note that \( \lambda_{i,0} = \lambda_i(D) \) and, if \( \overline{D} \) denotes the complement of \( D \), then \( \lambda_{0,i}(D) = \lambda_i(\overline{D}) \). We also have

\[
\lambda_{i,j} = \lambda_{i+1,j} + \lambda_{i,j+1}.
\] (7.3.1)

This leads to a version of Pascal’s triangle. If \( t = 2 \), then each entry in the triangle

\[
\begin{array}{ccc}
& b & \\
r & r & b - r \\
\lambda & r - \lambda & b - 2r + \lambda
\end{array}
\]

is the sum of the two entries immediately below it, and the \( j \)-th entry in row \( i \) is equal to \( \lambda_{i-j,j} \).

Given the recurrence in (7.3.1) an easy induction argument yields the following:

#### 7.3.1 Lemma. If \( D \) is a \( t \)-design and \( i + j \leq t \), then \( \lambda_{i,j} \) is determined by the parameters of \( D \). □

One corollary of this lemma is that the complement of a \( t \)-design is a \( t \)-design. You may show that

\[
\lambda_s(\overline{D}) = \lambda_{0,s} = \sum_{i=0}^{s} (-1)^i \binom{s}{i} \lambda_{i,0}.
\]

In some interesting cases we can compute \( \lambda_{i,j} \) even when \( i + j > t \). Suppose \( a_1, \ldots, a_k \) is a Steiner system with block size \( k \) and let \( a_1, \ldots, a_k \) be the points in the block \( \alpha \). Then we can compute \( k + 1 \) rows of the intersection triangle, because the number of blocks that intersect \( \alpha \) in \( a_1, \ldots, a_j \) is \( i \) when \( j \geq t \).
7.4 Complements and Incidence Matrices

Let \( W_{i,j} \) denote the 01-matrix with rows indexed by the \( i \)-subsets of \( \{1, \ldots, v\} \), columns by the \( j \)-subsets and with \((\alpha, \beta)\)-entry equal to 1 if \( \alpha \cap \beta = \emptyset \). Both parts of the next result are straightforward to prove, and are left as exercises.

7.4.1 Lemma. We have
(a) \( W_{i,k}W_{t,k}^T = \binom{v-t+i}{k-t} W_{i,t} \),
(b) \( W_{i,k}W_{t,k}^T = \binom{v-t+i}{k-i} W_{i,t} \).

7.4.2 Lemma. We have
(a) \( W_{t,k} = \sum_i (-1)^i W_{i,t} W_{i,k} \),
(b) \( W_{t,k} = \sum_i (-1)^i W_{i,t} W_{i,k} \).

Proof. We prove (a) and leave (b) as an exercise. Suppose that \( \alpha \) is a \( t \)-subset of \( V \) and \( \beta \) a \( k \)-subset. The \( \alpha\beta \)-entry of \( W_{t,k} \) is 1 or 0 according as \( \beta \) is contained in the complement of \( \alpha \), or not. The \( \alpha\beta \)-entry of \( W_{i,t} W_{i,k} \) is
\[
\binom{|\alpha \cap \beta|}{i},
\]
while the corresponding entry of the sum in (a) is
\[
\sum_i (-1)^i \binom{|\alpha \cap \beta|}{i} = \begin{cases} 1, & \text{if } \alpha \cap \beta = \emptyset; \\ 0, & \text{otherwise.} \end{cases}
\]
This completes the proof. \( \Box \)

7.4.3 Lemma. If \( t \leq k \leq v-t \), the matrices \( W_{t,k} \) and \( W_{t,k} \) have the same row space over the rationals.

Proof. We have \( W_{t,k} = \binom{k-t-i}{t-i}^{-1} W_{i,t} W_{i,k} \) and thus Lemma 7.4.2(a) implies that
\[
W_{t,k} = \left( \sum_i (-1)^i \binom{k-t-i}{t-i}^{-1} W_{i,t} W_{i,k} \right) W_{t,k}.
\]
Therefore each row of \( W_{t,k} \) is a linear combination of rows of \( W_{t,k} \). It is easy to verify that
\[
W_{i,t} W_{t,k} = \binom{v-k-i}{t-i} W_{i,k},
\]
whence Lemma 7.4.2(b) implies that
\[
W_{t,k} = \left( \sum_i (-1)^i \binom{v-k-i}{t-i}^{-1} W_{i,t} W_{i,k} \right) W_{t,k}.
\]
Consequently each row of \( W_{t,k} \) is a linear combination of rows of \( W_{t,k} \). \( \Box \)
7.4.4 Lemma. A design and its complement have the same strength.

Proof. From the proof of the previous lemma we know that there are matrices, $G$ and $H$ say, such that

$$W_{t,k} = GW_{t,k}, \quad W_{t,k} = HW_{t,k}.$$ 

Further the rows sums of $G$ are constant, as are the row sums of $H$. Hence if $W_{t,k}x = \lambda 1$ then $W_{t,k}x = \lambda G1 = c1$ for some constant $c$. Since $W_{t,v-k}$ is got from $W_{t,k}$ by permuting its columns, it follows that the complement of a design with strength $t$ has strength at least $t$. The lemma follows immediately.

Finally we note another important consequence of Lemma 7.4.3.

7.4.5 Theorem. The rank of $W_{t,k}$ is the minimum of its number of rows and its number of columns.

Proof. Assume $t \leq k \leq v - t$. We first consider the case where $v = t + k$. Then $W_{t,v-t}$ and $W_{t,k}$ are square of the same order. As $W_{t,v-t}$ is a permutation matrix, it is invertible. Since $W_{t,v-t}$ and $W_{t,v-t}$ have the same row space, they have the same rank and thus $W_{t,v-t}$ is invertible.

Now if $t \leq h \leq v - t$ then

$$W_{t,h}W_{h,v-t} = \binom{v-2t}{h-t}W_{t,v-t}.$$ 

Since the matrix on the right of this equation is invertible, it follows that the rows of $W_{t,h}$ are linearly independent.

7.5 Polynomials

If $M$ is a matrix and $p(x)$ is a polynomial, we define $p \circ M$ to be the matrix with the same order as $M$, with

$$(p \circ M)_{i,j} := p(M_{i,j}).$$ 

7.5.1 Lemma. Let $D$ be a $t$-design with incidence matrix $N$ and suppose $p(x)$ is a polynomial of degree, where $d \leq k$. If $M_r := N_r^TN_r$, then $p \circ M_1$ is a linear combination of the matrices $M_0, \ldots, M_d$.

Proof. There are scalars $c_{r,s}$ such that

$$x^r = \sum_s c_{r,s} \binom{x}{s}$$ 

and

$$(\binom{x}{s} \circ M_1)_{\alpha,\beta} = \binom{|\alpha \cap \beta|}{s} = (M_s)_{\alpha,\beta}.$$
7.5.2 Corollary. If \( \deg(p) = s \) and \( s \leq k \), then \( \text{rk}(p \circ M_1) \leq \binom{v}{s} \).

Proof. Assume \( i \leq j \). Since
\[
W_{i,j}W_{j,t} = \binom{t - i}{j - i}W_{i,t},
\]
we have
\[
W_{i,j}N_j = N_i
\]
and consequently \( \text{row}(N_j) \) is contained in \( \text{row}(N_i) \). Since \( \text{row}(N_i) = \text{row}(N_i^T N_i) \), it follows that \( \text{row}(M_i) \leq \text{row}(M_j) \). Therefore \( \text{row}(p \circ M_1) \) is contained in \( \text{row}(M_s) \), which has dimension \( \binom{v}{s} \). \( \square \)

The following result is also due to Ray-Chaudhuri and Wilson.

7.5.3 Theorem. If \( \mathcal{D} \) is a simple design on \( v \) points with degree \( s \), then \( |\mathcal{B}| \leq \binom{v}{s} \).

Proof. Let \( \Delta \) be the degree set of \( \mathcal{D} \) and let \( p \) be the monic polynomial of degree \( s \) with the elements of \( \Delta \) as its zeros. Then
\[
p \circ M_1 = p(k)I
\]
and accordingly
\[
b = \text{rk}(p(k))I = \text{rk}(p \circ M_1) \leq \binom{v}{s}.
\](\( \square \))

### 7.6 Gegenbauer Polynomials

Let \( G_s \) be the matrix representing orthogonal projection onto the columns of \( W_{s,k}^T \). Thus \( G_s = W_{s,k}^T(W_{s,k} W_{s,k}^T)^{-1}W_{s,k} \); we derive a more explicit expression for it.

7.6.1 Lemma. We have
\[
G_s = \sum_{i=0}^{s} (-1)^i \binom{s-i}{k-s-i} W_{i,k}^T W_{i,s}.
\]

Proof. Take \( G_s \) to be the matrix just defined. We have to prove that it is the required projection. We know that \( W_{i,k} \) and \( W_{i,k} \) have the same row space, and so it follows that the column space of \( G_s \) is contained in the column space of \( W_{s,k} \). Therefore \( G_s x \) lies in the column space of \( W_{s,k} \), for any vector \( x \), and so it will suffice to prove that \( G_s W_{s,k}^T = W_{s,k}^T \). By Lemma 7.4.2(a) we find that
\[
G_s W_{s,k}^T = \sum_{i} (-1)^i \binom{k-i}{s-i} W_{i,k}^T W_{i,s}.
\](7.6.1)

Since
\[
\binom{k-i}{s-i} W_{i,k}^T = W_{s,k}^T W_{i,s}
\]
7.7. A POSITIVE SEMIDEFINITE MATRIX

the right side of the previous equation is equal to

\[ W_{s,k}^T \sum_i (-1)^i W_{i,s} W_{i,s}. \]

By Lemma 7.4.2(b), the sum here is equal to \( W_{t,t} = I \), and thus the lemma follows.

\[ \]

7.6.2 Corollary. If \( \alpha \) and \( \beta \) are \( k \)-subsets of a \( v \)-set then

\[ (G_s)_{\alpha,\beta} = \sum_{i \geq 0} (-1)^i \frac{(k-i)(k-|\alpha \cap \beta|)}{(v-s-i)} \]

This corollary implies that the \((\alpha, \beta)\)-entry of \( G_s \) is a polynomial in \(|\alpha \cap \beta|\) with degree at most \( s \). We define the Gegenbauer polynomial \( g_s \) by

\[ g_s(x) = \binom{v}{k} \sum_{i \geq 0} (-1)^i \frac{(k-i)(k-x)}{(v-s-i)} \frac{(v-s-i)}{(k-s)}. \]

Note that \( g_s(k) = \binom{v}{s} \).

7.7 A Positive Semidefinite Matrix

We have seen already that a \( t \)-design on \( v \) points has at least

\[ \binom{v}{\lfloor \frac{t}{2} \rfloor} \]

blocks. We now prove a result which provides even more information (at somewhat greater cost). Let \( \Omega \) denote the set of all \( k \)-subsets of the \( v \)-set \( V \). If \( \Phi \subseteq \Omega \), let \( G_s(\Phi) \) denote the principal submatrix of \( G_s \) with rows and columns indexed by the elements of \( \Phi \). What we need is essentially an upper bound on the largest eigenvalue of \( G_s(\Phi) \), which we derive indirectly.

7.7.1 Lemma. If \( \Phi \) is a \( t \)-design and \( 2r \leq t \) then \(|\Phi|^{-1} \binom{v}{k} G_r(\Phi)\) is idempotent.

Proof. First recall that

\[ G_s = W_{s,k}(W_{s,k} W_{s,k}^T)^{-1} W_{s,k} \]

and therefore

\[ G_s(\Phi) = N_s^T (W_{s,k} W_{s,k}^T)^{-1} N_s. \]

If \( \Phi \) is a \( t \)-design and \( 2s \leq t \), then

\[ |\Phi|^{-1} N_s N_s^T = \binom{v}{k}^{-1} W_{s,k} W_{s,k}^T. \]

whence

\[ G_s(\Phi) = \left| \Phi \right|^{-1} \binom{v}{k} N_s^T (N_s N_s^T)^{-1} N_s. \]

Since \( N_s^T (N_s N_s^T)^{-1} N_s \) is the matrix that represents orthogonal projection onto \( \text{col}(N_s^T) \), we conclude that \(|\Phi|^{-1} \binom{v}{k} G_s(\Phi)\) is idempotent. \( \square \)
As the eigenvalues of an idempotent matrix are all equal to 0 or 1, the following follows immediately.

**7.7.2 Corollary.** If \( \Phi \) is a subset of \( \Omega \) with strength at least \( 2r \), then the matrix \( |\Phi|I - \binom{v}{k}G_r(\Phi) \) is positive semidefinite.

If a matrix is positive semi-definite then any principal submatrix of it is also positive semi-definite. Hence its diagonal entries are all non-negative. What are the diagonal entries of \( |\Phi|I - \binom{v}{k}G_r(\Phi) \)? The diagonal entries of \( G_r \) are all equal to \( \binom{v}{r} \) (why?), whence we see that if \( \Phi \) has strength \( 2r \) then

\[
|\Phi| \geq \binom{v}{r}.
\]

We have proved this already, as Theorem **7.4.5**. Fortunately we can say more.

**7.7.3 Theorem.** Let \( \Phi \) be a subset of \( \Omega \) with strength at least \( 2r \). If \( \alpha \) and \( \beta \) are distinct elements of \( \Phi \) with \( |\alpha \cap \beta| = i \) then

\[
|\Phi| \geq \binom{v}{r} + |g_r(i)|.
\]

**Proof.** Assume \( b = |\Phi| \) and suppose that \( \alpha \) and \( \beta \) are elements of \( \Phi \) such that \( |\alpha \cap \beta| = i \). Then \( |\Phi|I - \binom{v}{k}G_r(\Phi) \) has a \( 2 \times 2 \) submatrix equal to

\[
\begin{pmatrix}
 b - \binom{v}{r} & -g_r(i) \\
 -g_r(i) & b - \binom{v}{r}
\end{pmatrix}.
\]

As this submatrix must be positive semi-definite, its determinant is non-negative. Hence we find that

\[
\left( b - \binom{v}{r} - g_r(i) \right) \left( b - \binom{v}{r} + g_r(i) \right) \geq 0,
\]

which proves the theorem.

**7.7.4 Theorem.** Let \( \Phi \) be a subset of \( \Omega \) with strength at least \( 2r \). Then \( |\Phi| \geq \binom{v}{r} \) and, if equality holds, the degree set of \( \Phi \) is the set of zeros of \( g_r \).

**Proof.** Theorem **7.7.3** shows that \( |\Phi| \geq \binom{v}{r} \) and that, if equality holds, \( g_r(i) = 0 \) for any \( i \) in the degree set of \( \Phi \). This shows that the degree of \( \Phi \) is at most \( r \). If it is less than \( r \) then \( |\Phi| < \binom{v}{r} \), by Theorem **7.5.3** and therefore its degree is \( r \).

One consequence of the last theorem is that the degree set of a tight design is determined by \( v, k \) and \( r \). Further, most polynomials \( g_r \) do not have \( r \) integer zeros. Thus we obtain strong restrictions on the parameters of a tight design.
Chapter 8

Witt Designs

Our main goal in this section is to construct the Witt designs on 23 and 24 points; these are 4- and 5-designs respectively.

8.1 Codes

A code is a subspace of a vector space, along with an implied threat that Hamming distance will be mentioned very soon. The Hamming distance between two vectors is the number of coordinate positions in which they differ, and the weight of a vector is the number of non-zero entries in it—this is its the Hamming distance from the zero vector. The length of a code is the dimension of the vector space that contains it. The vector space $V$ is often $GF(q)^n$ but it could, for example, be the quotient ring $GF(q)[x]/(x^n - 1)$. In the latter case, subspaces which are invariant under multiplication by $x$ are cyclic codes.

If $V$ is a vector space and $v \in V$ then we define

$$v^\perp = \{ u \in V : v^T u = 0 \};$$

this is a subspace of $V$ and, if $v \neq 0$, its codimension in $V$ is 1. If $S \subseteq V$, then

$$S^\perp := \bigcap_{v \in S} v^\perp.$$

If $U$ is a code in $V$, then $U^\perp$ is the dual code of $U$. If $U \leq U^\perp$ we say that $U$ is self-orthogonal, and if $U = U^\perp$ that $U$ is self-dual.

Even if $v \neq 0$, we may find that $v \in v^\perp$—for example if $v = 1$ and $F = \mathbb{Z}_2$ and $\dim(V)$ is even. More generally, the intersection $U \cap U^\perp$ need not be empty. The following useful properties still hold though.

8.1.1 Lemma. If $U$ is a subspace of the vector space $V$, then

(a) $(U^\perp)^\perp = U$.

(b) $\dim(U^\perp) = \dim(V) - \dim(U)$. 

55
The codes of interest to us will arise usually as the column spaces of incidence matrices. We use $\text{rk}_p(N)$ to denote the rank of $N$, viewed as a matrix over $\mathbb{Z}_p$. We eliminate some uninteresting cases.

8.1.2 Lemma. Let $D$ be a 2-design with incidence matrix $N$ and let $n = r - \lambda$ be the order of $D$. If $p$ is a prime that does not divide $n$, then $\text{rk}_p(N) \geq v - 1$. If $p$ divides $n$, then $\text{rk}_p(N) \leq \frac{1}{2}(v + 1)$.

Proof. We have

$$NN^T = nI + \lambda J.$$ 

If $p$ is a prime that does not divide $n$, then $nI$ is invertible and $\text{rk}_p(J) = 1$. Therefore $\text{rk}_p(nI + \lambda J) \geq v - 1$.

If $p \mid n$ then

$$NN^T = \lambda J \mod p$$

and therefore $\text{rk}_p(NN^T) \leq 1$. It follows that (prove it!)

$$\text{rk}_p(N) \leq \frac{v + 1}{2}.$$

If $p$ does not divide $N$ and $\text{rk}_p(N) = v - 1$, you may prove that $\text{col}(N) = 1^\perp$.

Note that if $NN^T = \lambda J$ modulo $p$, then the space spanned by the differences of the rows of $N$ is self-orthogonal.

8.2 Perfect Codes

The ball of radius $e$ about a code word $w$ is the set of words that are at distance at most $e$ from $w$; we denote it by $B_e(w)$. The packing radius of a code $C$ is the greatest integer $e$ such that that the balls of radius $e$ are pairwise disjoint. If $C$ has packing radius $e$, then the minimum distance between two distinct code words is at least $2e + 1$. The covering radius of $C$ is the least integer $r$ such that each word lies in $B_r(w)$, for some code word $w$. A code is perfect if its packing radius is equal to its covering radius.

If $C$ is a code of length $n$ over an alphabet of size $q$ with packing radius $e$, then

$$|C| \leq \frac{q^n}{\sum_{i=0}^{e} \binom{n}{i}(q-1)^i};$$

this is the sphere packing bound, and follows trivially once we observe that the denominator is the size of a ball of radius $e$. A code is perfect if an only if equality holds in the sphere packing bound.

8.2.1 Lemma. Suppose $C$ is a perfect binary code of length $n$ and packing radius $e$ that contains the zero word. Then the supports of the codes words of weight $2e + 1$ form a design with parameters $(e + 1)-(n, 2e + 1, 1)$. 

\[ \square \]
8.3. THE BINARY GOLAY CODE

We offer an example. Let \( H \) be the matrix of \( GF(2) \) with the distinct non-zero binary vectors of length \( k \) as its columns. Thus \( H \) is \( k \times (2^k - 1) \). Let \( C \) be the kernel of \( H \). You may show that the rows of \( H \) are linearly independent, whence

\[
|C| = 2^{2^k-1-k}.
\]

Since the columns of \( H \) are distinct and non-zero, there are no words in \( C \) with weight one or two and therefore the packing radius of \( C \) is at least 1. The ball of radius 1 about a word in a binary code of length \( n \) has size \( n + 1 \), which is \( 2^k \) in our case. Hence the sphere-packing bound is tight and \( C \) is perfect. (In fact \( C \) is the binary Hamming code.) A perfect binary code with packing radius 1 gives rise to a Steiner triple system on \( 2^k - 1 \) points. Examples are known that are not Hamming codes.

8.3 The Binary Golay Code

A code is even if each word in it has even weight. It is doubly even if each word has weight divisible by four.

8.3.1 Lemma. Let \( G \) be a matrix over \( GF(2) \). If each row of \( G \) has even weight, then \( \text{row}(G) \) is an even code. If \( GG^T = 0 \) and each row of \( G \) is doubly even, then \( \text{row}(G) \) is doubly even.

We construct a \([24, 12, 8]\)-code over \( GF(2) \). Let \( N \) be the incidence matrix of the symmetric 2-(11, 6, 3) design (the complement of the Hadamard design). Define

\[
\hat{N} := \begin{pmatrix} 0 & 1^T \\ 1 & N \end{pmatrix}.
\]

and

\[
G = \begin{pmatrix} I & \hat{N} \end{pmatrix}.
\]

8.3.2 Lemma. The row space of \( G \) is a self dual binary code with with minimum distance eight.

Proof. Since any two distinct blocks of the 2-(11, 6, 3) design have exactly three points in common, it is easy to verify that any two rows of \( G \) are orthogonal. Since each row of \( G \) is doubly even, \( \text{row}(G) \) is a doubly even code. You may show that it does not contain any words of weight four.

8.3.3 Lemma. The words of weight eight in a binary \([24, 12, 8]\)-code form a 5-(24, 8, 1) design.

Proof. Let \( \alpha \) be a subset of \( V = \{1, \ldots, 24\} \) with size five. If \( x \) and \( y \) are two code words of weight eight with \( \alpha \) in their support, then \( x + y \) has weight at
most six. Hence each 5-subset of \( V \) is contained in the support of at most one code word of weight eight, and so the number of such words is at most

\[
\binom{24}{5} / \binom{8}{5} = 759
\]

and, if equality holds, we have our 5-design.

### 8.4 The Plane of Order Four

Let \( \mathcal{P} \) denote a projective plane of order four. We will show that all such planes are isomorphic, and more.

A \( k \)-arc in a projective plane is a set of \( k \) points, no three of which are collinear. A line which meets a \( k \)-arc in exactly one point is a tangent to it and a line which meets it in two points are secants. In a projective plane of order \( n \) there are \( n + 1 \) lines through each point, and it follows that there are exactly \( n + 2 - k \) tangents to each point in a \( k \)-arc. Hence we have:

**8.4.1 Lemma.** If there is a \( k \)-arc in a plane of order \( n \) then \( k \leq n + 2 \). If equality holds then \( n \) must be even.

**Proof.** Only the last claim requires attention. If an \((n + 2)\)-arc exists then the lines through any point not in the arc must meet it in zero or two points. Therefore the lines which meet the arc partition it into pairs.

We call an \((n + 1)\)-arc in a plane of order \( n \) an oval, and an \((n + 2)\)-arc is a hyperoval.

**8.4.2 Lemma.** The tangents to an \((n + 1)\)-arc in a plane of even order are concurrent.

**Proof.** Let \( C \) be an oval in a plane of order \( n \), where \( n \) is even. Let \( x \) be a point not on \( C \). Each line on \( x \) meets \( C \) in at most two points and, since \( |C| \) is odd, it follows that there must be a tangent to \( C \) through \( x \). Since each point on \( C \) lies on a tangent, it follows that each point on a secant to \( C \) lies on at least one tangent. But there are only \( n + 1 \) tangents to play with, therefore each point on a secant lies on a unique tangent to \( C \). Now let \( p \) be a point which lies on two tangents. Then \( p \) cannot lie on a secant and so each of the \( n + 1 \) lines on \( p \) are tangents.

The point through the tangents all pass is called the nucleus of the \((n + 1)\)-arc. An \((n + 1)\)-arc, together with its nucleus is a hyperoval.

**8.4.3 Lemma.** Any projective plane of order four contains 2520 4-arcs, and each 4-arc lies in a unique hyperoval.

**Proof.** A plane of order four contains \( 21 \cdot 20 \cdot 16 \) ordered 3-arcs. A point \( p \) in a 3-arc \( C_3 \) lies on three tangents. If \( q \) is a point distinct from \( q \) on one of these tangents and not on the line joining the two points in \( C_3 \setminus p \) may be adjoined
to \( C_3 \) to give an ordered 4-arc. There are nine such points \( q \), whence there are \( 21 \cdot 20 \cdot 16 \cdot 9 \) ordered 4-arcs.

Let \( C_4 \) be an ordered 4-arc and suppose that \( p \in C_4 \). Then \( p \) lies on two tangents. Any tangent to \( p \) meets the three secant to \( C \) which do not contain \( p \) in three distinct points. Thus there are exactly two points which may be added to \( C_4 \) to form an oval, and so there are exactly two ordered 5-arcs which contain \( C_4 \) and both of these extend to a hyperoval.

It follows from that there are \( 21 \cdot 20 \cdot 16 \cdot 9 \cdot 2 \) ordered 6-arcs in the plane of order four, hence exactly 168 hyperovals.

Let \( C \) be a hyperoval in a plane of order four. There are three secants through each point off \( C \), which partition its points into three pairs. Let the point set of \( C \) be

\[ \{1, 2, 3, 4, 5, 6\} . \]

Then the fifteen points off \( C \) determine fifteen distinct 1-factors of \( K_6 \). There are two classes of lines in our plane, the fifteen secants to \( C \) and the remaining six lines which are disjoint from it. These, we now admit, are called passants. The five 1-factors determined by the points on a passant form a 1-factorization of \( K_6 \) and distinct passants give distinct 1-factorizations. Each secant meets \( C \) in two points, and thus determines an edge of \( K_6 \). Thus the points of our plane now correspond to the 1-factors and the vertices of \( K_6 \), and the lines to the fifteen edges and six 1-factorizations of \( K_6 \). There are only six 1-factorizations of \( K_6 \), however.

A point of the plane is on a given passant if and only if the corresponding 1-factor lies in the passant’s 1-factorization. A point is on a secant if it is in the edge determined by the secant, or its 1-factor is one of the three containing that edge of \( K_6 \). It follows that the incidence structure with the vertices and 1-factors of \( K_6 \) as its vertices, and the edges and 1-factorizations as its vertices, is a projective plane of order four. Since this structure is unique, it follows that all planes of order four are isomorphic. (It also follows that all hyperovals in a plane of order four are equivalent under its collineation group.)

### 8.5 The Code of \( PG(2, 4) \)

Let \( P \) be a projective plane of order four and let \( B \) be its incidence matrix. We study the code over \( GF(2) \) generated by the columns of the matrix

\[ \hat{B} = (B \ 1) . \]

This is the extended binary code of \( PG(2, 4) \), which we denote by \( C \). (The binary code is the code generated by the rows of \( B \).) It is immediate that \( C \) is a self-orthogonal code of length 22.

#### 8.5.1 Lemma

The extended binary code of \( PG(2, 4) \) has dimension at least ten.
Proof. We first show that the minimum distance of $C^\perp$ is at least six. Suppose we have a minimal linearly dependent set $S$ of columns of $B$ with size at most five. The columns in $S$ correspond to lines in $PG(2,4)$. Since there are five points on any line of $PG(2,4)$, for any line $\ell$ in $S$ there is a point which lies on $\ell$, but no other line of $S$. This implies that the columns in $S$ cannot be linearly dependent.

By duality it follows that any set of five rows of $\hat{B}$ are linearly independent.

Next, it is not too hard to show that a set of six rows sums to zero if and only if the corresponding set of points forms a hyperoval. Hence there is no set of six rows, three of which correspond to collinear points, which sums to zero. Any set of three non collinear points lies in exactly three hyperovals, hence any code word which is a sum of three rows, corresponding to three non-collinear points, can be so expressed in exactly four ways. Thus we have found

$$1 + 21 + \binom{21}{2} + 210 + \frac{1}{4} \left( \binom{21}{3} - 210 \right) = 722$$

distinct code words in $C$. Since $722 > 512$, it follows that the dimension of $C$ is at least ten.

Our next step is to determine some information about the characteristic functions of the hyperovals.

8.5.2 Lemma. Let $C$ be the extended binary code of $PG(2,4)$. Then the characteristic function of any hyperoval lies in $C^\perp$.

Proof. Any line meets a hyperoval in an even number of points (zero or two, in fact) and so the characteristic function of a hyperoval is orthogonal to the characteristic function of any extended line, i.e., to any row of $\hat{B}$.

We are now left with two possibilities, either $C = C^\perp$ and $C$ has dimension 11, or $C$ is properly contained in $C^\perp$ and its dimension is 10. We verify that the latter is true.

8.5.3 Lemma. Let $C$ be the extended binary code of $PG(2,4)$. Then the characteristic function of any hyperoval does not lie in $C$.

Proof. Consider the subcode $C'$ of $C$ generated by the complements of the extended blocks. This code has index at most two in $C$ and, since all words in $C'$ are zero on their last coordinate, its index is exactly two. In particular this implies that $C'$ is also the subcode of $C$ generated by the words with last coordinate zero. Since $C'$ is a self-orthogonal code generated by words of weight divisible by four, all words in it have weight divisible by four. It follows that the characteristic polynomial of a hyperoval cannot lie in $C'$.

It follows that the code generated by the rows of $\hat{B}$ and the characteristic function of one oval is self-orthogonal code with dimension 11. As the codimension of $C$ in $C^\perp$ is two, there are exactly three codes of dimension 11 between $C$ and $C^\perp$.

***to be completed***
Chapter 9

Groups and Matrices

9.1 Group Matrices

Let $G$ be a finite group. We say that a matrix $M$ is a group matrix or a $G$-matrix if its rows and columns are indexed by the elements of $G$, and there is a function $\psi$ in $G$ such that

$$M_{g,h} := \psi(hg^{-1}).$$

Normally the first row and column of a group matrix will be indexed by the identity element and $\psi$ will take values in a field. A group matrix over $\mathbb{Z}_n$ is better known as a circulant. A circulant is a matrix where each row is the cyclic right shift of the row ‘above’ it.

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

Thus a circulant is a group matrix over $\mathbb{Z}_n$.

If $M$ is a $G$-matrix then so is $M^T$ (exercise). The set of all group matrices over $G$ with values in the ring $R$ will be denoted by $R[G]$. A difference set over $G$ is a $G$-matrix that is the incidence matrix of a symmetric design.

9.1.1 Lemma. If $M$ and $N$ are group matrices, then so are $MN$ and $M \circ N$.

It follows that the set of all group matrices over $G$ with values in $\mathbb{F}$ is a matrix algebra—a vector space of matrices that contains $I$ and is closed under matrix multiplication. It is also closed under transpose and under field automorphisms (for example, under complex conjugation when $\mathbb{F} = \mathbb{C}$.) If we need to be precise, it is a faithful representation of the group algebra of $G$ over $\mathbb{F}$ (but fortunately we do not need to be precise).

A function $\mu$ on $G$ is constant on conjugacy classes if

$$\mu(g^{-1}ag) = \mu(a)$$
for all \( g \) in \( G \). (Suppose for example that the elements of \( G \) were matrices and \( \mu \) is the trace function.) We say a group matrix is central if its defining function is constant on conjugacy classes.

**9.1.2 Lemma.** Let \( M \) be a group matrix based on the function \( \mu \). Then \( M \) commutes with all \( G \)-matrices if and only if \( \mu \) is constant on the conjugacy classes of \( G \).

There is a natural basis for the space of \( G \)-invariant matrices. Define \( P_x \), for \( x \in G \) by

\[
(P_x)_{g,h} = \begin{cases} 1, & g h^{-1} = x \\ 0, & \text{otherwise} \end{cases}
\]

We can check that \( P_x P_y = P_{xy} \) (exercise). Note that \( P_x \) is a permutation matrix. If \( e_g \) is the standard basis for \( \mathbb{F}^G \) (indexed by the elements of \( G \)), then

\[
P_g e_x = e_{xg^{-1}}.
\]

If \( \psi \) and \( \varphi \) are functions on \( G \), their convolution \( \psi \ast \varphi \) is given by

\[
(\psi \ast \varphi)(g) := \sum_{x \in G} \psi(x) \varphi(x^{-1} g)
\]

If \( M \) is a \( G \) matrix and \( N \) is an \( H \)-matrix, then \( M \otimes N \) is a group matrix for \( G \times H \).

### 9.2 Eigenvalues and Eigenvectors

We consider the space of group matrices over a finite abelian group \( G \). If \( \eta \) is a function on \( G \), then we define \( M(\eta) \) to be the matrix with columns

\[
P_g \eta, \quad g \in G.
\]

We have

\[
M(\eta) = \sum \eta(g) P_g^{-1}.
\]

The matrices \( P_g \) for \( g \) in \( G \) are normal and commute, hence they have a common orthonormal basis of eigenvectors. If \( \eta \) is a common eigenvector, there is a complex-valued function \( \lambda_g \) on \( G \) such that

\[
P_g \eta = \lambda_g \eta.
\]

This function is a homomorphism from \( G \) to the group of complex numbers of norm 1; such a function is a character of \( G \). Clearly each eigenvector determines a character of \( G \). Remarkably, the converse is also true:

\[
P_g \sum \lambda_x e_x = \sum \lambda_x e_{xg^{-1}} = \sum \lambda_{yg} e_y = \lambda_g \sum \lambda_y e_y.
\]
If $\lambda$ and $\mu$ are two characters of $G$, then the map

$$g \mapsto \lambda g \mu$$

is again a homomorphism from $G$ into the $\mathbb{C}^*$, and so it is a character. Hence the set of characters of $G$ forms a group, called the dual group of $G$ and denoted by $G^*$. It is isomorphic to $G$. (Prove this for cyclic groups, then show that $(G \times H)^* = G^* \times H^*$.)

Fix an isomorphism from $G$ to $G^*$, and let $\chi_g$ be the character assigned to $g$ in $G$. If $M$ is a $G$-matrix, then we define $\Phi(M)$ to be the matrix with rows and columns indexed by $G$, such that $(\Phi(M))_{g,h}$ is equal to the eigenvalue of $M$ on the character $\chi_{hg^{-1}}$. Thus $\Phi(M)$ is a $G$-matrix,

$$\Phi(I) = J$$

and, for any two group matrices $M$ and $N$,

$$\Phi(MN) = \Phi(M) \circ \Phi(N).$$

We also have

$$\Phi(J) = nI.$$

9.2.1 Theorem. We have $\Phi(M \circ N) = n^{-1}\Phi(M)\Phi(N)$ and $\Phi^2(M) = nM^T$.

### 9.3 Multipliers

A multiplier of a difference set in $G$ is an automorphism of $G$ that is also an automorphism of $D$.

9.3.1 Theorem. Suppose $D$ is a symmetric $(v, k, \lambda)$ design and $p | n$ and $p > \lambda$. If $S \subseteq V$ such that $|S| = k$ and its characteristic function $x_S \in \col_p(D)$, then $S$ is a block.

Proof. Since $p > \lambda$, we see that $p$ does not divide $k$. Let $\beta_1, \ldots, \beta_v$ denote the blocks of $D$. Since $x_S \in \col_p(D)$, we can write

$$x_S = \sum_i a_i x_{\beta_i}, \quad a_i \in GF(p).$$

Then

$$k = 1^T x_S = \sum_i a_i 1^T x_{\beta_i} = k \sum a_i \quad (9.3.1)$$

If $\beta$ is some block in $D$ we also have

$$\lambda = x_{\beta}^T x_S = \sum a_i x_{\beta}^T x_{\beta_i} = k \sum a_i = k \pmod{p}$$

and, since $p > \lambda$, this implies that $|S \cap \beta| \geq \lambda$.  


We now prove that if $S$ is a $k$-subset of $V$ and $|S \cap \beta| \geq \lambda$ for all blocks $\beta$, then $S$ is a block. We have
\[ NN^T = nI + \lambda J \]
and consequently
\[ N \left( N - \frac{\lambda}{k} J \right)^T = nI. \]
If $|S \cap \beta| \geq \lambda$ for all blocks $\lambda$ then $x_S^T N \geq \lambda 1^T$ and $x_S^T \left( N - \frac{\lambda}{k} J \right) \geq 0$. We have
\[ x_S^T \left( N - \frac{\lambda}{k} J \right) = ne_\beta \]
Let $a^T = x_S^T \left( N - \frac{\lambda}{k} J \right)$. Then,
\[ a^T = \sum a_\beta e_\beta^T \]
\[ = \frac{1}{n} \sum a_\beta ne_\beta^T \]
\[ = \frac{1}{n} \sum a_\beta x_\beta^T \left( N - \frac{\lambda}{k} J \right) \]
and therefore
\[ s_S^T \left( N - \frac{\lambda}{k} J \right) = \sum \frac{a_\beta}{n} x_\beta^T \left( N - \frac{\lambda}{k} J \right). \]
Since $N - \frac{k}{\lambda} J$ is invertible, we find that
\[ x_S = \frac{1}{n} \sum a_\beta x_\beta \]
where $a_\beta \geq c$ for all $\beta$. Since $x_S$ has weight $k$, only $a_\beta$ can be non-zero. So $x_S = x_\beta$ for some block $\beta$.

There is a useful way to view the second part of the above proof. If $M$ is a matrix, then the vectors $y$ such that $y^T M \geq 0$ form a convex cone, that is, a set of vectors in a real vector space closed under addition and multiplication by non-negative scalars. The set of all non-negative linear combinations of the columns of $M$ is also a convex cone, dual to the first. In the above proof we are arguing that the rows of $N$ generate the dual to the cone generated by the rows of $N - \frac{k}{\lambda} J$.

9.3.2 Theorem. Let $D$ be a symmetric design given by a difference set in the abelian group $G$. If $p \mid n$ and $p > \lambda$, then $p$ is a multiplier of the difference set.

Proof. We can assume $N \in F_p(G)$; $N$ is a sum of permutation matrices from $G$ with coefficients in $F_p$.

We can write $N$ as $\sum c_i P_i$ where $c_i = 0, 1$, and then
\[ N^p = \sum c_i^p P_i^p = \sum c_i P_i^p. \]
9.3. MULTIPLIERS

If $P \in G$ then $PN - NP$, as $\text{col}(N)$ is invariant under each element of $G$. Also,

$$N^p = N^{p^{-1}}N$$

whence we see that $\text{col}(N^p) \subseteq \text{col}(N)$. Each column of $N^p$ is the characteristic vector of a subset of size $k$, hence it must be the characteristic vector of a block of $D$.

9.3.3 Lemma. A multiplier of a symmetric design over a group $G$ fixes at least one block.

Proof. We can represent the action of the multiplier on points by the permutation matrix $P$. Then each column of $PN$ is a block, and $PN = NQ$ for some permutation matrix $Q$. Since $N^{-1}$ exists,

$$Q = N^{-1}PN$$

and therefore $\text{tr}(Q) = \text{tr}(P)$. Therefore the number of blocks fixed by $Q$ equals the number of points fixed by $P$. Since $P$ fixes 0, we are done.

9.3.4 Theorem. If $(v, k) = 1$ and $D$ is given by a difference set in an abelian group, there is a block which is fixed by every multiplier.

Proof. Let $\beta$ be a block and let $h$ be the product of its elements. If $g \in G$, then the product of the elements in the block $\beta g$ is $g^k h$. Since $k$ and $n$ are coprime, the map $g \mapsto g^k$ is an automorphism of $G$ and therefore there is a unique element $g$ of $G$ such that $g^k h = 1$. This shows that there is a unique block $\gamma$ such that the product of its elements if 1. Clearly this block is fixed by all multipliers.
Chapter 10

Mutually Unbiased Bases

10.1 Complex Lines and Angles

We want to work on problems related to sets of lines in complex (and real) space. A line is a 1-dimensional subspace and we can specify it by offering a basis, that is, a non-zero vector that spans it. Our complex inner product is given by

$$\langle x, y \rangle = \sum_i x_i \overline{y_i},$$

it is linear in the second variable. If $x$ and $y$ are unit vectors, the angle between the lines they span is determined by the real number

$$\langle x, y \rangle \langle y, x \rangle = \langle x, y \rangle \overline{\langle x, y \rangle} = |\langle x, y \rangle|^2.$$

Note that the value of this product does not change if we replace $x$ (or $y$) by another unit vector that spans the same line. Over $\mathbb{R}$ this product is equal to the square of the cosine of the angle between the lines.

If $x \neq 0$, then the matrix

$$P = (x, x)^{-1}xx^*$$

represents orthogonal projection onto the line spanned by $x$; this matrix is independent of the choice of vector that spans our line. If $x$ and $y$ are unit vectors and

$$P := xx^*, \quad Q := yy^*$$

then

$$\text{tr}(PQ) = \text{tr}(xx^*yy^*) = \text{tr}(y^*xx^*y) = \langle x, y \rangle \langle y, x \rangle.$$  

Note that if $M$ and $N$ are complex $d \times d$ matrices, then $\text{tr}(M^*N)$ is an inner product (on Mat$_{d \times d}(\mathbb{C})$) which we will denote by $\langle M, N \rangle$. 

67
10.2 Equiangular Lines

A set of lines is equiangular if the angle between any pair is the same.

10.2.1 Theorem. The size of an equiangular set of lines in $\mathbb{C}^d$ is at most $d^2$.

Proof. Let $P_1, \ldots, P_n$ be projections onto a set of $n$ equiangular lines in $\mathbb{C}^d$. Suppose that we have scalars $c_i$ such that

$$0 = \sum_i c_i P_i.$$

If we multiply each side of this by $P_r$ and take traces, we get

$$0 = c_r + \sum_{i \neq r} c_i a^2 = (1 - a^2)c_r + a^2 \sum_i c_i$$

from which it follows that $c_r$ is independent of $r$. So we have

$$0 = c_1 \sum_i P_i;$$

since $\text{tr}(P_i) = 1$ it follows that $c_i = 0$ for all $i$. $\square$

10.2.2 Theorem. If $da^2 < 1$ and there is a set of $n$ equiangular lines in $\mathbb{C}^d$ with squared cosine $a^2$, then

$$n \leq \frac{d - da^2}{1 - da^2};$$

Equality holds if and only if $\sum_i P_i = \frac{n}{d} I$.

Proof. Suppose

$$S := \sum_i P_i - \frac{n}{d} I$$

Then

$$S^2 = \sum_{i,j} P_i P_j - \frac{2n}{d} \sum_i P_i + \frac{n^2}{d^2} I$$

and, taking traces and noting that $\text{tr}(S^2) \geq 0$, we get

$$0 \leq n + n(n - 1)a^2 - \frac{n^2}{d}.$$ 

If we divide this by $n$ and rearrange, we get the bound stated in the theorem. Equality holds if and only if $S = 0$.

If we have set of $d^2$ equiangular lines in $\mathbb{C}^d$ with projections $P_i$, then there are scalars $c_i$ such that

$$I = \sum_i c_i P_i$$

and a variation on the argument in the proof of Theorem 10.2.1 yields that the $c_i$’s are all equal, and hence each is equal to $n/d$. So we have equality in the bound of the previous theorem, which yields that $a^2 = (d + 1)^{-1}$. 

10.3 Reality

Suppose $x_1, \ldots, x_n$ are unit vectors spanning a set of equiangular lines in $\mathbb{R}^d$ and

$$\sum_i P_i = \frac{n}{d} I.$$ 

Then if $U := (x_1 \cdots x_n)$ we find that

$$\frac{n}{d} I = \sum_i x_i x_i^T = U U^T.$$ 

Hence if $G := U^T U$ (that is, if $G$ is the Gram matrix of $x_1, \ldots, x_n$), then

$$G^2 = U^T U U^T U = \frac{n}{d} U^T U.$$ 

Equivalently the matrix

$$\frac{d}{n} G$$

is a projection with rank $d$, and therefore the eigenvalues of $G$ are $n/d$ (with multiplicity $d$) and 0 (with multiplicity $n - d$).

Note that $G$ is a symmetric matrix with diagonal entries equal to 1 and off-diagonal entries equal to $\pm a$; thus we can write

$$G = I + a S$$

where $S$ is a symmetric matrix with zero diagonal and off-diagonal entries $\pm 1$. It is the Seidel matrix of the set of lines. The eigenvalues of $S$ must be

$$-\frac{1}{a}, \frac{n - d}{da}$$

with respective multiplicities $n - d$ and $d$.

Since $S$ is an integer matrix, its eigenvalues are algebraic integers, and so $a^{-1}$ must be an algebraic integer. Hence all its algebraic conjugates are eigenvalues of $S$, each with multiplicity $n - d$. Therefore if $n \neq 2d$, then $a^{-1}$ must be an integer. (If $a^{-1}$ is irrational, it must be a quadratic irrational and $n = 2d$.)

Since $S$ has exactly two distinct eigenvalues its minimal polynomial is quadratic—thus we have

$$S^2 - a S - b I = 0$$

for suitable integers $a$ and $b$. We may normalize $S$ so that

$$S = \begin{pmatrix} 0 & 1^T \\ 1 & S_1 \end{pmatrix}$$

and now

$$S^2 = \begin{pmatrix} n - 1 & 1^T S_1 \\ S_1 1 & S_1^2 + J \end{pmatrix}.$$
Consequently $b = n - 1$, \quad $S_1 \mathbf{1} = a \mathbf{1}$

and $S^2 + J - aS - (n - 1)I = 0$, whence

\[ S^2 - aS - (n - 1)I = -J. \]

Using this we may show that the graph with adjacency matrix

\[ A = \frac{1}{2}(S + J - I) \]

is regular and has exactly three distinct eigenvalues. Therefore it is strongly regular.

### 10.4 Difference Sets and Equiangular Lines

**10.4.1 Lemma.** Let $G$ be an abelian group of order $n$, and let $\psi_1, \ldots, \psi_n$ be the characters of $G$. Suppose $N$ is a 01-group matrix over $G$. If $h^T$ is a row of $N$, then the number of angles between the lines spanned by the vectors $h \circ \psi_r$ is one less than the number of eigenvalues of $NN^T$.

**Proof.** If $\psi$ and $\varphi$ are characters of $G$, then

\[ \langle h \circ \psi, h \circ \varphi \rangle = \langle h, \overline{\psi} \circ \varphi \rangle \quad (10.4.1) \]

If $\chi$ is a character for $G$, then $N\chi = \lambda \chi$ for some $\lambda$ and therefore, since the entries of $\chi$ are complex numbers of norm 1,

\[ |\langle h, \chi \rangle| = |\lambda|. \]

So

\[ |\langle h \circ \psi, h \circ \varphi \rangle|^2 \]

is equal to the eigenvalue of $NN^T$ on $\overline{\psi}\varphi$.

Note that if the weight of the vector $h$ above is $d$, then the vectors $h \circ \psi$ lie a $d$-dimensional subspace of $\mathbb{C}^n$.

If $X$ is the incidence graph the design, then

\[ A(X) = \begin{pmatrix} 0 & N \\ N^T & 0 \end{pmatrix} \]

and

\[ A(X)^2 = \begin{pmatrix} NN^T & 0 \\ 0 & N^T N \end{pmatrix}. \]

It follows that number of angles is equal to the number of non-negative eigenvalues of $X$. 
If $N$ is a group matrix over $\mathbb{Z}_{n^2+n+1}$ and an incidence matrix for a projective plane of order $n$, then

$$NN^T = nI + J$$

which has eigenvalues $(n+1)^2$ and $n$ (with multiplicity 1 and $n^2+n$ respectively). Hence we obtain a set of $n^2+n+1$ equiangular lines in $\mathbb{C}^{n^2+n+1}$ whenever $n$ is a prime power. The size of this set of lines meets the relative bound, as you are invited to prove.

A complex matrix is flat if all its entries have the same absolute value. The vectors spanning the $d^2 - d + 1$ lines are flat; it can be shown that a set of flat equiangular lines in $\mathbb{C}^d$ has size at most $d^2 - d + 1$.

## 10.5 MUB’s

If $x_1, \ldots, x_d$ and $y_1, \ldots, y_d$ are two orthonormal bases in $\mathbb{C}^d$, we say that they are unbiased if there is a constant $\gamma$ such that for all $i$ and $j$,

$$\langle x_i, y_j \rangle \langle y_j, x_i \rangle = \gamma.$$

In other words, the angle between any two lines spanned by vectors in different bases is the same. A set of orthonormal bases is mutually unbiased if each pair of bases in it is unbiased. If $U$ and $V$ are $d \times d$ unitary matrices then their columns provide a pair of orthonormal bases, and these bases are unbiased if and only if the matrix $U^*V$ is flat. Note that $U^*V$ is itself unitary, and that its columns and the standard basis of $\mathbb{C}^d$ are unbiased.

The two bases

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

are mutually unbiased.

The angle between lines corresponding to vectors from distinct orthogonal bases is determined by $d$. To see this, suppose $x_1, \ldots, x_d$ and $y_1, \ldots, y_d$ are orthogonal and unbiased with $|\langle x_i, y_j \rangle|^2 = \gamma$. Then since $x_1, \ldots, x_d$ is an orthonormal basis

$$y_1 = \sum \langle x_i, y_1 \rangle x_i$$

and

$$\langle y_1, y_1 \rangle = \sum_i |\langle x_i, y_1 \rangle|^2 = d\gamma.$$

Hence $\gamma = d^{-1}$ (and $|\langle x_i, y_j \rangle| = d^{-1/2}$).

Our goal is to find mutually unbiased sets of bases with maximal size. How large can a mutually unbiased set of bases be? If $P$ and $Q$ are projections onto lines spanned by two vectors from a set of mutually unbiased bases, then $\langle P, Q \rangle$ is 0, 1 or $d^{-1}$. The Gram matrix of the projections onto lines spanned by vectors from a set of mutually unbiased bases is

$$G = I_m \otimes I_d + \frac{1}{d} (J_m - I) \otimes J_d.$$
We determine the rank of $G$ by counting its nonzero eigenvalues. The eigenvalues of $(J_m - I) \otimes J_d$ are

<table>
<thead>
<tr>
<th>eigenvalue</th>
<th>multiplicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(m-1)d$</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>$m(d-1)$</td>
</tr>
<tr>
<td>$-d$</td>
<td>$m-1$</td>
</tr>
</tbody>
</table>

Thus the eigenvalues of $I + \frac{1}{2} (J_m - I) \otimes J_d$ are

<table>
<thead>
<tr>
<th>eigenvalue</th>
<th>multiplicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>$m(d-1)$</td>
</tr>
<tr>
<td>0</td>
<td>$m-1$</td>
</tr>
</tbody>
</table>

Thus $\text{rk}(G) = 1 + md - m$ and therefore

$$1 + md - m \leq d^2,$$

from which it follows that $m \leq d + 1$.

**Note.** If we work in $\mathbb{R}^d$ we get

$$1 + md - m \leq \frac{d^2 + d}{2},$$

and then we find that $m \leq 1 + \frac{d}{2}$.

The columns of a unitary matrix form an orthonormal basis. In fact a matrix $H$ is unitary if and only if its columns form an orthonormal basis. Suppose $H$ and $K$ are unitary then the columns of $H$ and $K$ are unbiased if and only if all entries of $H^*K$ have absolute value $\frac{1}{\sqrt{d}}$. So $H^*K$ is flat and since it is a product of unitary matrices it is unitary. Note that $H$ and $K$ are unbiased if and only if $I$ and $H^*K$ are. Thus each flat unitary matrix gives a pair of unbiased bases in $\mathbb{C}^d$ (matrix, identity).

Suppose the columns of matrices $H_1, \ldots, H_m$ and $K_1, \ldots, K_m$ form mutually unbiased bases in $\mathbb{C}^d$ and $\mathbb{C}^e$ respectively. Then the Kronecker products

$$H_i \otimes K_i$$

give a set of $m$ mutually unbiased bases in $\mathbb{C}^{de}$. (This is very easily verified.) It follows that in any dimension there is a set of at least three mutually unbiased bases.

We describe a construction due to Beth and Wocjan. Let $\mathcal{A}$ be an $OA(n,k)$ and let $H$ be a flat unitary matrix of order $n \times n$. The dual incidence structure associated to $\mathcal{A}$ has $n^2$ points and $nk$ lines. The lines divide into $k$ parallel classes and each line contains exactly $n$ points. Let $N$ denote $\{1, \ldots, n\}$. The rows of $\mathcal{A}$ are indexed by ordered pairs from $N \times N$, and we assume that the row are listed in lexicographic order relative to the indexing pairs.
10.5.1 Theorem. Let $A$ be an $OA(n,k)$ and let $H$ be a flat $n \times n$ unitary matrix. Then the $n^2k$ vectors $n^{-1/2}(y\circ(1\otimes h))$, where $y$ varies over the incidence vectors of the points of $A$ and $h$ over the columns of $H$, form a set of $k$ mutually unbiased bases in $\mathbb{C}^{n^2}$.

There is an $OA(26,6)$, and so we obtain six mub’s in $\mathbb{C}^{676}$. We can find five mutually unbiased bases in $\mathbb{C}^4$ (and in $\mathbb{C}^{13^2}$), and our trivial product construction then yields five in $\mathbb{C}^{676}$.

10.6 Real MUB’s

We briefly consider the real case. This received almost no attention prior to the physicists’ work on the complex case.

We note first that a flat orthogonal matrix is a scalar multiple of a Hadamard matrix. It follows that if we have a real pair of mutually unbiased matrices in $\mathbb{R}^d$ then either $d = 2$ or $4 \mid d$.

10.6.1 Lemma. If there is a set of three mutually unbiased bases in $\mathbb{R}^d$, then $d$ is an even square.

Proof. Suppose $H$ and $K$ are $d \times d$ Hadamard matrices such that the columns of

$$I, \frac{1}{\sqrt{d}} H, \frac{1}{\sqrt{d}} K$$

are mutually unbiased. Then

$$\frac{1}{d} H^T K$$

must be a flat real orthogonal matrix and therefore

$$\frac{1}{\sqrt{d}} H^T K$$

is a Hadamard matrix. This implies that $\sqrt{d}$ must be rational.

10.6.2 Lemma. If there is a set of four mutually unbiased bases in $\mathbb{R}^d$, then $16 \mid d$.

Proof. Suppose we have four mutually unbiased bases in $\mathbb{R}^d$, the first of which is the standard basis, and assume that $d = 4s^2$. Then the last three bases come from three Hadamard matrices $H$, $K$ and $L$ such that if $x$, $y$ and $z$ respectively are columns from these three matrices, then

$$\langle x, y \rangle = \langle x, z \rangle = \langle y, z \rangle = 2s.$$

We consider the equation

$$\langle 1, (x + y) \circ (x + z) \rangle = \langle x + y, x + z \rangle.$$
Since $x$, $y$ and $z$ are $\pm 1$ vectors, the entries of $x + y$ and $x + z$ are $0$ and $\pm 2$ and therefore the left side above is divisible by $4$. On the other hand

$$\langle x + y, x + z \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle x, z \rangle + \langle y, z \rangle = 4s^2 \pm 2s \pm 2s \pm 2s$$

and therefore $s$ must be even.

10.7 Affine Planes

Let $V$ be a vector space of dimension two over $GF(q)$; we write its elements as pairs $(x, y)$. Let $[a, b]$ denote the set of points

$$\{(x, y) : y = ax + b\}.$$  

This a line in the affine plane over $GF(q)$, and as we vary $a$ and $b$ we get all lines except those parallel to the $y$-axis—the lines with infinite slope. It is easy to verify that this structure is a divisible semisymmetric design. Our problem is to show that there is an abelian group of automorphisms acting regularly on points and lines.

There are two obvious sets of automorphisms. Let $T_{u,v} : V \rightarrow V$ be given by

$$T_{u,v}(x, y) = (x + u, y + v).$$

We call the maps $T_{u,v}$ translations, they form an abelian group of order $q^2$. If $(x, y)$ is on the line $[a, b]$, then

$$y + v - (a(x + u) + b) = (y - ax - b) - (au - v)$$

and therefore $T_{u,v}(x, y)$ is on $[a, b - au + v]$. Thus we can define the image of $[a, b]$ under $T_{u,v}$ to be $[a, b - au + v]$, and with this definition $T_{u,v}$ is an automorphism of our incidence structure. We see that translations are automorphisms that each line to a parallel line. In particular you may show that the group of translations has $q$ orbits on lines.

We can also define dual translations $S_{u,v}$ by

$$S_{u,v}[a, b] = [a + u, b + v].$$

Then

$$y - (a + u)x - (b + v) = y + ux + v - ax - b$$

and so $S_{u,v}$ maps lines on $(x, y)$ to the lines on $(x, y + ux + v)$. Again we get a group of automorphisms, with $q$ orbits on points.

What we need though is an abelian group with one orbit on points and one orbit on lines. Define

$$R_{u,v} = T_{u,v}S_{u,0}.$$  

Then these $q^2$ automorphisms from an abelian group of order $q^2$ that acts transitively on point and on lines. Consequently we get a set of $q$ mutually unbiased bases in $\mathbb{C}^q$, that are all unbiased relative to the standard basis.

This construction does not make use of the fact that finite fields are associative, and we may use a commutative semifield in place of a field. All known examples of mub’s can be constructed in this way.
10.8 Products

If $H_1, \ldots, H_m$ is a set of unitary matrices describing a set of mub’s in $\mathbb{C}^d$ and $K_1, \ldots, K_m$ is a second set giving mub’s in $\mathbb{C}^e$, then the products

$$H_i \otimes K_i, \quad (i = 1, \ldots, m)$$

give us a set of mub’s in $\mathbb{C}^{de}$. This may not seem to be a very efficient construction, but in many cases it is the best we can do. If $d$ is a prime power then there is a mutually unbiased set of bases of size $d + 1$ in $\mathbb{C}^d$; hence this product construction guarantees the existence of a mutually unbiased set of three bases in any dimension. When $d \cong 2$ modulo four, there is no better bound known in general.

There is one construction, due to Beth and Wocjan, which is better than the product construction in some cases. Suppose we have an $OA(k, q)$ and a flat unitary matrix $H$ of order $q \times q$. Our array can be viewed as an incidence structure with $q^2$ lines and $kq$ points. Let $M$ be the incidence matrix of the dual; this has order $q^2 \times kq$. Then the $kd^2$ vectors

$$(1 \otimes He_i) \circ Me_j$$

form $k$ mutually unbiased bases in $\mathbb{C}^{q^2}$.

If $q = 26$, then the product construction provides five mub’s in $\mathbb{C}^{576}$. However there is an $OA(6, 26)$ and so Wocjan and Beth give us six mub’s.
Chapter 11

Association Schemes

11.1 Intersection Matrices

First an example. Let \( \mathcal{D} \) be a design with blocks of size \( k \), and define matrices \( A_0, \ldots, A_d \) with rows and columns indexed by the blocks of \( \mathcal{D} \), such that

\[
(A_i)_{\alpha,\beta} = \begin{cases} 
1, & |\alpha \cap \beta| = k - r; \\
0, & \text{otherwise.}
\end{cases}
\]

So the \( A_i \) are symmetric 01-matrices that sum to \( J \), and \( A_0 = I \). We call them the block intersection matrices of the design. Any automorphism of \( \mathcal{D} \) must commute with each of the matrices \( A_i \), and hence it must commute with any matrix in the algebra that they generate. Thus in some sense our design is most regular if the dimension of this algebra is as small as possible. Since \( A_0, \ldots, A_d \) are linearly independent, the lower bound on this dimension is one ore than the degree of \( \mathcal{D} \), and if equality holds then there are interesting consequences.

First, for each \( i \) and \( j \) the product \( A_i A_j \) belongs to the algebra, and therefore there must be scalars \( p_{i,j}(r) \) such that

\[
A_i A_j = \sum_{r=0}^{k} p_{i,j}(r) A_r.
\]

Second, from this we see that \( A_i A_j \) is symmetric, which implies that \( A_i \) and \( A_j \) commute.

An association scheme with \( d \) classes is a set \( \mathcal{A} \) of 01-matrices \( A_0, \ldots, A_d \) such that

(a) \( A_0 = I \).
(b) \( \sum_i A_i = J \).
(c) \( A_i^T \in \mathcal{A} \) for each \( i \).
(d) There are scalars \( p_{i,j}(r) \) such that \( A_i A_j = \sum r p_{i,j}(r) A_r \).

(e) \( A_i A_j = A_j A_i \) for all \( i \) and \( j \).

If \( A_i = A_i^T \), we say that the scheme is symmetric. This is the only case we will consider (and then (e) is redundant). The matrices \( A_1, \ldots, A_d \) can be viewed as adjacency matrices of graphs \( X_1, \ldots, X_d \). We will say that these graphs form an association scheme—this means an association scheme is a set of matrices or a set of graphs, whichever suits us. We use \( \mathbb{R}[A] \) to denote the vector space spanned by the matrices in the scheme; this is known as its Bose-Mesner algebra. Since \( J \in \mathbb{R}[A] \) and since \( \mathbb{R}[A] \) is commutative, each matrix \( A_i \) commutes with \( J \) and therefore each graph in the scheme is regular.

11.1.1 Theorem. Let \( D \) be a design with degree \( s \) and strength \( t \). If \( t \geq 2s - 2 \), then the block intersection matrices of \( D \) from an association scheme with \( s \) classes.

A design is said to be schematic if its intersection matrices form an association scheme. It is an exception for a design to be schematic. Any symmetric design is trivially schematic. A 2-design with degree two is schematic, and so each 2-(\( v, k, 1 \)) gives us an association scheme with two classes. A graph \( X \) and its complement form an association scheme with two classes if and only if \( X \) (and \( X \)) are strongly regular.

Using Hamming distance we can define “intersection matrices” for an orthogonal array. The analog of Theorem 11.1.1 holds for orthogonal arrays as well. If an orthogonal array has degree \( s \) and strength \( t \) and \( t \geq 2s - 2 \), then its intersection matrices form an association scheme with \( s \) classes. Accordingly each orthogonal array with index one and strength two gives rise to a pair of strongly regular graphs.

A finite set \( S \) of points on the unit sphere in \( \mathbb{R}^d \) has strength at least \( t \) if the average over the points in the set of a polynomial of degree at most \( t \) is equal to its average over the entire sphere. Then \( S \) is a 2-design if and only if

\[
\sum_{x \in S} xx^T = \frac{|S|}{d} I.
\]

The degree of \( S \) is the size of the set of inner products \( x^T y \), for distinct points \( x \) and \( y \). If \( S \) has degree \( s \) and strength \( t \) and \( t \geq 2s - 2 \), we get an association scheme with \( s \) classes.
Chapter 12

Incidence Structures

12.1 Definitions

An incidence structure \((P, B)\) consists of a set of points \(P\), a set of blocks \(B\) and an incidence relation on \(P \times B\). Thus a point and a block are either incident or not; in the first case we may say that the point lies in the block, or that the block lies on the point. What we are calling blocks may sometimes have another name, for example, in geometry the blocks are usually called lines. The point and block sets are disjoint. It is quite common to find that the blocks are defined to be subsets of the point set, but this is not a requirement.

Any graph determines an incidence structure where the vertices are the points and the edges are the blocks. A point is incident with the edges that contain it. A planar graph determines three incidence structures. The first is the one just described. The second has the vertices as points and the faces as its blocks. The third has the blocks as points and the edges as blocks.

If \(P\) and \(B\) are the point and block sets of an incidence structure, then we may view \(B\) as the point set and \(P\) as the block set of a second incidence structure. This is called the dual of the first incidence structure. The dual of the dual is the original structure.

If \((P, B)\) is an incidence structure, its incidence graph is the graph with vertex set \(P \cup B\), where a pair of vertices \(u\) and \(v\) are adjacent if one is a point and the other is a block incident with it. The incidence graph is is bipartite, with bipartition \((P, B)\). In fact we will view the incidence graph as bicolored: in addition to the graph itself we are provided with the ordered pair \((P, B)\) which specifies a 2-coloring of the bipartite graph. For example, the incidence graphs of an incidence structure and its dual have the same underlying bipartite graph, but they have different 2-colourings. The incidence graph provides a very useful tool for working with incidence structures.

An incidence structure \((P, B)\) is regular if each point is incident with the same number of blocks; it is uniform if each block is incident with the same number of points. The incidence structure formed by the vertices and edges of
a loopless graph is uniform—each edge is incident with exactly two vertices—but it is regular if and only if the underlying graph is regular. An incidence structure is thick if the minimum valency of its incidence graph is at least three.

We will say that an incidence structure is connected if its incidence graph is connected. A connected bipartite graph has a unique 2-coloring, and so a connected incidence structure is determined up to duality by its incidence graph.

An incidence structure is a partial linear space if each pair of points lies in at most one block. It is a linear space if each pair of points lies on exactly one line. A dual linear space is an incidence structure whose dual is a linear space. (This suggests, correctly, that the dual of a linear space need not be a linear space.)

12.1.1 Lemma. An incidence structure is a partial linear space if and only if its incidence graph has girth at least six.

Proof. Suppose $a$ and $b$ are points and $C$ and $D$ are blocks. Then the vertices $a, C, b, D$ form a 4-cycle in the incidence graph if and only if $a$ and $b$ are both incident with $C$ and $D$.

One corollary of this is that if an incidence structure is a partial linear space, then so is its dual. (As we noted above, the dual of a linear space might not be a linear space.)

Since the incidence graph of a partial linear space does not contain a copy of $K_{2,2}$ it cannot contain a copy of $K_{2,m}$ where $m \geq 2$. It follows that we cannot have two blocks incident with exactly the same set of points, and in this case it is natural to identify a block with the subset of $\mathcal{P}$ consisting of the points incident with it.

Suppose $(\mathcal{P}, \mathcal{B})$ is an incidence structure. If $a, b \in \mathcal{P}$, we may define the line through $a$ and $b$ to be the intersection of all blocks incident with both $a$ and $b$. The incidence structure formed from the points and lines just constructed is a partial linear space; it is usually interesting only if it is thick.

A subset $S$ of the points of a partial linear space is a subspace if any line that contains two points of $S$ is a subset of $S$.

Besides the incidence graph, there are two further graphs associated with a partial linear space. The vertices of the point graph are the points of the incidence structure, and two points are adjacent if and only if they are distinct and collinear, that is, there is a line that contains them both. The line graph has the lines as vertices, and two lines are adjacent if and only if the are distinct and a point in common. (Thus the line graph is the point graph of the dual of the partial linear space.)

12.2 Designs

We say that an incidence structure $(V, \mathcal{B})$ has strength at least $t$ if for $s = 1, \ldots, t$ there are constants $\lambda_1, \ldots, \lambda_t$ such that each $s$-subset of $V$ is incident with
exactly $\lambda_s$ elements of $B$. We take the view that $\lambda_0 = b$. A $t$-design is a uniform incidence structure with strength at least $t$. The usual convention is the “design” by itself means “2-design”. Note that the same design can, for example, be both a 2-design and 3-design. However the parameters will be different.

We have the following parameter relations

$$\binom{v}{s} \lambda_s = b \binom{k}{s}$$

or, equivalently

$$\binom{v}{r} = \binom{\lambda_0}{\lambda_s},$$

One consequence of this is that if our design has strength $t$ and $r, s \leq t$, then the ratio $\lambda_s/\lambda_t$ is determined by $v$ and $k$—thus it does not depend on the structure of the design.

A 2-design with $\lambda = 1$ is a partial linear space. A Steiner system is a $t$-design with $\lambda_t = 1$.

A design is simple if no two blocks are incident with the same set of points. Generally our designs will be simple, and you may assume that they are simple in the absence of any warning.

### 12.3 Matrices and Maps

The incidence matrix of an incidence structure $(P, B)$ has rows indexed by $P$, columns indexed by $B$ and its $ij$-th entry is 0 or 1 according as the $i$-th point is incident with the $j$-th block, or not. We see that incidence matrix depends on an ordering of the points and blocks, but this will not cause any problems. If the incidence matrix of an incidence structure is $B$, then the adjacency matrix of its incidence graph is

$$\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}.$$
permutations of \( \mathcal{P} \) and \( \mathcal{B} \) respectively, then \((\pi, \beta)\) is an automorphism if \(a^\pi\) is incident with \(B^\beta\) if and only if \(a\) is incident with \(B\).

Now suppose our incidence structure has \(v\) points and \(b\) blocks and let \(e_1, \ldots, e_v\) and \(f_1, \ldots, f_b\) respectively denote the standard bases of \(\mathbb{R}^v\) and \(\mathbb{R}^b\). If \(\pi\) is a permutation of the points, then there is a unique linear map of \(\mathbb{R}^v\) to itself that sends \(e_i\) to \(e_{i\pi}\) and the matrix that represents this linear map is a permutation matrix. The pair of permutation matrices \((P, Q)\) determines an automorphism of our incidence structure if and only if

\[PBQ^T = B.\]

(Taking the transpose of \(Q\) here is arbitrary, but useful.) If no two blocks are incident with the same set of points, then the block permutation of an automorphism is determined by the point permutation. In this case, a permutation \(\pi\) of \(V\) will be an automorphism if, when \(x \in V\) and \(\alpha \in B\), we have \(x\) is incident with \(\beta\) if and only if \(x^\pi\) is incident with \(\alpha^\pi\). (Here we get \(\alpha^\pi\) by applying \(\pi\) to each element of \(\alpha\).)
13.1 The Kronecker Product

If $A$ and $B$ are matrices over the same ring, we define their Kronecker product $A \otimes B$ to be the matrix we get by replacing each entry $A_{i,j}$ of $A$ with the matrix $A_{i,j}B$. Note that neither $A$ nor $B$ need be square. For example, if $x \in \mathbb{F}^m$ and $y \in \mathbb{F}^n$, then

$$x \otimes y^T = xy^T.$$  

If $A$ is an $m \times n$ matrix, then $\text{vec}(A)$ is the $mn \times 1$ matrix we get by stacking the columns of $A$ one on top of the other. So if $e_1, \ldots, e_n$ is the standard basis, then

$$\text{vec}(A) = \begin{pmatrix} Ae_1 \\ \vdots \\ Ae_n \end{pmatrix}.$$  

The Kronecker product is bilinear, i.e., it is linear in each variable. We also have

$$(A \otimes B)^T = A^T \otimes B^T.$$  

The following properties are fundamental.

13.1.1 Theorem. If $A$, $X$ and $B$ are matrices such that the product $AXB^T$ is defined, then

$$(I \otimes A)(X) = \text{vec}(AX), \quad (B \otimes I) \text{vec}(X) = (XB)^T.$$  

One consequence of this is that

$$(A \times B) = (I \otimes A)(B \otimes I) = (B \otimes I)(I \otimes A).$$  

It also follows that if $AC$ and $BD$ are defined, then

$$(A \otimes B)(C \times D) = AC \otimes BD.$$  

83
In particular if $Ax = \lambda x$ and $By = \mu y$, then

$$(A \otimes B)(x \otimes y) = \lambda x \otimes \mu y = \lambda \mu x \otimes y.$$ 

We have

$$(e_i \otimes g_k)^T (A \otimes B)(f_j \otimes h_\ell) = e_i^T A f_j g_k^T B h_\ell.$$ 

Because of this we can view the rows of $A \times B$ as being indexed by ordered pairs $(i, k)$, and the columns by ordered pairs $(j, \ell)$. Then

$$(A \otimes B)_{(i,k),(j,\ell)} = A_{i,j} B_{k,\ell}$$ 

Suppose $U$ and $V$ are vector spaces. The vector space spanned by the vectors

$$u \otimes v, \quad u \in U, \quad v \in V$$

is called the tensor product of $U$ and $V$, and is denoted by $U \otimes V$. (I will become upset if refer to this as the Kronecker product of $U$ and $V$.)

Suppose $P$ is the linear mapping on $V \otimes V$ defined by the requirement that

$$P(x \otimes y) = y \otimes x.$$ 

Prove that $P^2 = I$, that $P$ commutes with $A \otimes A^T$ and that $P(A \otimes A^T)$ is symmetric.

If $A$ and $B$ are $m \times n$ matrices, we define their Schur product $A \circ B$ by

$$(A \circ B)_{i,j} := A_{i,j} B_{i,j}$$

(It is sometimes called the bad-student’s product.) Show that $A \circ B$ is a principal submatrix of $A \otimes B$.

### 13.2 Normal Matrices

A matrix $M$ over $\mathbb{C}$ is normal if it commutes with its conjugate-transpose. Examples are Hermitian matrices and unitary matrices. If $A = L^* DL$ where $L$ is unitary and $D$ is diagonal, then $A^* = L^* \overline{D} L$; hence $A$ is normal. So a matrix that unitarily diagonalizable is normal.

The converse is true:

#### 13.2.1 Theorem. A matrix $M$ is unitarily diagonalizable if and only if it is normal.

#### 13.2.2 Theorem. Suppose $\mathcal{A}$ is a commutative algebra of $v \times v$ complex matrices. If $\mathcal{A}$ is closed under complex-conjugate, then there is a basis for $\mathbb{C}^v$ that consists of common eigenvectors for (the matrices in) $\mathcal{A}$.
13.3. POSITIVE SEMIDEFINITE MATRICES

This theorem fails for the commutative algebra consisting of the matrices of the form
\[
\begin{pmatrix}
a & b \\
0 & a
\end{pmatrix}, \quad a, b \in \mathbb{R}.
\]

A square matrix \( N \) is normal if and only if for all vectors \( z \) we have
\[
\langle Nz, Nz \rangle = \langle N^*z, N^*z \rangle. \tag{13.2.1}
\]

It is easy to verify that this holds if \( N \) is normal:
\[
\langle Nz, Nz \rangle = z^*N^*Nz = z^*NN^*z = \langle N^*z, N^*z \rangle.
\]

For the converse, show that (13.2.1) holds if and only if
\[
\langle Nz, Nw \rangle = \langle N^*z, N^*w \rangle
\]
for all \( z \) and \( w \), and then show that \( z^*(N^*N - NN^*)w = 0 \) for all \( z \) and \( w \) if and only if \( N^*N = NN^* \).

13.3 Positive Semidefinite Matrices

A complex matrix \( M \) is positive semidefinite if \( M = M^* \) and \( z^*Mz \geq 0 \) for all \( z \). If \( M \) is positive semidefinite and \( z^*Mz = 0 \) implies that \( z = 0 \), it is positive definite.

13.3.1 Theorem. If \( M \) is a Hermitian matrix, the following assertions are equivalent:

(a) \( M \) is positive semidefinite.

(b) \( M = N^*N \) for some matrix \( N \).

(c) The eigenvalues of \( M \) are non-negative. \( \square \)

A positive semidefinite matrix is positive definite if and only if it is invertible. The sum of two positive semidefinite matrices is positive semidefinite, and the sum of a positive definite and a positive semidefinite matrix is positive definite, hence invertible. Thus if \( r > \lambda \) and \( NN^T = (r-\lambda)I + \lambda J \), then \( NN^T \) is invertible because \( (r-\lambda)I \) is positive definite and \( \lambda J = \lambda 11^T \) is positive semidefinite.

13.3.2 Lemma. A Hermitian matrix \( M \) is positive semidefinite if and only \( \text{tr}(MX) \geq 0 \) for all positive semidefinite matrices \( X \). \( \square \)

(The set of \( n \times n \) positive semidefinite matrices is a convex cone; this lemma implies that this cone is self dual. Talk to Levent Tuncel.)

Prove that a principal submatrix of a positive semidefinite matrix is positive semidefinite. This implies that the diagonal entries of a positive semidefinite matrix are non-negative. Prove that if \( M \) is positive semidefinite and \( M_{r,r} = 0 \), then all entries in the \( r \)-th row and all entries in the \( r \)-th column of \( M \) are zero.

Prove that if \( M \) and \( N \) are positive semidefinite, so is \( M \otimes N \). Deduce that if \( M \) and \( N \) have the same order then \( M \circ N \) is also positive semidefinite. (This is an important result due to Schur.)
Chapter 14

Finite Fields

14.1 Arithmetic

We now start to develop the theory of finite fields. Let $F$ be a field. The identity element of $F$ generates an additive subgroup of $F$. If this subgroup is infinite, we say that $F$ has characteristic zero. If $F$ has characteristic zero, the subgroup generated by 1 is isomorphic to $\mathbb{Z}$; it follows that $F$ contains a subfield isomorphic to $\mathbb{Q}$. This is the prime subfield of $F$—the minimal subfield that contains 1. Both $\mathbb{R}$ and $\mathbb{C}$ are fields of characteristic zero.

Alternatively, the subgroup generated by 1 is finite. In this case the characteristic of $F$ is defined to be the order of this subgroup. If $p$ is prime and $F = \mathbb{Z}_p$, then the characteristic of $F$ is $p$. The field of rational functions over $\mathbb{Z}_p$ is infinite, but its characteristic is $p$.

**14.1.1 Theorem.** Let $F$ be a field. If the characteristic of $F$ is not zero, it is a prime number.

**Proof.** Suppose the characteristic of $F$ is $n$ and $n = ab$. Then $n \cdot 1 = 0$ and therefore

$$(a \cdot 1)(b \cdot 1) = n \cdot 1 = 0.$$ 

Since $F$ is a field, this implies that either $a \cdot 1 = 0$ or $b \cdot 1 = 1$. If $a \cdot 1 = 0$ then the order of the subgroup generated by 1 divides $a$. Therefore $a = n$ and $b = 1$. We conclude that $n$ must be a prime number.

If the characteristic of $F$ is positive, it follows that the prime field of $F$ is isomorphic to $\mathbb{Z}_p$, for some prime $p$. Hence $F$ is a vector space over $\mathbb{Z}_p$ for some prime $p$, and consequently

$$|F| = p^n$$

for some integer $n$.

**14.1.2 Corollary.** The order of a finite field is the power of a prime.  

87
We turn from addition to multiplication. The non-zero elements of a field form an abelian group under multiplication, when the field is finite we have a much stronger assertion.

14.1.3 Theorem. If \( F \) is finite, its non-zero elements form a cyclic group under multiplication.

Proof. Let \( A \) denote the group formed by the non-zero elements of \( F \). Let \( m \) denote the exponent of \( A \), that is, the least integer \( m \) such that \( a^m = 1 \) for all \( a \) in \( A \). Then every element of \( A \) is root of the polynomial \( t^m - 1 \). Since \( t^m - 1 \) has at most \( m \) roots, \( |A| \leq m \). On the other hand, by Lemma ??, we see that \( A \) contains an element of order \( m \), and therefore it is cyclic. \( \square \)

We use \( F^* \) to denote the set of non-zero elements of \( F \). A generator of the cyclic group \( F^* \) is usually called a primitive element of \( F \).

14.1.4 Corollary. If \( F \) is a finite field of order \( q \) and \( a \in F \), the minimal polynomial of \( a \) divides \( t^q - 1 - 1 \).

Proof. The non-identity elements of \( F \) form a cyclic group of order \( q - 1 \), and therefore \( a^{q-1} - 1 = 0 \). So the minimal polynomial of \( a \) divides \( t^q - 1 - 1 \). \( \square \)

14.2 Automorphisms

We study the automorphism of finite fields. Let \( F \) be a field. A map \( \sigma : F \to F \) is an automorphism if it is a bijection and, for all \( a \) and \( b \) in \( F \),

\[(a + b)\sigma = a\sigma + b\sigma, \quad (ab)\sigma = a\sigma b\sigma.\]

The most familiar example is the operation of complex conjugation on the complex numbers. We describe a second example. Let \( F \) be a field with characteristic \( p \). If \( i \) is an integer and \( 0 < i < p \) then the binomial coefficient

\[\binom{p}{i} = 0\]

modulo \( p \). Hence if \( x \) and \( y \) belong to \( F \), then

\[(x + y)^p = x^p + y^p.\]

Since \((xy)^p = x^py^p\), it follows that the \( p \)-th power map \( x \mapsto x^p \) is an automorphism of \( F \). It is known as the Frobenius automorphism of \( F \).

If \( \gamma \) is an automorphism of the field \( F \), then \( \text{fix}(\gamma) \) is the subset

\[\{a \in F : a^\gamma = a\} \]

We say that \( \text{fix}(\gamma) \) is the set of elements of \( F \) fixed by \( \gamma \). If \( \Gamma \) is a group of automorphisms of \( F \), then \( \text{fix}(\Gamma) \) denotes the set of elements of \( F \) fixed by each element of \( \Gamma \). Hence

\[\text{fix}(\Gamma) = \bigcap_{\gamma \in \Gamma} \text{fix}(\gamma).\]
It is easy to verify that $\text{fix}(\gamma)$ is a subfield of $\mathbb{F}$, and therefore $\text{fix}(\Gamma)$ is a subfield too. For example, if $\gamma$ is complex conjugation on $\mathbb{C}$, then $\text{fix}(\gamma) = \mathbb{R}$.

**14.2.1 Lemma.** Let $\mathbb{F}$ be a field of characteristic $p$, and let $\tau$ be the Frobenius automorphism of $\mathbb{F}$. Then $\text{fix}(\gamma)$ is the prime subfield of $\mathbb{F}$.

**Proof.** We have $a^\tau = a$ if and only if $a^p - a = 0$. Therefore $\text{fix}(\tau)$ is a subfield of $\mathbb{F}$ consisting of the roots of $t^p - t$, and therefore it is the prime subfield.

If $\gamma \in \text{Aut}(\mathbb{F})$, then for any $a$ in $\mathbb{F}$

$$a^\gamma = (1a)^\gamma = 1^* a^\gamma$$

and so $1^\gamma = 1$. Now

$$(1 + 1)^\gamma = 1^\gamma + 1^\gamma = 1 + 1,$$

and a very simple induction argument yields that each element of the prime subfield is fixed by $\gamma$.

**14.2.2 Theorem.** Let $\mathbb{F}$ be a finite field of characteristic $p$ and order $p^n$ and let $\tau$ be the Frobenius automorphism of $\mathbb{F}$. Then $\mathbb{F}$ has a subfield of order $p^k$ if and only if $k \mid n$. If $k \mid n$, then there is a unique subfield of order $p^k$; it is the fixed field of $\tau^k$.

**Proof.** Let $K$ be a subfield of $\mathbb{F}$ with order $q = p^k$. Then $\mathbb{F}$ is a vector space over $\mathbb{K}$, and therefore there is an integer $d$ such that

$$p^n = |\mathbb{F}| = |\mathbb{K}|^d = p^{kd}.$$ 

This shows that $k \mid n$. Each element of $\mathbb{K}$ is a root of $t^{p^k} - t$, since this polynomial has at most $p^k$ roots in $\mathbb{F}$, there is at most one field of order $p^k$.

If $a \in \mathbb{F}$ then $a^{\tau^k} = a$ if and only if

$$a^{p^k} - a = 0.$$ 

Accordingly $\text{fix}(\tau^k)$ consists of roots of

$$t^{p^k} - t.$$ 

If $k \mid n$, then $t^{p^k} - t$ divides $t^n - t$ and therefore it has exactly $p^k$ roots in $\mathbb{F}$. Thus $\text{fix}(\tau^k)$ is a subfield of order $p^k$.$\square$

**14.2.3 Corollary.** If $\mathbb{F}$ is a finite field of characteristic $p$ and order $p^d$, then the Frobenius automorphism of $\mathbb{F}$ has order $d$. $\square$

**14.2.4 Theorem.** Let $q = p^d$, where $p$ is prime. There is a unique field of order $q$, which is the splitting field for the polynomial $t^d - t$. $\square$
Proof. Let $F$ denote the splitting field for $t^q - t$ over the field $\mathbb{Z}_p$. Then the $q$ roots of $t^q - t$ in $F$ form the set of elements of $F$ fixed by the $q$-th power map $a \mapsto a^q$, and therefore they form a subfield of $F$. Since this subfield contains all roots of $t^q - q$, it is the splitting field for $t^q - t$ and therefore this subfield equals $F$. This shows that a field of order $q$ exists.

Suppose $E$ is a field of order $q$. Since $F^*$ is cyclic, the elements of $F$ give $q$ distinct roots for $t^q - t$. Since no subfield of $E$ can contain all these roots, $E$ is a splitting field for $t^q - t$. As all splitting fields for a polynomial are isomorphic, this shows that there is a unique field of order $q$ (up to isomorphism).

Let $q$ be a prime power and let $E$ be an extension of degree $d$ of a field $F$ with order $q$. Let $a$ be a primitive element in $E$ and let $\psi$ be its minimal polynomial over $F$. Then $\psi$ is irreducible over $F$ and

$$E = F(a) \cong F[t]/(\psi).$$

This implies that $\deg(\psi) = n$. By the theorem, for each positive integer $d$ there is a finite field $E$ of order $q^d$, and this field has a subfield of order $q$. We conclude that for each positive integer $d$, there is an irreducible polynomial in $F[t]$ with degree $d$.

### 14.3 Squares

We describe the basic results concerning squares in finite fields.

An element in a field $F$ that can be written as $a^2$ for some $a$ is, naturally, called a square in $F$. We determine the squares in finite fields.

First let $F$ be a finite field of characteristic two. Then $F^*$ is a cyclic group of odd order. Every element in a group of order $k$ is square. For suppose $x$ is a group element with odd order $k$. Then $k + 1$ is even and so

$$x = (x^{(k+1)/2})^2,$$

therefore $x$ is square.

Now consider a field $F$ of order $q$, where $q$ is odd. Then $F^*$ is a cyclic group of even order. Let $a$ be a primitive element of $F$, that is, a generator of $F^*$. The non-zero squares in $F$ are precisely the even powers of $a$, and these form a subgroup of $F^*$ of order $(q - 1)/2$. Thus exactly half the non-zero elements of $F$ are squares.

Denote the set of non-zero squares in $F$ by $S$. Since the index of $S$ in $F^*$ is two, the quotient $F^*/S$ is isomorphic to the subgroup of $\mathbb{Z}$ formed by the set $\{1, -1\}$. Therefore map from $F^*$ to $\mathbb{Z}$ that assigns 1 to each square and $-1$ to each non-square is a homomorphism. It follows that the product of a non-zero square with a non-square is not a square, and the product of two non-squares is a square.

As $a$ has order $q - 1$, we see that $a^{(q-1)/2}$ is a root of $t^2 - 1$ and is not equal to 1. Hence $a^{(q-1)/2} = -1$ and from this we deduce that $-1$ is a square if and
Only if \((q - 1)2\) is even. In other words, \(-1\) is a square if and only if \(q \equiv 1 \mod 4\).

Finally we prove that each element of \(\mathbb{F}^*\) is the sum of two squares. Consider the set \(S + S\). Since the order of \(S\) does not divide \(q\), we see that \(S\) is not an additive subgroup of \(\mathbb{F}\) and therefore there is an element \(b\) in \((S + S) \setminus S\). Since \(b\) is not a square, the multiplicative coset \(bS\) is the set of non-zero non-squares in \(\mathbb{F}^*\) and, as \(b\) is the sum of two squares, every element in \(bS\) is the sum of two squares.
## Index

- **G-matrix**, 61
- **s-arc**, 27
- **t-design**, 83
- **2-design**, 1

 affine resolvable design, 42
 amply regular, 74
 angle, 67
 association scheme, 79
 automorphism, 83
 automorphism group, 83

 ball of radius $e$, 56
 bicolored, 81
 bilinear form, 8
 biplane, 39
 block design, 1
 block graph, 39
 block intersection matrices, 79
 blocks, 1
 Bose-Mesner algebra, 80

 central, 62
 character, 82
 characteristic zero, 89
 circle, 30
 circulant, 61
 class graph, 45
 code, 55
 commutative semifield, 77
 complete design, 2
 complex Hadamard matrix, 25
 congruent, 9
 connected, 82
 convex cone, 64
 convolution, 62
 core, 24
 covering radius, 56
 cyclic, 55
 degenerate projective plane, 27
 degree, 45
 difference set, 2, 61
 distance, 44
 doubly even, 57
 dual, 81
 dual code, 54
 dual group, 63
 dual linear space, 82
 dual translations, 77

equiangular, 68
equivalent, 9, 33
even, 57
extension, 48
fibres, 75
flag, 2
flat, 71
flip operator, 20
Frobenius automorphism, 90

generalized Hadamard matrices, 25
graph of an array, 34
group matrix, 61

Hadamard design, 19
Hamming distance, 55
Higman-Sims graph, 41
homomorphism, 83
hyperplanes, 7

incidence graph, 81
incidence matrix, 83
incidence structure, 81

92
INDEX

isomorphic quadratic spaces, 10
isomorphism, 83

Kronecker product, 19, 85

Latin square of order $n$, 33
Legendre function, 24
length, 55
line graph, 82
line through $a$ and $b$, 82
linear space, 27, 82

Menon, 22
monomial, 17
monomial matrix, 17
monomially equivalent, 17
multiple of a symmetric design, 39
multiplier, 63
mutually unbiased, 71

non-degenerate, 8
non-singular, 11
non-trivial, 2
normal, 86
normalized, 17

order, 12, 17, 28
orthogonal array, 33

packing radius, 56
parallel, 20
parallel class, 42
partial geometry, 36
partial linear space, 27, 82
perfect, 56
point divisible, 74
point graph, 82
positive definite, 87
positive semidefinite, 87
primitive, 40
primitive element, 90
projective plane, 8, 27
proper, 36

quadratic character, 24
quadratic form, 8
quadratic space, 10

quasi-residual, 39
quasi-symmetric, 39

radical, 10
regular, 74, 81
residual, 29
residual design, 39
resolvable, 42

schematic, 80
Schur product, 20, 86
Seidel matrix, 69
self-dual, 55
self-orthogonal, 55
semisymmetric design, 74
simple, 83
square, 92
Steiner system, 47, 83
Steiner triple systems, 3
strength, 82
strongly regular, 37
strongly resolvable designs, 39
subspace, 82
sum, 10
symmetric, 8, 80
symmetric design, 2, 8

tensor product, 20, 86
thick, 28, 82
tight, 49
translate, 2
translations, 77
transversal design, 76
triangle-free, 40
trivial, 2, 37
type-II, 24

unbiased, 71
uniform, 81

weight, 55
width, 44
Witt cancellation, 11